Review of Economics of the Household 2, 413–441, 2004


# Get High and Get Stupid: The Effect of Alcohol and Marijuana Use on Teen Sexual Behavior*

MICHAEL GROSSMAN mgrossman@gc.cuny.edu
*City University of New York Graduate Center and National Bureau of Economic Research, 365 Fifth Avenue, New York, NY 10016, USA*

ROBERT KAESTNER† kaestner@uic.edu
*University of Illinois at Chicago and National Bureau of Economic Research, Institute for Government and Public Affairs, 815 West Van Buren Street, Suite 525, Chicago, IL 60607, USA*

SARA MARKOWITZ markow@newark.rutgers.edu
*Rutgers University at Newark and National Bureau of Economic Research, Department of Economics, Rutgers University, Newark, NJ 07102, USA*

**Abstract.** Numerous studies have documented a strong correlation between substance use and teen sexual behavior, and this empirical relationship has given rise to a widespread belief that substance use causes teens to engage in risky sex. This causal link is often used by advocates to justify policies targeted at reducing substance use. Here, we argue that previous research has not produced sufficient evidence to substantiate a causal relationship between substance use and teen sexual behavior. Accordingly, we attempt to estimate causal effects using two complementary research approaches. Our findings suggest that

†To whom correspondence should be addressed.
*Research for this paper was supported by grant number 5 R01 DA12692 from the National Institute on Drug Abuse to the National Bureau of Economic Research. We are indebted to Inas Rashad and Nasreen Khan for research assistance. We wish to thank Donald Kenkel, Jody Sindelar, David Salkever, David Bishai, Eric Slade, and two anonymous referees for helpful comments and suggestions. Our research is based in part on the Add Health project, a program project designed by J. Richard Udry (PI) and Peter Bearman, and funded by grant PO1-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation by the National Cancer Institute; the National Institute of Alcohol Abuse and Alcoholism; the National Institute on Deafness and Other Communication Disorders; the National Institute on Drug Abuse; the National Institute of General Medical Sciences; the National Institute of Mental Health; the National Institute of Nursing Research; the Office of AIDS Research, NIH; The Office of Behavior and Social Science Research, NIH; the Office of the Director, NIH; the Office of Research on Women's Health, NIH; the Office of Population Affairs, DHHS; the National Center for Health Statistics, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Center for Disease Control and Prevention, DHHS; the Office of Minority Health, Office of Public Health and Science, DHHS; the Office of the Assistant Secretary for Planning and Evaluation, DHHS; and the National Science Foundation. Persons interested in obtaining data files from The National Longitudinal Study of Adolescent Health should contact Add Health Project, Carolina Population Center, 123 West Franklin Street, Chapel Hill, NC 27516-3997 (email: addhealth@unc.edu). This paper has not undergone the review accorded to official NBER publications; in particular, it has not been submitted for approval by the Board of Directors. Any opinions expressed are those of the authors and not those of NIDA or NBER.

substance use is not causally related to teen sexual behavior, although we cannot definitively rule out that possibility.

Received March 24, 2004; Accepted September 24, 2004

**Keywords:** sexual behavior, substance use

**JEL Classification:** I10, I11

## 1. Introduction

Recent years have witnessed a widespread public policy concern with the sexual behavior of teenagers. Interest in this issue has been brought on mainly by the problems of teen pregnancy and out-of-wedlock birth. Teen rates of pregnancy and out-of-wedlock birth in the U.S. are high by historical standards and high relative to other developed countries. Between 1940 and 1994, the rate of out-of-wedlock birth among teenagers aged 15–19 increased from 7 per 1000 to 46 per 1000, and even the much heralded decline since 1994 leaves the rate of out-of-wedlock birth at about where it was in 1990 (Stephanie J. Ventura and Christine A. Bachrach, 2000). Internationally, rates of teen pregnancy and out-of-wedlock birth in the U.S. are nearly twice that of Britain and Canada (Jacqueline E. Darroch et al., 2001).

While some may argue about the moral aspects, most of the policy concern is over the consequences of teen pregnancy and out-of-wedlock birth. It is a widely held belief that an out-of-wedlock birth curtails educational opportunities and reduces socioeconomic attainment in adulthood, although the evidence on this point is not uniform.[1] In addition, the sexual activity that is an antecedent to pregnancy is by definition risky since it increases the chances of contracting sexually transmitted diseases (STDs) including HIV/AIDS. In the U. S., rates of STDs among adolescents are high relative to adults: among women, those aged 15–19 had the highest rates of gonorrhea, chlamydia and human papillomavirus (HPV); similarly high (relative to adults) rates of STDs are found among young men (CDC, 2002). In addition, STD rates of teenagers in the U.S. are an order of magnitude higher than STD rates of teens in other developed countries; for example, the gonorrhea rate among teens in the U.S. is 10 times higher than it is in Canada and England (Darroch et al., 2001).

What factors account for the high rates of pregnancy, out-of-wedlock birth and STDs among teenagers in the U.S.? One of the most important factors is their sexual behavior. Teenagers are sexually active: approximately half of all high school students have had sexual intercourse at some time in their life; 36% of them have had sexual intercourse in the last three months; and among those sexually active, approximately half have had two or more sexual partners in the last year (MMWR, 2000; Darroch et al., 2001). Teenagers also have relatively low rates of contraceptive use: among sexually active students only 58% report the use of a condom during last sexual intercourse (MMWR, 2000). In addition, Darroch et al. (2001) report that 20% of sexually active teenage girls report using no contraceptive method during last

intercourse. These and other statistics have led the Department of Health and Human Services (2000) to emphasize reductions in risky sexual behavior by teenagers as an important mechanism to achieve the year 2010 goals for unintended pregnancy (out-of-wedlock birth) and STDs.

To summarize, statistics on pregnancy, out-of-wedlock birth, and STDs illustrates the magnitude of the public health problem associated with the sexual behavior of teenagers. The health and development of teenagers are being adversely affected by their sexual behavior. One factor that may be contributing to this problem is substance use, which is strongly correlated with sexual activity and contraceptive use. For example, in 1999, 25% of sexually active students had used alcohol or drugs at last sexual intercourse (MMWR, 2000). Importantly, however, previous research has not established whether or not the association between substance use and sexual behavior is causal and the absence of such information undermines the formulation of effective public policies. In this paper, we investigate whether or not alcohol and marijuana are causal factors affecting sexual activity and risky sexual behavior.

To accomplish this goal we rely on two different statistical methodologies. The first is a regression model based on maximum likelihood estimation and the assumption that conditional on other measured factors, substance use and sexual activity are random variables with a bivariate normal joint probability distribution. This procedure, commonly referred to as a bivariate probit model, provides a way to control for the effect of unmeasured characteristics that may confound the causal relationship between substance use and sexual activity. The second statistical approach is an individual fixed-effects regression model estimated by ordinary least squares. This approach uses a pre- and post-test research design. The key aspect of this procedure is the use of longitudinal data that allows us to control for hard to measure time-invariant, personal characteristics that may confound estimates of the relationship between substance use and sexual behavior of teenagers.

## 2. Substance use and risky sexual behavior

A variety of analytical frameworks designed to explain risky and deviant behavior by adolescents has been developed by psychologists and sociologists.[2] Within the broad perspectives of the above theories, there are compelling a priori reasons to expect effects of illegal drugs and alcohol on risky sexual behavior (for example, Barbara C. Leigh and Ron Stall, 1993; Edward O. Laumann et al., 1994; Thomas R. Eng and William T. Butler, 1997). These substances may be employed as a way to ''break the ice'' with a new partner or to heighten the degree of sexual pleasure. At the same time, their use may interfere with judgment and decision making, leading to more unplanned sexual activity with a larger number of partners, improper application of a given method of birth control, or failure to use any type of contraceptive device. Substance abuse may also affect sexual behavior through its impact on partner selection and the level of communication within the relationship. Laurie Schwab Zabin and Hayward (1994) report that contraceptive use among teens and young

adults is greatest when sexual activity is planned and when partners have developed a close relationship. Substance use may affect these important determinants of birth control. Interestingly, this possibility suggests that substance use may affect contraception and sexual behavior even if it does not immediately precede sexual intercourse because of its earlier effects on partner selection and partner communication. Finally, risky sexual behavior may be an unintended consequence of experimentation with drugs and alcohol, which typically occurs when individuals are teenagers. At this stage of the life cycle, they may lack information about the potential harmful consequences of their actions and/or may discount these consequences very heavily.

Research on the relationship between substance use and risky sex among teenagers is especially compelling because the use of these substances rises from age 12 to approximately age 29, peaks among older teenagers and young adults, and has been increasing during most of the 1990s (Bridget F. Grant et al., 1991; National Institute on Drug Abuse, 1991; Johnston, O'Malley, and Bachman, 1999). Given this and the plausible mechanisms specified above linking substance use to sexual behavior, it is surprising that the Healthy People 2010 report (DHHS, 2000) fails to mention risk reduction objectives for substance use in conjunction with its year 2010 health status objectives for STDs. Risk reduction goals for drugs and alcohol are specified to achieve other health objectives associated with motor vehicle accident mortality, cirrhosis mortality, drug-related deaths, and drug-related hospital emergency department visits.

Numerous social scientists have studied the relationship between risky sexual behavior or premarital pregnancy and the use of marijuana, cocaine, and alcohol.[3] Although Leif C. Crowe and William H. George (1989) point out that alcohol and drugs may suppress sexual response in some subjects, the overwhelming majority of these researchers have documented a positive relationship between risky sex or premarital pregnancy and substance use.[4] They have not, however, demonstrated that the association between the two outcomes implies causality from substance use to risky sexual behavior. There are several competing explanations of the association.

According to Richard Jessor and Shirley L. Jessor's (1977) problem behavior theory, the two outcomes are manifestations of a common personality trait. This suggests that risky sex and substance use are associated because both are related to an unmeasured third variable such as a thrill-seeking personality. Leigh and Stall (1993) indicate that this interpretation is supported by the finding in many studies that cigarette smoking is highly correlated with risky sex. It is hard to argue that smoking causes temporary lapses in judgment or decision making. Thus, controlling for this third factor, substance use and sex will not be related. Alternatively, a teenager who chooses to have many sexual partners may use drugs and alcohol to cope with society's negative view of such behavior (M. Lynne Cooper, J. B. Skinner, and William H. George, 1990). In effect, the teenager consumes these substances to lower the psychic costs of risky sex. In this scenario, risky sex causes, but is not caused by, substance use, and accounting for this reverse causality will eliminate the dependence of sexual activity on substance use. Finally, substance use and sex may

simply be complementary consumption goods to a teenager, and like any complementary goods, they will be positively related. If this characterization is correct, then substance use is not a cause of teen sexual behavior, although changes in substance use will be associated with changes in sexual activity.

The differences among the alternative explanations of the substance use-risky sex relationship are crucial for the design of public policy. If substance use causes risky sex, then one appropriate policy would be to find ways to reduce substance use among teenagers and young adults. These policies, however, will not be successful in discouraging risky sex if both variables are caused by a common personality trait or if risky sex causes substance use. In the former case, public policies addressing either substance abuse or risky sex should focus on factors that affect the common personality trait. If risky sex causes substance use, the focus should be on unique determinants of risky sex such as sex education classes and free condom availability in schools.[5]

The large numbers of social science studies cited above have not established that substance use causes risky sex for a number of reasons. First, many of these studies use non-representative samples such as students at one or two high schools, or students in one metropolitan area.[6] Thus, the results of these studies may not be applicable to other localities as the relationship between substance use and sexual practices may differ across cities or regions. Moreover, most of these studies fail to control for a variety of family background and personal factors that may confound estimates of the relationship between substance use and sexual practices. Finally, no prior study has recognized the possibility that sexual activity may cause substance use (i.e., reverse causality).

In summary, many previous studies have demonstrated that substance use and sexual activity are positively related among teenagers and young adults, but only a few recent studies have attempted to estimate a causal effect.[7] In this study, the authors report only weak evidence that substance use is associated with teen sexual behavior, which suggests that substance use is not causally related to teen sexual behavior. However, as we describe below, this study has a major weakness that raises questions as to the correctness of its conclusions. Thus, the previous literature leaves unanswered the critical policy question of whether substance use causes risky sexual behavior and its consequences.

## 3. Research design and statistical methods

Our objective is to obtain estimates of the causal effects of substance use on sexual activity and risky sex. This task is complicated by two factors. First, an individual's sexual behavior and substance use may depend on a common set of personal and social factors, some of which are unmeasured. Second, the two behaviors may influence each other directly. We refer to these factors as statistical endogeneity and structural endogeneity, respectively. The research design and statistical methods we use are intended to address these problems.

The empirical models we use are based on linear specifications of the structural demand function for sexual activity ($S_t$), and the reduced form demand function for substance use ($D_t$). These demand functions can be derived from a simple utility maximizing model of behavior in which both sexual activity and substance use are valued by the teen and substance use is an input into the production of sexual activity. Specifically, we assume a production function of the form

$$S_t = F(D_t, T_t, ), \tag{1}$$

where $T_t$ is the time allocated to this activity. There may be other inputs, but we suppress them for simplicity. Note that the time input is not limited to the number of hours spent having intercourse, but also includes the time spent searching for partners and becoming acquainted with them.

The demand functions on which we focus are

$$S_t = \alpha_1 D_t + \alpha_2 q_t + \alpha_3 w_t + \alpha_4 v_t + \alpha_5 u + \varepsilon_t, \tag{2}$$

$$D_t = \gamma_1 q_t + \gamma_2 w_t + \gamma_3 v_t + \gamma_4 p_t + \gamma_5 u + \gamma_6 \varepsilon_t + \gamma_7 \eta_t. \tag{3}$$

In equations (2) and (3), u stands for measured and unmeasured determinants such as a thrill-seeking personality or a social environment that encourages deviant behavior. Other determinants of sexual activity are prices [e.g., condoms ($q$), time ($w$)] and income ($v$). The subscript $t$ is an index of time or age since we sometimes estimate these models on panel data, and $\varepsilon_t$ and $\eta_t$ are disturbance terms that are uncorrelated with each other and with $u$, $q_t$, $w_t$, and $v_t$. Intercepts have been suppressed for convenience. The determinants summarized by $u$ are assumed to be time-invariant and uncorrelated with $q_t$, $w_t$, and $v_t$.

Equation (2) specifies a conditional demand function for sexual activity (Robert A. Pollak, 1969; M. J. Browning, 1983) because the quantity of drugs rather than the price of drugs is one of its arguments. This equation can be viewed as the one that results from replacing the time input ($T_t$) in the production function with its determinants other than the price of drugs. Our hypothesis is that $D_t$ has a positive coefficient in the equation because $D$ has a positive marginal product in the production of $S_t$. That hypothesis could be explored directly by estimating the production function given by equation (1). We do not attempt to do so because the time input and other inputs that may enter the production function are not observed.

We acknowledge that the parameter of $D_t$ in equation (2) may be positive even if $D_t$ is not an input into the production of $S_t$. That would be the case if $S_t$ and $D_t$ were complements in consumption in the sense that a reduction in the price of $D_t$ causes both $D_t$ and $S_t$ to increase. We believe, however, that it is plausible to interpret a positive effect in the context of production theory because that interpretation is suggested by the literature on the relationship between substance use and sexual behavior that we have summarized.

The parameter of interest is $\alpha_1$, but simple ordinary least squares (OLS) estimates of equation (1) may be biased for two reasons: first, because unmeasured

components of $u$ are likely to be correlated with substance use (statistical endogeneity), and second because substance use may be correlated with $\varepsilon_t$ due to its causal dependence on sexual activity (structural endogeneity). Both of these possible correlations are reflected in equation (3), the reduced form demand function for substance use. Therefore, an alternative estimation strategy is necessary. We use two: a bivariate probit regression model, and an individual, fixed-effects regression model.

The bivariate probit model is based on the assumption that the unmeasured determinants of equations (2) and (3) have a joint, bivariate normal distribution. This procedure is applicable when sexual activity ($S_t$) and substance use ($D_t$) are measured as dichotomous variables, for example, a dummy variable indicating whether or not a teen had sexual intercourse in the last year and a dummy variable indicating whether or not a teen had an episode of binge drinking. The bivariate probit regression model accounts for the fundamental statistical problem associated with equation (2): the correlation between substance use (conditional on measured characteristics) and the unmeasured determinants ($u$ and $e_t$) of sexual activity. Importantly, it theoretically has the potential to control for both statistical and structural endogeneity since both problems result in a correlation between $D_t$ and $e_t$ in equation (2).

Equations (2) and (3) can be jointly estimated, but in this case, identification comes from functional form restrictions. In practice, however, the bivariate probit regression model performs relatively poorly (imprecise estimates) in these circumstances. To be effective, the bivariate probit model requires valid exclusion restrictions—variables that determine substance use, but do not affect sexual behavior. Substance use prices and control policies are often used for such purposes (e.g., Daniel I. Rees et al., 2001). However, for this identification strategy to be effective, it is necessary to have data that span several years, so that there is sufficient within-state variation in these policies, and relatively large samples to insure precise estimates. Most data sets, including the National Longitudinal Survey of Youth (NLSY97), do not have sufficient sample sizes, nor do they cover sufficiently long periods to be useful in this regard. This point is ignored by Rees et al. (2001) and the poor performance of the bivariate probit model in their case is revealed by the fact that the standard errors of the estimates of the effect of substance use from the bivariate probit model are 8–12 times as large as the corresponding standard errors from the univariate probit model.[8] Thus, the analysis of Rees et al. (2001) has little power to detect reliably true effects, and consequently, few definitive conclusions can be drawn from their analysis. Similar criticisms apply to the analysis of Sen (2002), which is similar to Rees et al. (2001) except for the data. An alternative identification strategy that does not rely on exclusion restrictions has been suggested by Joseph G. Altonji, Todd E. Elder, and Christopher R. Taber (2000). In this case, identification is achieved under the assumption of "equal selection between observed and unobserved variables."

The intuition behind the Altonji, Elder, and Taber (2000) strategy can be illustrated as follows. Re-write equations (2) and (3) in matrix form:

$$S = \alpha D + W'\Gamma \tag{4}$$

$$D = X\beta + u \tag{5}$$

Note that equation (4) represents the fully specified model and as such, includes no error term. The non-substance use determinants of sexual activity ($W'\Gamma$) can be divided into two parts: the observed and unobserved determinants, or:

$$S = \alpha D + X'\gamma + \varepsilon \tag{6}$$

where $X'\gamma$ is the observed component of sexual activity and ε is the unobserved component. The identification problem is that substance use is likely correlated with the unmeasured component, but substance use is also likely correlated with the measured component. These correlations can be expressed as follows:

$$\text{Proj}(D|X'\gamma, \varepsilon) = \phi_c X'\gamma + \phi_\varepsilon \varepsilon. \tag{7}$$

Equation (7) assumes that the correlation between substance use and the measured component of sexual activity is equal to the correlation between substance use and the unmeasured component of sexual activity. This is the equal selection rule, and it is justified for example, if the measured variables were chosen randomly from a large set of possible determinants, which is a reasonable assumption given that most secondary data sets used for economic analyses were not devised for the specific research question under investigation (Altonji, Elder, and Taber 2000).

Altonji, Elder, and Taber (2000) show that under certain conditions, the correlations in equation (7) are equal to the following:[9]

$$\phi_c = \frac{\text{Cov}(D, X'\gamma)}{\text{Var}(X'\gamma)} = \frac{\text{Cov}(X\beta + u, X'\gamma)}{\text{Var}(X'\gamma)} = \frac{\text{Cov}(X\beta, X'\gamma)}{\text{Var}(X'\gamma)}$$

and

$$\phi_c = \frac{\text{Cov}(D, \varepsilon)}{\text{Var}(\varepsilon)} = \frac{\text{Cov}(X\beta + u, \varepsilon)}{\text{Var}(\varepsilon)} = \frac{\text{Cov}(u, \varepsilon)}{\text{Var}(\varepsilon)}, \tag{8}$$

so

$$\frac{\text{Cov}(X\beta, X'\gamma)}{\text{Var}(X'\gamma)} = \frac{\text{Cov}(u, \varepsilon)}{\text{Var}(\varepsilon)}.$$

The important point here is that left hand side of the last equality in equation (8) can be estimated using observed data, and it is equal to the correlation coefficient in the bivariate (standard) normal distribution. This equality can be used to identify the model, since it provides an estimate of $\rho$—the correlation between the errors in the bivariate probit model. Estimation proceeds in steps. Initially, we assume that $\rho$—the correlation between the errors in the bivariate probit model—is zero and obtain estimates of

$$\frac{\mathrm{Cov}(X\beta, X'\gamma)}{\mathrm{Var}(X'\gamma)}$$

to use as an estimate of $\rho$. We then re-estimate the model to obtain a new estimate of

$$\frac{\mathrm{Cov}(X\beta, X'\gamma)}{\mathrm{Var}(X'\gamma)}$$

and $\rho$. We continue the process until the estimate of $\rho$ converges.

The second identification strategy we employ is an individual, fixed-effects (FE) regression model. The identifying assumption of the FE procedure is that the correlation between substance use and the unmeasured determinants of sexual activity in equation (2) is due to an unmeasured, time-invariant, person-specific effect ($u$). Longitudinal data can be used to eliminate the effect of this variable, for example, by taking first differences of the data.

$$S_t = \alpha_1 D_t + \alpha_2 q_t + \alpha_3 w_t + \alpha_4 v_t + \alpha_5 u + \varepsilon_t, \tag{9}$$

$$S_{t-1} = \alpha_1 D_{t-1} + \alpha_2 q_{t-1} + \alpha_3 w_{t-1} + \alpha_4 v_{t-1} + \alpha_5 u + \varepsilon_{t-1}, \tag{10}$$

$$\Delta S = \alpha_1 \Delta D + \alpha_2 \Delta q + \alpha_3 \Delta w + \alpha_4 \Delta v + \Delta\varepsilon. \tag{11}$$

The important aspect of equation (11) is that the influence of ($u$) on sexual behavior has been eliminated. This solves the statistical endogeneity problem, and if this was the only impediment to obtaining causal estimates, FE estimation will produce unbiased estimates of the structural parameter of the effect of substance use on sexual activity. We acknowledge that the FE procedure does not address structural endogeneity. We view it as a useful complement to the bivariate probit model because it reveals how the effect of substance use on sexual behavior changes when only time-invariant factors are taken into account.

## 4. Data

The data for the analysis comes from the 1997 cohort of the NLSY97. These data represent a panel data collection of approximately 8500 youths who were between the ages of 12 and 16 as of December 31, 1996.[10] These youths were interviewed in 1997 for the first time and were re-interviewed in 1998 and 1999. Information about sexual activity was obtained for all youths age 15 and over, so by 1999, information about sexual activity of almost all the respondents in the NLSY97 has been obtained.[11] As noted we use two estimation strategies. The sample used to estimate the bivariate probit model consists of all respondents with valid information on sexual activity, substance use and other variables used in the regression model. For this analysis, a teen is included in the sample only one time—the first time such information is available for that teen. A similar sample is used to estimate the FE model, but in this case, only those respondents with 2 years of information are included in

the sample. We take the first two observations for each individual if there are more than two.

The following dependent variables are used: (1) the number of occasions of sexual intercourse in the past year, and (2) the number of times birth control was used in the past year. Using the first two variables, we define a measure of risky sex, which is equal to one if the person was sexually active in the last year and used birth control less than 90% of the time, as measured by the ratio of the number of times used birth control to the number of occasions of sexual intercourse. For those respondents who were not sexually active, risky sex is set to zero. Approximately 40% of the sexually active respondents are defined to have engaged in risky sex by our definition, and this is in line with estimates of contraceptive use from other surveys (e.g., Darroch et al., 2001).

Our explanatory variables of interest are alcohol and marijuana use. We measure alcohol use in two ways: whether or not a respondent has drank alcohol in the last 30 days, and whether or not a respondent has had an episode of binge drinking in the past 30 days. Binge drinking is defined as having five or more drinks at one time. Marijuana use is measured as a dummy variable equal to one if a respondent used marijuana in the past 30 days. We recognize that the alcohol and marijuana use information is based on the past 30 days whereas the information on sexual behavior refers to the past year. This undoubtedly creates some measurement error, but we view the measures of alcohol and marijuana use as markers for the use of such substances and indicative of a youth's overall involvement in the use of these substances. For example, binge drinking in the past 30 days represents a greater involvement in alcohol use than does an indicator of any use in the past 30 days. Therefore, we should observe a dose-response type of relationship if there is a true underlying causal relationship between alcohol use and sexual behavior.

A variety of other personal and family characteristics were included in the regression models. Two model specifications were used. The first includes only a limited set of covariates: race/ethnicity dummy variables, age dummy variables, highest grade completed dummy variables, weight, height, weight*height, health (poor) status, and year dummy variables. The second specification adds the following covariates: percent of families in poverty in county of residence, mother's education, dummy variable for missing mother's education, mother's age at first birth, dummy variable for missing mother's age at first birth, indicator of two-parent family at age 12, dummy variables for current family structure, indicator of whether or not respondent destroyed property, indicator of whether or not respondent stole goods, indicator of whether or not respondent smoked, number of cigarettes smoked in last 30 days, several dummy variables indicating peer behaviors (church attendance, college plans, cut school, and sports participation), and dummy variables for school environment (cheating and perceptions of discipline). We use two specifications to assess the sensitivity of estimates of the effect of alcohol and marijuana use on sexual behavior to the inclusion of measured characteristics. For example, if estimates are relatively unaffected by the addition of several variables that are strongly correlated with sexual behavior, it is reasonable to conclude that there is

relatively little selection on either observed or unobserved variables. On the other hand, if the addition of a few variables results in relatively large changes in estimated effects, it is reasonable to conclude that selection is a significant problem, and then it is necessary to decide whether or not there is more or less selection on observed characteristics than unobserved characteristics. We discuss these issues in more detail below.

## 5. Results

### *5.1. Descriptive Analysis*

Table 1 presents sample means and proportions of selected variables by gender and alcohol use, as measured by binge drinking in the past 30 days. The figures indicate that sexual activity is more prevalent among teens who binge drink compared to teens who do not binge drink: more than half of teens who binge drink have had sexual intercourse in the past year whereas only 20% of teens who do not binge drink have had sexual intercourse in the last year. Similar differences in sexual activity by binge drinking status are observed for risky sex, and tables in the appendix show that other measures of substance use are also strongly correlated with sexual activity. These findings are consistent with the large number of previous studies that demonstrate a significant correlation between sexual activity and substance use among teenagers.

The figures in Table 1, however, also reveal that substance use is correlated with other characteristics of teenagers. Teens who are binge drinkers are more likely to have destroyed property, stole goods, and smoked cigarettes than are teens who do not binge drink. Similarly, binge drinkers have peer groups that are less likely to attend religious services, more likely to cut school, and less likely to play a sport than teens who are not binge drinkers. It is clear from the data that there is ''positive selection'' on observed characteristics—teens who drink are more likely to be sexually active, and more likely to engage in other risky (e.g., smoke cigarettes) behaviors than teens who do not drink. This selection on observed characteristics strongly suggests that there will be selection on unobserved characteristics, and provides evidence to question whether or not the often found association between substance use and sexual activity is causal. It may be the case that a common set of underlying determinants cause both outcomes. For example, teens with a thrill-seeking personality are expected to be more likely to drink and use marijuana and more likely to engage in sexual activity. It is precisely this possibility that this analysis is intended to address. We now turn to the multivariate results.

### *5.2. Bivariate probit analysis*

Estimates of the effect of alcohol and marijuana use on sexual activity are presented in Table 2. Estimates were obtained from maximum likelihood univariate and bivariate probit regression models using respondents from the NLSY97 who are

*Table 1.* Sample means of selected characteristics by gender and past 30 day binge drinking respondents age 15–17 in NLSY97.

| | Males binge drinker past 30 days | | Females binge drinker past 30 days | |
|---|---|---|---|---|
| Variable | Yes | No | Yes | No |
| Had sex past 12 months | 0.547** | 0.216 | 0.505** | 0.198 |
| Risky sex past 12 months | 0.236** | 0.070 | 0.264** | 0.077 |
| Black | 0.130** | 0.267 | 0.123** | 0.281 |
| Hispanic | 0.234 | 0.209 | 0.248** | 0.208 |
| Age (in months) | 193.4** | 190.6 | 191.8** | 191.1 |
| Height (in inches) | 69.11** | 68.65 | 64.38 | 64.26 |
| Weight (in pounds) | 156.25 | 154.10 | 131.53 | 131.99 |
| Poor health | 0.266** | 0.224 | 0.393** | 0.309 |
| Highest grade completed | 8.90** | 8.73 | 8.96 | 8.92 |
| Two bio. parents age 12 | 0.456** | 0.474 | 0.423 | 0.447 |
| Mother's education | 12.45 | 12.49 | 12.34 | 12.39 |
| Mother's age at 1st birth | 22.60 | 22.92 | 23.08 | 22.80 |
| Pct. fam. in poverty (county) | 10.49** | 11.08 | 10.15** | 11.03 |
| Destroyed property | 0.620** | 0.352 | 0.443** | 0.181 |
| Stole goods > \$50 | 0.624** | 0.356 | 0.554** | 0.277 |
| Peers attend church | 0.209** | 0.252 | 0.220 | 0.258 |
| Peers cut school | 0.473** | 0.357 | 0.555** | 0.454 |
| Peers play sport | 0.578 | 0.597 | 0.541** | 0.603 |
| Smoked cigarettes | 0.846** | 0.388 | 0.868** | 0.405 |
| Marijuana use past 30 days | 0.479** | 0.069 | 0.449** | 0.060 |
| Number of observations | 640 | 3364 | 455 | 3425 |

Notes:
1. Number of observations listed represents the maximum number. For several variables the actual sample size is slightly less because of missing information.
2. For variables with significant amount (e.g., 10% of sample) of missing information, we assigned the sample mean and created a variable indicating that for this observation the information was missing. Variables falling into this category are: mother's education, mother's age at first birth, weight, and proportion of peers that attend church regularly.
3. Asterisks (**) next to a number indicates that the difference between drinkers and non-drinkers is statistically significant ($p < 0.05$).

between the ages of 15 and 17. All models were estimated separately by gender. For each of the three measures of substance use listed in the table, three separate estimates are obtained. The first column shows estimates from a univariate probit regression that includes only a limited set of covariates (see notes to table); column two shows estimates from a univariate probit regression that includes an extended set of covariates (see notes to table); and column three presents estimates from a bivariate probit regression that includes the larger set of independent variables.

Estimates in column one indicate that alcohol and marijuana use are positively associated with sexual activity and risky sex for both males and females. The magnitudes of the estimates are very large. For example, males who are binge drinkers

*Table 2*. Estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse and risky sex in last 12 months respondents age 15–17 in NLSY97.

| | Males | | | Females | | |
|---|---|---|---|---|---|---|
| *Sexual intercourse* | Univariate probit | Univariate probit | Bivariate probit | Univariate probit | Univariate probit | Bivariate probit |
| Alcohol use past 30 days | 0.889** | 0.587** | 0.431 | 0.722** | 0.316** | −0.776 |
| | (0.050) | (0.057) | (0.502) | (0.051) | (0.060) | (0.488) |
| | [0.301] | [0.182] | | [0.230] | [0.087] | |
| Rho ($\rho$) | | | 0.091 | | | 0.646** |
| | | | (0.291) | | | (0.291) |
| Binge drinking past 30 days | 1.046** | 0.688** | −0.316 | 0.897** | 0.444** | −0.341 |
| | (0.060) | (0.067) | (0.284) | (0.067) | (0.075) | (0.411) |
| | [0.377] | [0.228] | | [0.315] | [0.132] | |
| Rho ($\rho$) | | | 0.551** | | | 0.434 |
| | | | (0.147) | | | (0.223) |
| Marijuana use past 30 days | 1.090** | 0.609** | −0.271 | 1.241** | 0.657** | −0.137 |
| | (0.062) | (0.072) | (0.323) | (0.071) | (0.081) | (0.459) |
| | [0.397] | [0.203] | | [0.446] | [0.207] | |
| Rho ($\rho$) | | | 0.482** | | | 0.431 |
| | | | (0.169) | | | (0.239) |
| *Risky sex* | | | | | | |
| Alcohol use past 30 days | 0.654** | 0.376** | −0.507 | 0.689** | 0.372** | −0.413 |
| | (0.060) | (0.070) | (0.590) | (0.061) | (0.072) | (0.943) |
| | [0.116] | [0.050] | | [0.124] | [0.047] | |
| Rho ($\rho$) | | | 0.365 | | | 0.272 |
| | | | (0.344) | | | (0.553) |
| Binge drinking past 30 days | 0.794** | 0.483** | −0.751 | 0.810** | 0.473** | −1.523** |
| | (0.067) | (0.076) | (0.540) | (0.074) | (0.082) | (0.664) |
| | [0.166] | [0.073] | | [0.180] | [0.070] | |
| Rho ($\rho$) | | | 0.549 | | | 0.898** |
| | | | (0.314) | | | (0.272) |
| Marijuana use past 30 days | 0.787** | 0.407** | −0.610 | 1.027** | 0.596** | −0.810 |
| | (0.069) | (0.080) | (0.523) | (0.074) | (0.085) | (0.601) |
| | [0.168] | [0.061] | | [0.245] | [0.095] | |
| Rho ($\rho$) | | | 0.435 | | | 0.642 |
| | | | (0.304) | | | (0.323) |
| Additional covariates | No | Yes | Yes | No | Yes | Yes |
| N | 3926 | 3727 | 3727 | 3792 | 3635 | 3635 |

Notes:

1. Standard errors in parentheses.
2. Marginal effects in brackets. Marginal effects were approximated using the discrete change from zero to one.
3. Risky sex is equal to zero if a respondent did not have sex in past 12 months, or if they did have sex, they used birth control (unspecified type) more than 90% of the time.
4. Basic set of covariates are: race/ethnicity dummy variables, age dummy variables, highest grade completed dummy variables, weight, height, weight*height, health (poor) status, and year dummy variables.
5. Additional covariates are: percent of families in poverty in county of residence, mother's education,

dummy variable for missing mother's education, mother's age at first birth, dummy variable for missing mother's age at first birth, indicator of two-parent family at age 12, dummy variables for current family structure, indicator of whether or not respondent destroyed property, indicator of whether or not respondent stole goods, indicator of whether or not respondent smoked, number of cigarettes smoked in last 30 days, several dummy variables indicating peer behaviors (church attendance, college plans, cut school, and sports participation), and dummy variables for school environment (cheating and perceptions of discipline).

have a probability of having sex in the past year that is 38% points greater than males who are not binge drinkers. The similar figure for females is 32% points. These two estimates are close to the unadjusted differences evident in Table 1. There is also a dose-response effect for alcohol use: binge drinking has a larger effect than an indicator of any drinking in the past 30 days. Among males, effects of marijuana use on sexual activity and risky sex are similar in magnitude to the effects of binge drinking, but among females, marijuana use has even a larger effect on sexual activity and risky sex than does binge drinking. Overall, the estimates in column one are consistent with the numerous studies that show a strong, positive association between substance use and sexual behavior.

Column two presents estimates from an expanded regression model that includes several additional explanatory variables. Estimates in this column are approximately half the size of those in column one. The diminished magnitudes of the estimates indicate that there is considerable selection on observable characteristics, a finding consistent with the simple descriptive statistics in Table 1. The significant selection on observable characteristics raises the likelihood that there is also selection on unobserved characteristics. In fact, the pseudo R-square of these regression is approximately 0.25, so a significant amount of unexplained variation in sexual behavior remains. Thus, it is likely that the estimates in column two represent upper bound estimates of the effects of substance use on sexual behavior.

The third column presents estimates from a bivariate probit regression. Theoretically, this estimation procedure accounts for any selection on unobserved variables, but as noted earlier, this procedure is most efficacious when there are valid exclusion restrictions. In the current case, there are no exclusion restrictions and the model is identified by functional form restrictions. Consequently, we place little credence in these estimates and show them primarily to illustrate their poor performance. In general, estimates of the effects of substance use in column three are negative and not statistically significant. The sign reversal is due to the strong positive selection on unobserved variables indicated by estimates of the correlation coefficient ($\rho$), which are large and positive, but usually not statistically significant. The imprecision of the estimates of the effects of substance use and the correlation coefficients illustrate the weakness of this estimation strategy. These estimates are very large, but not statistically significant, suggesting that this estimation strategy has little power to detect true effects.

We next turn to the estimates obtained using the identification strategy suggested by Altonji, Elder, and Taber (2000). In this model, identification is achieved by assuming that the amount of selection on unobserved variables is equal to the amount of selection

on observed variables, or the equal selection rule. In Table 3, we present these estimates along with several estimates from constrained bivariate probit regressions for which values of the correlation coefficient ($\rho$) are fixed. We assume that there is positive selection on unobserved characteristics, which is consistent with evidence in Table 1 and estimates of such selection in Table 2, and set $\rho$ equal to 0.1, 0.2, 0.3, 0.4, and 0.5. Estimates from these models reveal how much selection on unobserved characteristics is necessary to eliminate the positive association between substance use and sexual behavior. We assume that lower bound estimates of the effects of substance use on sexual behavior are zero since there is little reason to expect that substance use reduces sexual activity and risky sexual behavior.

Estimates in Table 3, which are for the male sample, indicate that a relatively small amount of selection on unobserved characteristics can eliminate the positive association between substance use and sexual behavior. For example, a correlation

*Table 3*. Estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse and risky sex in last 12 months male respondents age 15–17 in NLSY97.

| *Sexual intercourse* | Univariate probit $\rho=0$ | Constrained bivariate probit $\rho=0.1$ | Constrained bivariate probit $\rho=0.2$ | Constrained bivariate probit $\rho=0.3$ | Constrained bivariate probit $\rho=0.4$ | Constrained bivariate probit $\rho=0.5$ | Bivariate probit $\rho=\frac{\mathrm{Cov}(X'\beta, X'\gamma)}{\mathrm{Var}(X'\gamma)}$ |
|---|---|---|---|---|---|---|---|
| Alcohol use past 30 days | 0.587** (0.057) | 0.416** (0.057) | 0.244** (0.057) | 0.069 (0.056) | −0.108** (0.054) | −0.288** (0.052) | −0.146** (0.054) [$\rho=0.42$] |
| Binge drinking past 30 days | 0.688** (0.067) | 0.511** (0.067) | 0.332** (0.066) | 0.151** (0.065) | −0.033 (0.063) | −0.219** (0.061) | −0.259** (0.060) [$\rho=0.52$] |
| Marijuana use past 30 days | 0.609** (0.072) | 0.430** (0.072) | 0.250** (0.071) | 0.067 (0.070) | −0.118 (0.068) | −0.305** (0.066) | −0.868** (0.053) [$\rho=0.80$] |
| *Risky sex* | | | | | | | |
| Alcohol use past 30 days | 0.376** (0.070) | 0.207** (0.069) | 0.037 (0.068) | −0.133** (0.067) | −0.304** (0.066) | −0.477** (0.063) | −0.960** (0.054) [$\rho=0.77$] |
| Binge drinking past 30 days | 0.483** (0.076) | 0.309** (0.075) | 0.134 (0.074) | −0.040 (0.073) | −0.214** (0.071) | −0.389** (0.069) | −1.060** (0.051) [$\rho=0.89$] |
| Marijuana use past 30 days | 0.407** (0.080) | 0.232** (0.080) | 0.056 (0.079) | −0.118 (0.078) | −0.292** (0.076) | −0.464** (0.073) | NA [$\rho > 1$] |
| Additional covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes:
1. $\rho = (\mathrm{Cov}(X'\beta, X'\gamma))/(\mathrm{Var}(X'\gamma))$ is calculated by a two−step procedure. An initial value for rho was chosen and then the bivariate probit model was estimated. Rho was then re−calculated using estimates of $X'\beta$ and $X'\gamma$ from the bivariate probit. This procedure was repeated until estimates of rho converged.
2. See notes to Table 2.

coefficient ($\rho$) of between 0.2 and 0.3 eliminates the positive associations between alcohol or marijuana use in the past 30 days, and the probability of having sexual intercourse in the last year. For the risky sex outcome, a correlation coefficient ($\rho$) of between 0.1 and 0.2 is sufficient to eliminate statistically significant associations. Thus, unless there is very little selection on unobserved characteristics, there is little likelihood that the positive associations between substance use and sexual activity that are often found are causal.

One method for assessing how much selection there is on unobservable variables is to assume that it is equal to the selection on observed variables. This is the suggestion of Altonji, Elder, and Taber (2000). In the last column of Table 3, we present estimates from a bivariate probit model for which the equal selection rule is used to identify the model. As can be seen, there is a significant amount of selection on observed characteristics, and if this amount of selection also characterizes unobserved variables, the positive associations between substance use and sexual behavior are eliminated.[12] In fact, estimates indicate that selection on unobservable variables would have to be less than half the amount of selection on observed variables for the positive associations between substance use and sexual behavior to remain.[13] While we have included some very powerful correlates of sexual behavior and substance use in the model, for example cigarette consumption, it is unlikely, given the relatively large amount of unexplained variation that remains, that there is not additional selection on unobservable factors. We cannot definitively say how much selection remains, but it would have to be significantly less than the amount of selection on observables for there to be a positive association between substance use and sexual activity and substance use and risky sex.

Table 4 presents estimates similar to those in Table 3, but in this case for the female sample. Here again, we see that a relatively small amount of selection on unobserved variables will eliminate the positive associations between substance use and sexual behavior. In this case, a correlation coefficient ($\rho$) of between 0.1 and 0.2 is sufficient to eliminate most of the positive associations between substance use and sexual behavior. Applying the equal selection rule reveals that the amount of selection on observed variables is substantial and if the same amount of selection characterized unobserved variables, the positive associations between substance use and sexual behavior are eliminated. In fact, the amount of selection on unobservable variables would have to be approximately one-fifth of the amount of selection on observed characteristics for there to be any positive associations between substance use and sexual behavior.[14]

### *Fixed-effects estimates*

The second approach we use to obtain causal estimates is an individual, fixed effects regression model. This method yields unbiased estimates of the causal effect of substance use on sexual behavior if the only unobserved variable is a time-invariant, person-specific effect, for example, a thrill-seeking motivation. We estimate the

*Table 4*. Estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse and risky sex in last 12 months female respondents age 15–17 in NLSY97.

| Sexual intercourse | Univariate probit $\rho=0$ | Constrained bivariate probit $\rho=0.1$ | Constrained bivariate probit $\rho=0.2$ | Constrained bivariate probit $\rho=0.3$ | Constrained bivariate probit $\rho=0.4$ | Constrained bivariate probit $\rho=0.5$ | Bivariate probit $\rho=\frac{\mathrm{Cov}(X'\beta,X'\gamma)}{\mathrm{Var}(X'\gamma)}$ |
|---|---|---|---|---|---|---|---|
| Alcohol use past 30 days | 0.316** (0.060) | 0.147** (0.060) | −0.022 (0.059) | −0.192** (0.058) | −0.361** (0.056) | −0.530** (0.054) | −0.635** (0.053) [$\rho=0.56$] |
| Binge drinking past 30 days | 0.444** (0.075) | 0.264** (0.075) | 0.084 (0.074) | −0.097 (0.072) | −0.279** (0.071) | −0.461** (0.068) | −0.611** (0.065) [$\rho=0.58$] |
| Marijuana use past 30 days | 0.657** (0.081) | 0.477** (0.080) | 0.295** (0.079) | 0.110 (0.078) | −0.077 (0.076) | −0.267** (0.073) | −0.969** (0.054) [$\rho=0.86$] |
| *Risky sex* | | | | | | | |
| Alcohol use past 30 days | 0.372** (0.072) | 0.203** (0.072) | 0.034 (0.071) | −0.136 (0.070) | −0.308** (0.068) | −0.480** (0.066) | −0.793** (0.060) [$\rho=0.68$] |
| Binge drinking past 30 days | 0.473** (0.082) | 0.296** (0.082) | 0.120 (0.081) | −0.056 (0.079) | −0.231** (0.077) | −0.406** (0.074) | −0.791** (0.065) [$\rho=0.72$] |
| Marijuana use past 30 days | 0.596** (0.085) | 0.419** (0.085) | 0.241** (0.084) | 0.062 (0.082) | −0.118 (0.080) | −0.298** (0.077) | NA [$\rho > 1$] |
| Additional covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes:
1. $\rho = (\mathrm{Cov}(X'\beta, X'\gamma))/(\mathrm{Var}(X'\gamma))$ is calculated by a two-step procedure. An initial value for rho was chosen and then the bivariate probit model was estimated. Rho was then re-calculated using estimates of $X'\beta$ and $X'\gamma$ from the bivariate probit. This procedure was repeated until estimates of rho converged.
2. See notes to Table 2.

model by OLS using a sample of respondents who have been interviewed twice and who were age 15–17 at the time of their first interview.[15]

Cross-sectional and fixed-effect estimates are presented in Table 5. The left panel shows estimates for males, and the right panel shows estimates for females. We present both univariate probit and OLS cross-sectional estimates to show that the method of estimation makes little difference. Column three in each panel lists the fixed-effects estimates. Cross-sectional estimates in Table 5 confirm earlier findings and indicate that there is a strong positive association between substance use and sexual activity, and substance use and risky sex for both males and females. Again, the magnitudes of the cross-sectional estimates are very large: binge drinking is associated with a 21% point increase in the probability that a teenage male had sexual intercourse in the last year, and a 15% point increase in the probability that a teenage female had sexual intercourse in the last year. These estimates are quite similar to those in Table 2, which

*Table 5*. Fixed-effect estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse and risky sex in last 12 months respondents age 15–17 in NLSY97.

| | Males | | | Females | | |
|---|---|---|---|---|---|---|
| Sexual intercourse | Univariate probit | OLS | OLS fixed-effects | Univariate probit | OLS | OLS fixed-effects |
| Alcohol use past 30 days | 0.584** (0.062) [0.192] | 0.166** (0.017) | 0.052** (0.013) | 0.310** (0.066) [0.089] | 0.081** (0.017) | 0.022 (0.014) |
| Binge drinking past 30 days | 0.671** (0.073) [0.233] | 0.212** (0.021) | 0.075** (0.016) | 0.466** (0.084) [0.145] | 0.147** (0.023) | 0.082** (0.017) |
| Marijuana use past 30 days | 0.574** (0.078) [0.200] | 0.200** (0.023) | 0.049** (0.018) | 0.573** (0.090) [0.184] | 0.202** (0.024) | 0.050** (0.020) |
| *Risky sex* | | | | | | |
| Alcohol use past 30 days | 0.358** (0.075) [0.052] | 0.060** (0.013) | 0.038** (0.013) | 0.383** (0.080) [0.051] | 0.066** (0.013) | 0.040** (0.014) |
| Binge drinking past 30 days | 0.449** (0.081) [0.072] | 0.094** (0.015) | 0.079** (0.016) | 0.438** (0.091) [0.066] | 0.101** (0.017) | 0.086** (0.017) |
| Marijuana use past 30 days | 0.285** (0.087) [0.044] | 0.068** (0.017) | 0.035** (0.018) | 0.558** (0.095) [0.091] | 0.156** (0.019) | 0.076** (0.020) |
| Additional covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 3076 | 3076 | 6375 | 3034 | 3034 | 6184 |

Notes:
1. See notes to Table 2.

is important since it establishes that any observed differences between fixed effects and OLS estimates are not due to sample composition.

FE estimates are presented in column three. These estimates are dramatically smaller than the cross-sectional estimates; often only one-third to one-quarter the size of the cross-sectional estimates. For example, binge drinking is now associated with a 7% point increase in the probability that a male will have had sexual intercourse; in Table 2, the similar estimate was 22% points. However, all of the FE estimates are positive and all but one are statistically significant. Thus, controlling for time-invariant, person-specific effects substantially reduces the magnitudes of the effects of substance use, but does not eliminate the positive association between substance use and sexual behavior. The question that this approach leaves unanswered is how important are unmeasured, time-varying effects? In this regard the bivariate probit model is superior because it accounts for all unmeasured factors, but that analysis was hampered by its inability to definitively assess the amount of selection on unobserved variables.

*5.4 Results using national longitudinal survey of adolescent health (Ad-Health)*

To investigate the robustness of our estimates, we obtained estimates of the effect of substance use on sexual behavior using a sample of young adults drawn from the National Longitudinal Survey of Adolescent Health (Ad-Health). Ad-Health, described in detail by Peter S. Bearman, Jo Jones, and J. Richard Udry (1997) and conducted by the University of North Carolina Population Center, is made up of approximately 12,000 youth in the seventh through twelfth grades from 80 schools who were interviewed in 1995 and 1996.[16] We use the public use sample, which is a 50% random sample. Importantly, Ad-Health contains similar data about sexual behavior, substance use, and other personal and family information as that contained in the NLSY97. Therefore, we are able to estimate models that are nearly equivalent to those used in the analysis of the NLSY97. We limit the Ad-Health sample to youths aged 15–17 since this was the age range for the NLSY97 sample.

Table 6 presents estimates of the effect of substance use on the probability of having sexual intercourse in the last year. Estimates in the top panel refer to the male sample, and those in the bottom panel refer to the female sample. Estimates in the first column confirm the well established fact that sexual activity and substance use are positively correlated. For both males and females, alcohol and marijuana use are positively associated with sexual intercourse in the last year. The estimates in column one of Table 6 have magnitudes of comparable size to those in Tables 3 and 4, although there is less evidence of a dose-response effect for alcohol use in Table 6 since the estimates of the effect of binge drinking are approximately the same size as the estimates of the effect of any alcohol use in past 30 days. It is also the case that a relatively small amount of selection on unobservable variables eliminates the positive effects of substance use. A correlation coefficient ($\rho$) of 0.2 is, in all but one case, sufficient to eliminate statistically significant associations between substance use and sexual intercourse. Finally, if we assume that the amount of selection on unobserved variables is equal to the amount of selection on observed variables, all estimates of the effect of substance use are negative (last column), and estimates of the correlation among the errors derived under the equal selection rule are large and positive. These results are very similar to those obtained using the NLSY97 sample. These results indicate that only if there is relatively little selection on unobserved variables—approximately one-fourth of the amount of selection on observed variables—will the positive association between substance use and sexual activity remain.

FE estimates for the Ad-Health sample are presented in Table 7. In general, the FE estimates are positive and small in magnitude, and only one of six are statistically significant. Controlling for time-invariant person characteristics greatly reduces the effect of substance use on sexual activity. For example, the estimate of the effect of binge drinking on sexual intercourse for males is 0.041; binge drinkers probability of sexual intercourse in the last year is 4% points higher than non-binge drinkers. This represents a 12% increase relative to the

*Table 6*. Estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse in last 12 months male and female respondents age 15–17 in Ad-Health.

| | Univariate probit $\rho = 0$ | Constrained bivariate probit $\rho = 0.1$ | Constrained bivariate probit $\rho = 0.2$ | Constrained bivariate probit $\rho = 0.3$ | Constrained bivariate probit $\rho = 0.4$ | Constrained bivariate probit $\rho = 0.5$ | Bivariate probit $\rho = \frac{\mathrm{Cov}(X'\beta, X'\gamma)}{\mathrm{Var}(X'\gamma)}$ |
|---|---|---|---|---|---|---|---|
| *Males* | | | | | | | |
| Alcohol use past 30 days | 0.506** (0.083) | 0.336** (0.083) | 0.165** (0.082) | −0.081 (0.080) | −0.183** (0.078) | −0.356** (0.076) | −0.936** (0.060) [$\rho = 0.83$] |
| Binge drinking past 30 days | 0.494** (0.094) | 0.320** (0.094) | 0.145 (0.093) | −0.032 (0.091) | −0.210** (0.089) | −0.390** (0.086) | −0.587** (0.081) [$\rho = 0.61$] |
| Marijuana use past 30 days | 0.398** (0.102) | 0.222** (0.102) | 0.045 (0.100) | −0.134 (0.099) | −0.314** (0.096) | −0.495** (0.093) | NA [$\rho > 1$] |
| *Females* | | | | | | | |
| Alcohol use past 30 days | 0.402** (0.083) | 0.232** (0.083) | 0.061 (0.082) | −0.111 (0.080) | −0.283** (0.078) | −0.456** (0.076) | −0.993** (0.061) [$\rho = 0.82$] |
| Binge drinking past 30 days | 0.422** (0.107) | 0.243** (0.107) | 0.061 (0.106) | −0.122 (0.104) | −0.307** (0.101) | −0.493** (0.098) | −0.815** (0.089) [$\rho = 0.67$] |
| Marijuana use past 30 days | 0.423** (0.109) | 0.247** (0.109) | 0.069 (0.107) | −0.111 (0.106) | −0.292 (0.103) | −0.476** (0.099) | −1.346** (0.064) [$\rho = 0.96$] |
| Additional covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes:

1. $\rho = (\mathrm{Cov}(X'\beta, X'\gamma))/(\mathrm{Var}(X'\gamma))$ is calculated by a two-step procedure. An initial value for rho was chosen and then the bivariate probit model was estimated. Rho was then re-calculated using estimates of $X'\beta$ and $X'\gamma$ from the bivariate probit. This procedure was repeated until estimates of rho converged.
2. See notes to Table 2.

unadjusted mean of non-binge drinkers. In sum, estimates in Table 7 confirm our earlier findings from the NLSY97 analysis, and establish that time-invariant person effects account for a large portion of the covariance between substance use and sexual behavior. Indeed, in Ad-Health, the fixed effects estimates are not statistically significant, although they are positive. The lack of statistical significance may reflect the smaller sample sizes associated with Ad-Health, but the estimates in Table 7 are smaller in magnitude than those in Table 5.

## 6. Conclusions

In this study, we have examined the effect of alcohol and marijuana use on sexual behavior of teens. We were particularly interested in investigating whether the widely

*Table 7*. Fixed-effect estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse in last 12 months respondents age 15–17 in Ad-Health.

| | Males | | | Females | | |
|---|---|---|---|---|---|---|
| Sexual intercourse | Univariate probit | OLS | OLS fixed-effects | Univariate probit | OLS | OLS fixed-effects |
| Alcohol use past 30 days | 0.414** (0.092) [0.156] | 0.142** (0.029) | 0.028 (0.029) | 0.319** (0.092) [0.120] | 0.108** (0.030) | 0.062** (0.028) |
| Binge drinking past 30 days | 0.418** (0.103) [0.160] | 0.147** (0.034) | 0.041 (0.031) | 0.312** (0.121) [0.120] | 0.109** (0.038) | 0.010 (0.034) |
| Marijuana use past 30 days | 0.330** (0.113) [0.127] | 0.118** (0.038) | 0.036 (0.036) | 0.416** (0.120) [0.160] | 0.141** (0.038) | -0.010 (0.037) |
| Additional covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 1208 | 1208 | 2318 | 1269 | 1269 | 2451 |

Notes:
1. See notes to Table 2.

documented positive association between substance use and risky sexual behavior would remain in analyses that controlled for both measured and unmeasured heterogeneity. Toward this end we used two statistical approaches—bivariate probit and individual fixed effects—to assess the strength of the association between substance use and sexual behavior. Under certain conditions, these approaches will yield unbiased estimates of causal effects of substance use on sexual behavior.

Overall, estimates in Tables 2 through 4, and Table 6 suggest that it is unlikely that the often found positive associations between substance use and sexual activity and risky sexual behavior reflect true causal relationships. It appears more likely that these positive associations reflect the influence of omitted variables. This conclusion is based on the following evidence. First, there was a significant amount of selection on observable variables that suggests that selection on unobserved variables is likely important. Second, a significant amount of unexplained variation in sexual behavior remained even after the addition of several powerful predictors of such behavior; the pseudo R-square statistic for most of the regressions was approximately 0.25. The relatively large amount of unexplained variation and the substantial amount of selection on observable variables suggests that there is at least some selection on unobservable characteristics. Third, using the equal selection rule of Altonji, Elder, and Taber (2000) eliminated the positive association between substance use and sexual behavior. In fact, the amount of selection on unobservable variables would have to be significantly less than the amount of selection on observable variables for there to be a positive association between substance use and sexual behavior; among men, the amount of selection on unobservable variables would have to be approximately half the amount of selection on observable variables; and among females the

amount of selection on unobservable variables would have to be approximately one-fourth the amount of selection on observable variables. Thus, Tables 2–4 provide substantial evidence to suggest that the positive association between substance use and sexual behavior of teens is not causal, particularly for females. However, the absence of a truly definitive way to estimate the amount of selection on unobservable variables leaves open the possibility that there is a small positive association between substance use and sexual behavior.

FE estimates also show that there is significant selection on unobservable factors that confound estimates of the relationship between substance use and sexual behavior. In this case, accounting for a person-specific effect dramatically reduced estimates of the association between substance use and sexual behavior. Nevertheless, a significant positive association sometimes remained. Thus, it is not possible to rule out that this represents a true causal effect. However, the FE estimation strategy does not control for time-varying factors that are unmeasured. For young people, it is likely that there are many personality and behavioral factors that are still developing and influencing decisions related to sexual behavior and substance use. Thus, the FE methodology may not be adequate.

In sum, we believe it is reasonable to view the FE estimates as upper bound estimates and zero as lower bound estimates. We set the lower bound at zero because there is little theoretical support to justify a negative effect of substance use on sexual activity and risky sex. Evidence derived from the bivariate probit analysis strongly suggests that the truth lies closer to the lower bound than the upper bound, particularly for females, but as noted, this is not a definitive approach. Therefore, we conclude as follows:

- 1Alcohol or marijuana use in the past 30 days is associated with between 0 and 5% point (25% of non-drinker mean) increase in the probability that a teenage male has had sexual intercourse in the last year.
- Binge drinking in the past 30 days is associated with between 0 and 8% point (36% of non-drinker mean) increase in the probability that a teenager (male or female) has had sexual intercourse in the last year.
- Alcohol use in the past 30 days is associated with between 0 and 2% point (12% of non-drinker mean) increase in the probability that a teenage female has had sexual intercourse in the last year.
- Marijuana use in the past 30 days is associated with between 0 and 5% point (25% of non-drinker mean) increase in the probability that a teenage female has had sexual intercourse in the last year.

We have clearly shown that there is a significant amount of "selection" that confounds estimates of the association between substance use and sexual behavior. This should give pause to researchers and policymakers who suggest that a reduction in substance use will significantly reduce risky sexual behavior and reduce teen pregnancy, out-of-wedlock birth and STDs. Indeed, our best guess is that substance use has no causal effect on sexual activity and risky sex.

*Appendix Table A.1*. Sample means of selected characteristics by sex and past 30 day alcohol use respondents age 15–17 in NLSY97.

| | Males alcohol use past 30 days | | Females alcohol use past 30 days | |
|---|---|---|---|---|
| Variable | Yes | No | Yes | No |
| Had sex past 12 months | 0.445** | 0.198 | 0.389** | 0.172 |
| Risky sex past 12 months | 0.176** | 0.064 | 0.187** | 0.063 |
| Black | 0.156** | 0.283 | 0.160** | 0.305 |
| Hispanic | 0.220 | 0.210 | 0.216 | 0.212 |
| Age (in months) | 192.5** | 190.5 | 191.5 | 191.1 |
| Height (in inches) | 68.94** | 68.63 | 64.41** | 64.22 |
| Weight (in pounds) | 155.9 | 153.8 | 130.87 | 132.35 |
| Poor health | 0.248 | 0.224 | 0.361** | 0.301 |
| Highest grade completed | 8.893** | 8.699 | 8.985** | 8.903 |
| Two bio. parents age 12 | 0.472 | 0.470 | 0.439 | 0.445 |
| Mother's education | 12.55 | 12.45 | 12.63** | 12.28 |
| Mother's age at 1st birth | 22.83 | 22.88 | 23.34** | 22.63 |
| Pct. fam. in poverty (county) | 10.49** | 11.18 | 10.23** | 11.21 |
| Destroyed property | 0.563** | 0.328 | 0.347** | 0.157 |
| Stole goods > \$50 | 0.573** | 0.328 | 0.489** | 0.238 |
| Peers attend church | 0.222** | 0.255 | 0.234 | 0.261 |
| Peers cut school | 0.432** | 0.352 | 0.510** | 0.449 |
| Peers play sport | 0.589 | 0.595 | 0.579 | 0.603 |
| Smoked cigarettes | 0.754** | 0.344 | 0.766** | 0.336 |
| Binge drinking past 30 days | 0.554** | 0.000 | 0.408** | 0.000 |
| Marijuana use past 30 days | 0.355** | 0.045 | 0.302** | 0.027 |
| Number of observations | 1163 | 2848 | 1118 | 2766 |

Notes:

1. Number of observations listed represents the maximum number available. For several variables the actual sample size is slightly less because of missing information.
2. For variables with significant amount (e.g., 10% of sample) of missing information, we assigned the sample mean and created a variable indicating that for this observation the information was missing. Variables falling into this category are: mother's education, mother's age at first birth, weight, and proportion of peers that attend church regularly.
3. Asterisks (**) next to a number indicates that the difference between drinkers and non-drinkers is statistically significant ($p < 0.05$).

*Appendix Table A.2*. Sample means of selected characteristics by sex and past 30 day marijuana use respondents age 15–17 in NLSY97.

| | Males marijuana use past 30 days | | Females marijuana use past 30 days | |
|---|---|---|---|---|
| Variable | Yes | No | Yes | No |
| Had sex past 12 months | 0.605** | 0.217 | 0.633** | 0.187 |
| Risky sex past 12 months | 0.253** | 0.071 | 0.328** | 0.072 |
| Black | 0.240 | 0.246 | 0.184** | 0.273 |
| Hispanic | 0.210 | 0.214 | 0.180 | 0.217 |
| Age (in months) | 193.00 | 190.76 | 191.98 | 191.09 |
| Height (in inches) | 68.84 | 68.70 | 64.48 | 64.25 |
| Weight (in pounds) | 152.91 | 154.69 | 133.48 | 131.75 |
| Poor health | 0.317** | 0.218 | 0.456** | 0.302 |
| Highest grade completed | 8.80 | 8.75 | 8.95 | 8.92 |
| Two bio. parents Age 12 | 0.332** | 0.492 | 0.342** | 0.456 |
| Mother's education | 12.36 | 12.49 | 12.62 | 12.36 |
| Mother's age at 1st birth | 22.47 | 22.93 | 22.94 | 22.82 |
| Pct. fam. in poverty (County) | 10.35 | 11.08 | 9.66 | 11.07 |
| Destroyed property | 0.682** | 0.351 | 0.490** | 0.178 |
| Stole goods > \$50 | 0.725** | 0.348 | 0.653** | 0.269 |
| Peers attend church | 0.204** | 0.252 | 0.164** | 0.264 |
| Peers cut school | 0.519** | 0.354 | 0.647** | 0.444 |
| Peers play sport | 0.549** | 0.601 | 0.494** | 0.608 |
| Smoked cigarettes | 0.876** | 0.398 | 0.942** | 0.403 |
| Binge drinking past 30 days | 0.570** | 0.096 | 0.498** | 0.072 |
| Number of observations | 542 | 3470 | 412 | 3478 |

Notes:

1. Number of observations listed represents the maximum number. For several variables the actual sample size is slightly less because of missing information.
2. For variables with significant amount (e.g., 10% of sample) of missing information, we assigned the sample mean and created a variable indicating that for this observation the information was missing. Variables falling into this category are: mother's education, mother's age at first birth, weight, and proportion of peers that attend church regularly.
3. Asterisks (**) next to a number indicates that the difference between drinkers and non-drinkers is statistically significant ($p < 0.05$).

*Appendix Table A.3.* Estimates of the effect of alcohol and marijuana use on the probability of having sexual intercourse in last 12 months respondents age 15–17 in NLSY97 — Lifetime alcohol use > 0.

| | Univariate probit $\rho=0$ | Constrained biv. probit $\rho=0.1$ | Constrained biv. probit $\rho=0.2$ | Constrained biv. probit $\rho=0.3$ | Constrained biv. probit $\rho=0.4$ | Constrained biv. probit $\rho=0.5$ | Bivariate probit $\rho=\frac{\mathrm{Cov}(X'\beta, X'\gamma)}{\mathrm{Var}(X'\gamma)}$ |
|---|---|---|---|---|---|---|---|
| *Males* | | | | | | | |
| Alcohol use past 30 days | 0.435** (0.065) | 0.272** (0.065) | 0.106 (0.064) | −0.063 (0.063) | −0.234** (0.062) | −0.407** (0.059) | 0.190** (0.065) [$\rho=0.15$] |
| Binge drinking past 30 days | 0.573** (0.070) | 0.404** (0.070) | 0.232** (0.069) | 0.058 (0.068) | −0.118 (0.066) | −0.296** (0.064) | 0.002 (0.067) [$\rho=0.33$] |
| Marijuana use past 30 days | 0.508** (0.078) | 0.335** (0.077) | 0.160** (0.076) | −0.017 (0.075) | −0.196** (0.073) | −0.376** (0.071) | −0.722** (0.063) [$\rho=0.69$] |
| *Females* | | | | | | | |
| Alcohol use past 30 days | 0.133** (0.065) | −0.029 (0.065) | −0.193** (0.064) | −0.356** (0.063) | −0.520** (0.062) | −0.683** (0.060) | −0.156** (0.065) [$\rho=0.18$] |
| Binge drinking past 30 days | 0.354** (0.077) | 0.181** (0.076) | 0.007 (0.076) | −0.168** (0.074) | −0.343** (0.072) | −0.518** (0.070) | −0.250** (0.073) [$\rho=0.35$] |
| Marijuana use past 30 days | 0.571** (0.084) | 0.395** (0.084) | 0.217** (0.083) | 0.037 (0.081) | −0.146 (0.079) | −0.330** (0.076) | −0.828** (0.064) [$\rho=0.77$] |
| Additional covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes:

1. $\rho = (\mathrm{Cov}(X'\beta, X'\gamma))/(\mathrm{Var}(X'\gamma))$ is calculated by a two-step procedure. An initial value for rho was chosen and then the bivariate probit model was estimated. Rho was then re-calculated using estimates of $X'\beta$ and $X'\gamma$ from the bivariate probit. This procedure was repeated until estimates of rho converged.
2. See notes to Table 2.

## Notes

1. This point is the subject of some disagreement. For example, see Hayes, 1987; Hotz, McElroy and Sanders, 1999.
2. These include problem-behavior theory (Jessor and Jessor, 1977), developmental-stage theory (Kandel 1989), social learning theory (Akers, 1977), self-derogation theory (Kaplan, 1975), life-course theory (Ensminger, Brown, and Kellman, 1982), and social control theory (Hirschi, 1969). Most of these theories define certain types of deviant behavior with respect to age and examine progressions from one type to another over the early life cycle. All of them consider determinants specific to the individual and his or her home, school, and community environment, but they place different emphasis on the importance of these broad classes of factors and variables within each class.
3. See Harvey and Beckman (1986); Yamaguchi and Kandel (1987); Mott and Haurin (1988); Elliot and Morse (1989); Kandel (1989); Biglan et al. (1990); Ensminger (1990); Hingson et al. (1990); Leigh (1990); Rosenbaum and Kandel (1990); Gold et al. (1991); Gold et al. (1992); Gold and Skinner (1992); Orr, Beiter, and Ingersoll (1991); Shafer and Boyer (1991); Mensch and Kandel

(1992); Strunin and Hingson (1992); Leigh (1993); Leigh and Stall (1993); Shafer et al. (1993); Cooper, Peirce, and Huselid (1994); Laumann et al. (1994); Lowry et al. (1994); Senf and Price (1994); Donovan and McEwan (1995); de Gaston, Jensen, and Weed (1995); Graves and Leigh (1995); Harvey and Spigner (1995); Fergusson and Lynskey (1996); Fortenberry et al. (1997); Jakobsen et al. (1997); Kaestner (1997, 1998); and Kowaleski-Jones and Mott (1998). This list excludes studies that describe sexual behavior among specific groups of substance abusers such as alcoholics and intravenous drug users.

4. See Harvey and Beckman (1986); Leigh (1990); Gold et al. (1991); Gold and Skinner (1992); Leigh (1993); Senf and Price (1994); Donovan and McEwan (1995); and Fortenberry et al. (1997) for exceptions.
5. If sex and substance use are complements in consumption, then policies that reduce alcohol consumption will also reduce risky sexual behavior. However, the efficiency of such policies depends on the strength of the cross-price effects. It may be more efficient to directly target the causes of risky sexual behavior.
6. See Ensminger (1990); Hingson et al. (1990); Orr, Beiter, and Ingersol (1991); Shafer and Boyer (1991); Strunin and Hingson (1992); Harvey and Spigner (1995); Fergusson and Lynskey (1996); and Jakobsen et al. (1997).
7. See Rees et al. (2001); Sen (2002); and Rashad and Kaestner (2004).
8. In addition, Rees et al. (2001) do not adjust the standard errors for possible within-state clustering, which can be particularly important in their case since the excluded instruments are measured at the state level and have only state variation. Rashad and Kaestner (2004) re-estimate the Rees et al. (2000) model correcting for this and show that the instruments are jointly significant at the 0.10 level.
9. The necessary conditions are: random selection of observed variables, large number of determinants (i.e., large W) of sexual activity, and independence of observed and unobserved variables.
10. Blacks and Hispanics are over sampled.
11. Note that respondents who turn 15 in 1999 will not have the two years of information necessary to carry out the FE estimation.
12. In addition to the results shown, we calculated the correlation coefficient ($\rho$) excluding age and race, which are strictly exogenous characteristics and not variables that can be ''selected''. The results were similar to those reported.
13. We also estimated similar models for a sample limited to those who have consumed alcohol sometime in their life. The results from these analyses are similar to those reported ands are in the appendix.
14. Estimates for a sample limited to drinkers are in the appendix (see note 11).
15. OLS was chosen because it produces consistent estimates, are easy to implement, and easy to interpret. As will be shown in Table 5, univariate probit regressions and OLS regressions yield very similar estimates of the effect of substance use on sexual behavior.
16. Blacks from well-educated families (with a parent with a college degree) and Hispanics were over sampled.

## References

Akers, Ronald L. (1977). ''Deviant Behavior: Social Learning Approach.'' (2nd Edn.), Belmont, California: Wadsworth.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. (2000). ''Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.'' National Bureau of Economic Research Working Paper No. 7831.

Barnes, Grace M., Michael P. Farrell, and Sarbani Banerjee. (1994). ''Family Influences on Alcohol Abuse and Other Problem Behaviors Among Black and White Adolescents in a General Population Sample.'' *Journal of Research on Adolescence* 4, 183–202.

Bach, Peter B., and John Lantos. (1999). "Methadone Dosing, Heroin Affordability, and the Severity of Addiction." *American Journal of Public Health* 89, 662–665.

Bearman, Peter S., Jo Jones, and J. Richard Udry. (1997). "The National Longitudinal Study of Adolescent Health: Research Design." Chapel Hill, North Carolina: University of North Carolina.

Biglan, Anthony et al. (1990). "Social and Behavioral Factors Associated with High-Risk Sexual Behavior among Adolescents." *Journal of Behavioral Medicine* 13, 245–261.

Brown, Sarah S. and Leon Eisenberg, eds. (1995). "The Best Intentions: Unintended Pregnancy and the Well-Being of Children and Families." Washington, DC: National Academy Press.

Browning, M. J. (1983). "Necessary and Sufficient Conditions for Conditional Cost Functions." *Econometrica* 51(3), 851–856.

Bureau of the Census, U.S. Department of Commerce. (1997). *Statistical Abstract of the United States 1997*. Washington, DC: U.S. Government Printing Office.

Centers for Disease Control and Prevention, U.S. Department of Health and Human Services (USDHHS). (1996). *AIDS Public Information Data Set*. Atlanta, Georgia: CDC.

Centers for Disease Control and Prevention, (USDHHS). (1997a). *Sexually Transmitted Disease Surveillance 1996*. Atlanta, Georgia: CDC.

Centers for Disease Control, USDHHS. (1997b). *Summary of Notifiable Diseases, United States 1996. Mortality and Morbidity Weekly Report* 45, 3–87.

Chamberlain, Gary. (1980). "Analysis of Covariance with Qualitative Data." *Review of Economic Studies* 47, 225–238.

Chesson, Harrell, Paul Harrison, and William J. Kassler. (forthcoming). "Sex Under the Influence: The Effect of Alcohol Policy on Sexually Transmitted Disease Rates in the U.S." *Journal of Law and Economics*.

Cooper M.Lynne, Robert S. Peirce, and Rebecca Farmer Huselid. (1994). "Substance Use and Sexual Risk Taking Among Black Adolescents and White Adolescents." *Health Psychology* 13, 251–262.

Cooper M.Lynne, J. B. Skinner, and William H. George. (1990). "Alcohol Use and Sexual Risk-Taking Among Adolescents: Methodological Approaches for Addressing Causal Issues." In Daniela Seminara, Ronald Ross Watson, and Albert Pawlowski, (eds.), *Alcohol, Immunomodulation, and AIDS*. New York: Alan R. Liss, 11–90.

Council of Europe. (1996). *Recent Demographic Developments in Europe*. Brussels: Council of Europe Publications.

Crowe, Leif C. and William H. George. (1989). "Alcohol and Human Sexuality: Review and Integration." *Psychological Bulletin* 105, 374–386.

Darroch, Jacqueline E., Jennifer Frost, Susheela Singh, and the Study Team. (2001). "Teenage Sexual and Reproductive Behavior in Developed Countries: Can More Progress Be Made?" *Occasional Report No. 3*, New York: Alan Guttmacher Institute.

de Gaston, Jacqueline F., Larry Jensen, and Stan Weed. (1995). "A Closer Look at Adolescent Sexual Activity." *Journal of Youth and Adolescence* 465–477.

Donovan, Catherine, and Robert McEwan. (1995). "A Review of the Literature Examining the Relationship between Alcohol Use and HIV-Related Sexual Risk-Taking in Young People." *Addiction* 90, 319–328.

Elliot, Delbert S. and Barbara J. Morse. (1989). "Delinquency and Drug Use as Risk Factors in Teenage Sexual Activity." *Youth and Society* 21, 32–60.

Eng, Thomas R. and William T. Butler, eds. (1997). "The Hidden Epidemic: Confronting Sexually Transmitted Diseases." Washington, DC: National Academy Press.

Ensminger, Margaret E. (1990). "Sexual Activity and Problem Behaviors among Black, Urban Adolescents." *Child Development* 60, 2032–2046.

Ensminger, Margaret E., C. Hendricks Brown, and Sheppard G. Kellam. (1982). "Sex Differences in Antecedents of Substance Use Among Adolescents." *Journal of Social Issues* 38, 25–42.

Fergusson, David M. and Michael T. Lynskey. (1996). "Alcohol Misuse and Adolescent Sexual Behaviors and Risk Taking." *Pediatrics* 98, 91–96.

Fleming, Douglas T. et al. (1997). "Herpes Simplex Virus Type 2 in the United States, 1976 to 1994." *New England Journal of Medicine* 337, 1105–1111.
Fortenberry, J.Dennis et al. (1997). "Sex under the Influence: A Diary Self-Report Study of Substance Use and Sexual Behavior among Adolescent Women." *Sexually Transmitted Diseases* 24, 313–321.
Gold, Ron S. and Michael J. Skinner. (1992). "Situational Factors and Thought Processes Associated with Unprotected Intercourse in Young Gay Men." *AIDS* 6, 1021–1030.
Gold, Ron S. et al. (1991). "Situational Factors and Thought Processes Associated with Unprotected Intercourse in Gay Men." *Psychology and Health* 5, 259–278.
Gold, Ron S. et al. (1992). "Situational Factors and Thought Processes Associated with Unprotected Intercourse in Heterosexual Students." *AIDS Care* 4, 305–323.
Grant, Bridget F. et al. (1991). "Prevalence of DSM-III-R Alcohol Abuse and Dependence." *Alcohol Health & Research World* 15, 91–96.
Graves, Karen L. and Barbara L. Leigh. (1995). "The Relationship of Substance Use to Sexual Activity among Young Adults in the United States." *Family Planning Perspectives* 27, 18–33.
Harvey, S.Marie and Linda J. Beckman. (1986). "Alcohol Consumption, Female Sexual Behavior and Contraceptive Use." *Journal of Studies on Alcohol* 47, 327–332.
Harvey, S.Marie and Clarence Spigner. (1995). "Factors Associated with Sexual Behavior among Adolescents: A Multivariate Analysis." *Adolescence* 30, 253–264.
Hayes, Cheryl. (1987). "Risking the Future: Adolescent Sexuality, Pregnancy and Childbearing Vol. 10." Washington, DC: National Academy Press.
Henshaw, Stanley K. (1998). "Unintended Pregnancy in the United States." *Family Planning Perspectives* 30, 24–29, 46.
Hingson, Ralph W. et al. (1990). "Beliefs about AIDS, Use of Alcohol and Drugs, and Unprotected Sex among Massachusetts Adolescents." *American Journal of Public Health* 80, 295–299.
Hirschi, Travis. (1969). "Causes of Delinquency." Berkeley, California: University of California Press.
Hotz, Joseph V., Susan McElroy, and Seth Sanders. (1999). "Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Very Natural Experiment," NBER Working Paper No. 7397.
Jakobsen, Reidar et al. (1997). "Noncoital Sexual Interactions and Problem Behavior among Young Adolescents: The Norwegian Longitudinal Health Behaviour Study." *Journal of Adolescence* 20, 71–83.
Jessor, *Richard and Shirley L. Jessor*. (1977). "Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth." New York: Academic Press.
Kaestner, Robert. (1997). "The Effects of Cocaine and Marijuana Use on Marriage and Marital Stability." *Journal of Family Issues* 18(2), 145–173.
Kaestner, Robert. (1998). "Drug Use, Culture, and Welfare Incentives: Correlates of Family Structure and Out-of-Wedlock Birth." *Eastern Economic Journal* 24, 395–416.
Kandel, Denise. (1989). "Issues of Sequencing of Adolescent Drug Use and Other Problem Behaviors." *Journal of Drug Issues* 53, 55–76.
Kaplan, Howard B. (1975). "Self-Attitudes and Deviant Behavior." Goodyear, Pacific Palisades: CA.
Kaplan, Howard B. (1980). "Deviant Behavior in Defense of Self." New York: Academic Press.
Kowaleski-Jones, Lori and Frank L. Mott. (1998). "Sex, Contraception and Childbearing among High-Risk Youth: Do Different Factors Influence Males and Females?" *Family Planning Perspectives* 30, 163–169.
Laumann, Edward O. et al. (1994). "The Social Organization of Sexuality: Sexual Practice in the United States." Chicago: University of Chicago Press.
Leigh, Barbara C. (1990). "Alcohol and Unsafe Sex: An Overview of Research and Theory." *Alcohol, Immunomodulation, and AIDS* 35–46.
Leigh, Barbara C. (1993). "Alcohol Consumption and Sexual Activity as Reported with a Diary Technique." *Journal of Abnormal Psychology* 102, 490–493.
Leigh, Barbara C. and Ron Stall. (1993). "Substance Use and Risky Sexual Behavior for Exposure to HIV: Issues in Methodology, Interpretation, and Prevention." *American Psychologist* 48, 1035–1045.

Lowry, Richard et al. (1994). "Substance Use and HIV-Related Sexual Behaviors among US High School Students: Are They Related?" *American Journal of Public Health* 84, 1116–1120.
Mensch, Barbara and Denise B. Kandel. (1992). "Drug Use as a Risk Factor of Premarital Teen Pregnancy and Abortion in a National Sample of Young White Women." *Demography* 29, 409–429.
Mott, Frank L. and R. Jean Haurin. (1988). "Linkages between Sexual Activity and Alcohol and Drug Use among American Adolescents." *Family Planning Perspectives* 20, 128–136.
Orr, Donald P., Mary Beiter, and Gary Ingersoll. (1991). "Premature Sexual Activity as an Indicator of Psychosocial Risk." *Pediatrics* 87, 141–147.
Pollak, Robert A. (1969). "Conditional Demand Functions and Consumption Theory." *Quarterly Journal of Economics* 83(1), 60–78.
Public Health Service, USDHHS. (1995). *Healthy People 2000 Midcourse Review and 1995 Revisions*. Washington, DC: US Government Printing Office.
Rashad, Inas, and Robert Kaestner. (2004). "Teenage Sex, Drugs and Alcohol: Problems Identifying the Cause of Risky Behaviors." *Journal of Health Economics* 23(3), 493–503.
Rees, Daniel I., Laura M. Argys, and Susan Averett. (2001). "New Evidence on the Relationship Between Substance Use and Adolescent Sexual Behavior." *Journal of Health Economics* 20, 835–845.
Rosenbaum, Emily and Denise B. Kandel. (1990). "Early Onset of Adolescent Sexual Behavior and Drug Involvement." *Journal of Marriage and the Family* 52, 783–798.
Sen, Bisakha. (2002). "Does Alcohol-Use Increase the Risk of Sexual Intercourse Among Adolescents? Evidence From the NLSY97." *Journal of Health Economics* 21, 1085–1093.
Senf, Janet H. and Carol Q. Price. (1994). "Young Adults, Alcohol and Condom Use: What Is the Connection?" *Journal of Adolescent Health* 15, 238–244.
Shafer, Mary-Ann and Cherrie B. Boyer. (1991)."Psychosocial and Behavioral Factors Associated with Risk of Sexually Transmitted Diseases, Including Human Immunodeficiency Virus Infection among Urban High School Students." *Journal of Pediatrics* 119, 826–833.
Shafer, Mary-Ann et al. (1993). "Relationship between Drug Use and Sexual Behaviors and the Occurrence of Sexually Transmitted Diseases among High-Risk Male Youth." *Sexually Transmitted Diseases* 30, 307–313.
Sonenstein, Freya L. et al. (1998). "Changes in Sexual Behavior and Condom Use among Teenage Males: 1988 to 1995." *American Journal of Public Health* 88, 956–959.
Strunin, Lee and Ralph Hingson. (1992). "Alcohol, Drugs, and Adolescent Sexual Behavior." *International Journal of Addictions* 27, 129–146.
Turner, Charles F. et al. (1998). "Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology." *Science* 280, 867–873.
Ventura, Stephanie J., Sally C. Curtin, and T.J. Mathews. (1998). "Teenage Births in the United States: National and State Trends, 1990– 96." National Vital Statistics System. Hyattsville, Maryland: National Center for Health Statistics.
Ventura, Stephanie J. and Christine A. Bachrach. (2000). "Nonmarital Childbearing in the United States, 1940-99." *National Vital Statistics Reports* 48, 16. Hyattsville, MD: National Center for Health Statistics.
Yamaguchi, Kazuo and Denise B. Kandel. (1987). "Drug Use and Other Determinants of Premarital Pregnancy and Its Outcome." *Journal of Marriage and the Family* 49, 257–270.
Zabin, Laurie Schwab and Sarah C. Hayward. (1994). "Adolescent Sexual Behavior and Childbearing." New York: Sage Publications.