SUBSTANCE USE, RISKY SEXUAL BEHAVIOR, AND EMPLOYMENT AMONG YOUNG PEOPLE

Courtney McCole Wicher Keeler

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Public Health.

Chapel Hill
2012

Approved by:

Marisa Domino, Ph.D.

Gary Cuddeback, M.S.W., M.P.H., Ph.D.

Carolyn Halpern, Ph.D.

James Porto, Jr., Ph.D.

Sally Stearns, Ph.D.

ABSTRACT

COURTNEY KEELER: Substance use, risky sexual behavior, and employment among young people
(Under the direction of Marisa Domino)

Using data from the National Longitudinal Study of Adolescent Health, this work examines the impact of individual substance use, peer substance use, and depression on risky sexual behaviors, rape victimization among women, and labor market outcomes. The data are nationally representative of American youth. Although Fagan (1993) hypothesized that substance use not only increases the probability of perpetrating violent crimes but also the probability of becoming a victim of violent crime, the impact of substance use and depression on rape victimization remains largely uninvestigated. Previous research often neglects the concurrent impact of depression and the role of peer substance use in shaping the outcomes of interest. I fill these gaps by controlling for individual and peer substance use as well as depression. I use zero-inflated negative binomial, linear probability, and two-part models to investigate these relationships. Given the endogeneity of depression and substance use, analyses incorporate instrumental variable approaches. The results suggest that neither substance use nor depression have a causal impact on the risky sexual behavior, rape, or labor market outcomes. The analyses do indicate, however, that peer substance use influences the observed health and employment outcomes. As a result, health providers may want to consider a patient's social environment when devising prevention and treatment plans

ACKNOWLEDGEMENTS

The dissertation process has been one of the most inspiring and difficult experiences of my life, and I could not have done this alone. First and foremost, I would like to thank my advisor, Dr. Marisa Domino. She has supported and nurtured my intellectual interests, read countless drafts, and helped shape my initial idea into this final product. I came into the Health Policy and Management Ph.D. program as a somewhat timid individual. Under Marisa's guidance, I have grown into a much stronger woman and finally found my own voice. I would also like to thank Dr. Gary Cuddeback, Dr. Carolyn Halpern, Dr. James Porto, Jr., and Dr. Sally Stearns who have guided me throughout this process and from whom I have learned so much.

My friends and family have been so very supportive the last few years. Thank you for giving me a better prospective on things and always reminding me that there is a light at the end of the tunnel. Most importantly, I would like to thank my husband, Daniel, for accompanying me on this journey and being proud of me every step of the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Many adolescents and young adults experiment with substances sometime during their youth. Data from the National Household Survey of Drug Abuse suggest that in 2001, substance use was widespread, with 20.8% of adolescents and 31.9% of young adults reporting some illicit drug use in the past year. Nevertheless, substance use, and accompanying worries regarding self-medication and dependence, can be dangerous if not lethal during these formative years.

Peer substance use behavior closely interacts with an individual's own health and employment outcomes. While substance use has physical effects on individuals, researchers and policymakers often overlook the externalities associated with the consumption of alcohol and drugs. Substance use not only directly and indirectly impacts a consumer's health, but individual consumption potentially generates negative externalities that can affect the health of a third party. As a result, peer substance use potentially impacts individual health and other behavioral outcomes. The impact of peer substance use is central to these analyses.

Sexual behaviors and labor market outcomes can be defined in a variety of ways. In this work, the sexual outcomes of interest include the number of sexual partners, sexually transmitted disease status, and being a victim of rape; the labor market outcomes of interest include employment status and wage rate.

Among young adults, sexual activity presents a major public health challenge. Sexually transmitted diseases are particularly concerning, especially for younger age groups. The 2009 Sexually Transmitted Disease Surveillance 2009 report, published by the Division of STD Prevention within the Centers for Disease Control, details STD rates within the United States population. Although the overall rates of gonorrhea and syphilis continued to decline between 2005 and 2009, chlamydia rates increased during the same time period (37.6% for men and 20.3% for women); moreover, the report illustrates the relatively high risk of sexually transmitted disease facing adolescents and young adults compared to other groups. This risk is particularly acute for females. In 2009, the age-specific rates of chlamydia and gonorrhea were highest among women aged 15–19 and 20–14 years and men aged 20–24 years. The age-specific rates of primary and secondary syphilis were also highest among individuals ages 20-24 and 24–29. These statistics suggest that sexual intercourse itself is a risky undertaking for adolescents and young adults, especially if individuals use no protective measures.

Rape also threatens health, placing individuals in danger of serious psychological distress and, for females, unplanned pregnancy. According to national statistics published by the Centers for Disease Control, 11% of high school females and 4% of high school males have reported being raped. Similar statistics suggest that roughly 20% to 25% of women are raped or experience a rape attempt during college (CDC, 2007).

At first blush, the link between sexual behavior and labor market experiences appears tenuous. Risky sexual behaviors are a health outcome; labor market behaviors are not. Nevertheless, both risky sexual behavior and employment present major hurdles confronting

adolescents and young adults. While risky sexual behavior is a major public health challenge for young people, adolescents and young adults also face issues of unemployment and underemployment as they progress into adulthood, particularly in tough economics conditions.

Substance use can also affect labor market outcomes. The Bureau of Labor Statistics reported in February 2010 that the national unemployment rate in the United States was 9.7%. Given the current recessionary atmosphere, predictors of positive labor market outcomes are important. What is more, unemployment itself may adversely affect health outcomes, as most individuals acquire health insurance through their employers, although this may become less of an issue under health care reform.

While researchers have explored the relationships among substance use, sexual behavior, and labor market outcomes in separate pieces, the findings from these studies are far from conclusive. While some researchers find evidence that substance use is significantly associated with sexual behavior and labor market outcomes, others do not. The direction of these relationships remains unclear. A fresh approach is needed.

Previous research has largely neglected the direct influence of depression and peer behavior in studies examining the impact of substance use on sexual risk-taking and employment; these studies also ignore the possible role that depression and peer behavior play in mediating the relationship between substance use and the outcomes of interest.

Both existing literature and intuition justify the inclusion of depression and peer substance use indicators when modeling substance use. In omitting peer substance use and indicators of depression from the analysis, substance use variables suffer from omitted

variable bias. As a result, the true relationship between substance use and the outcome of interest are partially masked in the existing work.

This work incorporates measures of depression and peer behavior. First, I control for depressive symptoms. Second, I control for individual (or own) substance use and peer substance use. While important, the impact of peer behavior is difficult to assess. To do so, I rely on social interaction theory (Akerlof, 1997) and models of status-seeking and conformist behavior (Akerlof, 1997).

Depression and individual substance use variables are potentially endogenous since the error term contains unobserved personal characteristics and other behaviors that are likely correlated with these variables. Endogeneity may also be driven by reverse causality (i.e., sexual risk-taking and observed labor market drive observed substance use or depression outcomes rather than the other way around). By identifying variation in the endogenous right-hand side variable separately from the error term, instrumental variable approaches help establish causality.

Three separate aims are addressed in this study. Crucially for all analyses, I control for depressive symptoms and peer substance use:

1. Estimate the causal effect of substance use on sexual activity. This aim uses two dependent variables: (a) reported number of sexual partners and (b) reported sexually transmitted disease status in the year prior;

2. Estimate the causal effect of substance use on the probability of a woman reporting having ever been raped at the time of the interview; and

3. Estimate the causal effect of substance use during early adulthood on later labor market outcomes.

I hypothesize that substance use, depression, and peer substance use are positively associated with reported the number of partners, one-year STD status, and rape, as well as negatively associated labor market outcomes

The data for this analysis come from the National Longitudinal Study of Adolescent Health (hereafter referred to as Add Health). Add Health surveyed respondents from early adolescence into adulthood. The first wave of Add Health was administered in 1994. The latest and fourth wave occurred between 2007 and 2008.

II. BACKGROUND AND SIGNIFICANCE

I begin this section by discussing the prevalence of substance use and mental health conditions, depression in particular, among adolescents and young adults. Next, I survey the existing literature, investigating risky sexual behaviors, labor market outcomes, depression, and substance use. I end with a discussion of several economic models germane to this work, including rational addiction models and social interaction theory.

A. Prevalence of substance use in adolescent and young adult populations

Substance use has taken a backseat in public health research. As Steven Schroeder, a distinguished professor of Health and Health Care at the University of California at San Francisco, noted, "Despite their huge health toll, substance abuse disorders remain underappreciated and underfunded." (Schroeder, 2005) Nevertheless, this field of research is important, not only because of the direct effect of substance use on an individual's health, but also because of the indirect and often ignored external consequences.

The National Survey on Drug Use and Health (NSDUH) allows researchers to examine the prevalence of substance use among adolescents and young adults. NSDUH is administered at the household level. Institutionalized youths in juvenile detention centers and other custodial settings are not interviewed. As a result, this survey does not represent adolescents with severe mental illness, those receiving inpatient patient care, or youth in detention facilities.

Data from the 2010 NSDUH confirm the relatively high risk of substance use and abuse among adolescents and young adults. NSDUH estimates that, in 2010, 22.6 million Americans could be identified as current illicit drug users.[1] Marijuana was the most popular illicit drug. Rates of current illicit drug use are estimated to be highest among individuals 18 to 20 followed by individuals 21 to 25. Similar statistics relating to binge drinking suggest that heavy drinking rates are highest among individuals 21 to 25 (SAMHSA, 2011).

B. Co-occurrence of substance use and mental health conditions in adolescent and young adult populations

The prevalence of substance use and mental health disorders is high and often co-occurring. Depression is one of the most common mental health disorders in the United States. The National Co-morbidity Survey Replication (NCS-R), which is nationally representative of the U.S. population, found that the 12-month prevalence of major depressive disorder was 7%, surpassed only by specific and social phobias. The lifetime prevalence of mood disorders and substance disorders is also quite high, at 21% and 15%, respectively (NIMH).

Taking the findings from the NCS-R, the National Epidemiologic Study of Alcohol and Related Conditions, and the National Survey on Drug Use and Health together, a 2007 report by SAMHSA'S Co-Occurring Center for Excellence estimated that more than 9% of adults surveyed had a substance use disorder, 9% had a diagnosable mood disorder, and more than five million had a co-occurring substance use disorder and serious mental illness (CSAT). Research shows that the co-occurring substance use disorder often results from an

[1] The NDSUH classifies an individual as a current illicit drug user if "the had used an illicit drug during the month prior to the survey interview."

existing mental disorder, often anxiety or bipolar mood disorders, which begins in adolescence (NIMH).

Substance use may function as a form of self-medication. The self-medication hypothesis suggests that individuals will use nicotine, alcohol, and other drugs to alleviate undesirable affect states (Khantzian, 1985; Khantzian 1997). The emotional and psychological ramifications of depression predispose individuals towards substance use. According to Khantzian (1997), "clinical observations and empirical studies that focus on painful affect and subjective states of distress more consistently suggest that such states of suffering are important psychological determinants in using, becoming dependent upon, and relapsing to addictive substances."

While the above statistics relate to adults, the association between substance abuse and mental health, and depression in particular, also holds for adolescents and young adults. In 2005, 8.8% of youths 12 to 17 in the National Survey of Drug Use and Health (NSDUH) had a major depressive episode (MDE) (SAMHSA, 2007a). The NSDUH shows that youths are twice as likely to initiate alcohol use if they had a MDE in the last year relative to youths who did not (29.2% versus 14.5%). These youths are also twice as likely to initiate illicit drug use (16.1% versus 6.9%).

The link between substance use and depression continues into adulthood. During 2005 and 2006, NSDUH data found that 33.7% of young adults aged 18 to 25 that experienced an MDE within the last year initiated alcohol use relative to 24.8% of young adults who did not experience such an episode. Twelve percent of young adults with a MDE initiated illicit drug use relative to 5.8% of other individuals. The most popular illicit substance was marijuana, followed by cocaine. In each case, consumption of the drug was

substantially higher in the MDE group (SAMHSA, 2007b). Related statistics suggest that young males aged 18 to 25 who experienced serious psychological distress over the past year are more likely to engage in binge drinking and illicit drug use (SAMHSA, 2006).

C. Substance use as a cause of mental illness

Although many articles discuss the association of alcohol and drug use with mental health disorders, the discussion about the directionality of this relationship is limited. Much of the existing literature surrounds cannabis use, schizophrenia, and psychosis in adolescents and young adults. Research suggests that cannabis use may be related to schizophrenia (Arseneault et al., 2004; Andreasson et al., 1987), poor mental health (van Ours and Williams, 2010; Fergusson and Horwood, 1997), and psychosis (Fergusson et al., 2005; Henquet et al., 2004). Moreover, results indicate a "dose-response relationship" among cannabis use and poor mental health (van Ours and Williams, 2010) and psychosis (Henquet et al., 2004).

The validity of the statistical methodology used in these studies and the underlying mechanism defining the causal link are weak. Many unobserved and omitted variables confound the estimation process. Some researchers fail to use appropriate statistical procedures to establish causality. For example, many studies do not use instrumental variables or dynamic modeling techniques in their statistical analyses (Henquet et al., 2004; Fergusson and Horwood, 1997; Andreasson et al., 1987). Therefore, even though mental health status clearly plays a mediating role in defining the relationship between substance use and associated behaviors, the mechanism remains unclear. Indeed, even the directionality of the relationship between mental health and substance use is ambiguous.

D. Consequences of substance use

Research demonstrates the relationship between substance use and adverse sexual outcomes and other negative behaviors. As French and colleagues (2000) discuss, alcohol and illicit drug use have been linked to a series of undesired outcomes, such as unprotected sexual activity, needle exchange (Gunn et al. 1995; Chirgwin et al. 1991; Cottler et al., 1990; Rolfs et al., 1990) and sexually transmitted diseases (Cheeson et al., 2000), including HIV (Allen et al., 1992; Bruenau et al., 1997; Kral et al., 1998). Sexually transmitted disease presents a particular health risk for young people. For many sexually transmitted diseases, including chlamydia, gonorrhea, and syphilis, adolescents and young adults between the ages of 15 and 30 have the highest age-specific diagnosis rates compared to other age groups (CDC, 2010). The high prevalence of STDs within this age group enhances the relative risks associated with sexual activity.

Substance use has also been tied to aggressive and otherwise reckless behaviors, including abusive behaviors aimed at wives (Markowitz, 2000) and children (Markowitz and Grossman, 2000) and driving under the influence (Phelps, 1987). The relationship with driving under the influence has been found to be particularly strong for young people (Phelps, 1987; Peck et al., 2008).

Automobile accidents resulting from high blood alcohol concentration are perhaps the most common type of fatality resulting from alcohol consumption. In 2008, the number of deaths from alcohol-impaired driving crashes was estimated to be 1.25 fatalities per 100 million vehicle miles traveled (NHTSA, 2009). In total, 11,773 individuals died from alcohol-impaired crashes in the same year (NHTSA, 2008).

In addition, researchers have observed a more general relationship between substance use and criminal behavior. The 1997 National Longitudinal Survey of Youth highlights the connection between delinquency and substance use. Substance users experience higher proportions of delinquent behavior, the sale of drugs, days of suspension, and vandalism. Similar statistics hold for more dramatic behaviors, including major theft, attack/assault, and carrying a handgun (McCurley and Snyder, 2008).

Finally, substance use has also been shown to increase health expenditures. French and colleagues (2000) find that chronic drug users and injecting drug users experience significantly more inpatient and emergency care and less outpatient care relative to non-drug users; however, the magnitude of the overall effect was small. Being a chronic drug user is associated with 0.08 more inpatient visits, 0.24 more emergency room visits, and 0.16 fewer outpatient visits compared to non-drug users over approximately a one-year period. Given the relative expense of emergency and inpatient care, it is not surprising that, on average, chronic and injecting drug users spend $1,000 more in total health care costs than non-drug users. The sample population for this study consists of 1,570 respondents who filled out a self-reported questionnaire collected between 1996 and 1997. While the data may be nationally representative of low-income individuals with substance use disorders, the results may not be nationally representative of substance users as a whole.

D.1. The relationship between substance use and risky sexual behaviors

Unprotected sexual behavior is characterized by a lack of contraception. Fontanet and colleagues (1998) define an unprotected sex act as any sexual act in which condoms are not used or sexual acts in which a condom tore or slipped in or out. While birth control does protect against unplanned pregnancy, its use alone does not protect against all risks

associated with sexual intercourse – namely sexually transmitted diseases, which are a central concern for this age group (CDC, 2010).

Substance use and poor mental health may instigate risky sexual behaviors. The current literature focuses almost exclusively on alcohol and marijuana use, leaving the impact of harder drug use on sexual risk-taking largely unknown. Mental health is often unaddressed by these analyses. Of the work described in the next section, only DeSimone (2010) incorporates any controls for mental health status. DeSimone uses measures of perceived weight and whether or not an individual planned a suicide in the past year as proxies for depression in examining the impact of binge drinking on risky sexual behavior among college students; nevertheless, his proxies lack a clinical foundation and may not even reflect depressive tendencies.

The debate continues as to whether a link between risky sexual behavior and substance use exists. While some researchers find evidence associating heavy alcohol and marijuana use with increased sexual activity among young adults (Cooper et al., 1994), others find little evidence of such a link (Rees et al., 2001). Given the high prevalence of sexually transmitted disease within this age group, an increase in sexual activity presents a risk in and of itself.

Investigating the impact of alcohol and marijuana use on a variety of risky sexual behaviors, some studies have concluded that a causal relationship between substance use and risky behaviors is unlikely (Grossman, Kaestner, and Markowitz, 2004; Grossman and Markowitz, 2005). Nevertheless, some literature suggests that the impact of alcohol use on the probability of engaging in sexual intercourse depends on how much alcohol one consumes (e.g., Sen, 2002). Among college students, binge drinking has been linked to

measures of risky sexual behavior, including sex with multiple partners, both with and without a condom (DeSimone, 2010).

These studies are far from consistent in their definitions of substance use and sexual risk-taking. Researchers have relied on varying measures of "heavy" alcohol and marijuana use as well as different definitions of sexual activity and "unprotected" sexual intercourse.

Research in this area has also neglects the potential impact of substance use on the probability of being raped or the probability of raping someone else. Fagan (1993) argues that substance use increases an individual's likelihood of being a victim of violent crime. Carpenter and Dobkin (2010) suggest that the pharmacological effects of alcohol may increase the likelihood of victimization, proposing that copious alcohol consumption produces sedative effects, making individuals at higher risk for attack.

D.1.a. Possible mechanisms through which substance use influences risky sexual behaviors

Several mechanisms linking substance use and risky behavior have been proposed. Some researchers hypothesize that alcohol and marijuana use may increase the likelihood of these risky behaviors by increasing sexual aggression, lowering inhibitions, and/or diminishing an individual's ability to assess risk (Rees et al., 2001).

The durational effects of this mechanism remain understudied. Most of the work in risky behaviors focuses on short-term consequences of substance use. Some researchers have found a negative association between past substance use and current employment (discussed later in this chapter), which suggests that the effects of substance use may linger over several years. Nevertheless, the durational effects of substance use in the context of risky sexual behavior remain uncertain.

Markowitz (2000) hypothesizes an alternative mechanism. Individuals may use alcohol or drugs to remove responsibility for otherwise unacceptable behavior (Markowitz, 2000). Carpenter and Dobkin (2010) refer to this as the "excuse motive." Under this hypothesis, substance use and risky behaviors are associated, but substance use does not cause risky sexual behaviors. Markowitz (2000) discusses how individuals may use alcohol consumption to diminish the personal blame incurred from spousal abuse.

Grossman, Kaestner, and Markowitz (2004) and Grossman and Markowitz (2005) argue that Jessor and Jessor's (1977) problem behavior theory may be a key piece of the puzzle. The theory suggests that a common third variable causes both substance use and risky sexual behavior. Grossman and Markowitz (2005) give the example of the tendency towards thrill seeking as such a variable. From a statistical standpoint, excluding these variables from the analysis would produce omitted variable bias. This is one reason why Grossman, Kaestner, and Markowitz (2004) as well as Grossman and Markowitz (2005) use person-level fixed effects analysis in their work.

In summary, the mechanism relating substance use and risky sexual behavior remains ambiguous. While researchers have developed several explanations linking substance use and sexual risk-taking, the precise mechanism by which alcohol and drug use influences riskiness remains uncertain and likely varies by individual.

D.1.b. Risky sexual behaviors and mental health

Researchers often omit mental health indicators from analyses of substance use and risky sexual behavior. Although it is difficult to unravel the individual impact of mental health and substance use, mental health does play an important role in shaping sexual risk-taking. Research suggests that, among young people, stress and depressed mood increase the

probability of having sex without a condom (Brown et al., 2006; Brooks et al., 2002). Emotional distress, including depression, may predispose an individual towards a variety of risky sexual behaviors, including unprotected sex and unplanned pregnancy (Kirby, 2002). In addition to unprotected sexual activity and unplanned pregnancy, the number of sexual partners is positively correlated with psychological disturbance (Tubman et al., 1996).

D.1.c. Rape and mental health

While research suggests that substance use and victimization are associated (Fagan, 1993), I have found no work suggesting that depression would lead to rape victimization. In fact, examining the timing of events, one would except that victimization would lead to depression, not the other way around. Frank and Stewart (1984) provide the following advice, "Clinicians and researchers as well should be alert to the presence of post-rape depression and tailor their interventions and research strategies accordingly." Nevertheless, depression is likely correlated with substance use and other observable and unobservable factors that influence an individual's likelihood of victimization. As a result, if depression does impact the incidence of rape, excluding it from an analysis of substance use and rape may generate omitted variable bias.

D.2. The relationship between substance use and labor market outcomes

Substance use during youth may have long-term consequences, potentially having a demonstrable effect on future labor market outcomes. Many researchers have analyzed this relationship, employing various definitions of substance use and labor market outcomes. Some researchers find evidence of a negative effect of substance use on labor market outcomes (e.g., DeSimone, 2002; French et al., 2001; Buchmuellar and Zuekas, 1998) while others have not (e.g., van Ours, 2006; MacDonald and Pundey, 2000).

The discrepancies in the direction of the association between substance use and labor market outcomes could result from several factors: differences in labor market outcomes studied, differences in instrumental variables used, age differences among sample populations, confusion between substance use and substance abuse, varying definitions of substance use and the frequency/level of use, different lagged periods of substance use, the sporadic inclusion of gender and human capital measures, and the absence of indicators of mental health and peer behavior.

D.2.a. Differences in labor market outcomes studied

Different definitions of labor supply may result in inconsistent findings (MacDonald and Pundey, 2000). Labor market outcomes previously studied in the context of substance use include hours worked (Zarkin et al., 1998), unemployment (MacDonald and Pundey, 2000), productivity and workplace achievement as measured by occupational class (MacDonald and Pundey, 2000), labor force participation (French et al., 2001), and employment. Employment classification spans a spectrum of values, from having worked at least one hour in the last year (DeSimone, 2002) to being employed at the time of the interview (French et al., 2001) to having a full-time job (van Ours, 2006).

D.2.b. Age differences in the sample population

Although the choice of age range may be arbitrary and determined by data availability, sample size concerns, or similar considerations, the age range likely impacts the findings, in that the impact of substance use on labor market outcomes varies by age. In fact, some inconsistencies in findings across studies may result from dramatic differences in the sample populations (Kandel et al., 1993).

In most cases, substance use is defined as current substance use, meaning the durational effects of substance use are under-investigated. Limited research investigates the impact of age of onset or the durational effects of substance use on employment outcomes (e.g., van Ours, 2006).

Studies of rational addiction demonstrate that younger and older individuals approach substance use quite differently.[2] As a result, the effect of substance use on employment is likely to vary across age groups. Studies investigating the impact of current and, to a more limited extent, prior substance use on employment for younger populations may find differing results than studies investigating the same relationship in older populations. In studies relying on broad age ranges, the opposing effects of substance use for different age groups may cancel each other out completely.

Researchers have studied the relationship between substance use and employment for a variety of age groups. For instance, some of the previous age ranges studied include 14 to 22 (DeSimone, 2002), 18 to 24 (Zarkin et al., 1998), 26 to 50 (van Ours, 2006), and 26 to 59 (French et al., 2001).

Given that substance use is likely to have a distinctive impact on employment for different age groups, studies with smaller age ranges will likely paint a more realistic and accurate picture of the impact of substance use on employment for that particular group.

[2] A concept discussed later in this section, rational addiction is an economic theory that hypothesizes that substance abuse is the result of a utility maximization process.

D.2.c. Varying definitions of substance use

Many researchers have relied on broad definitions of substance use.[3] In this work, however, I chose to define substance use narrowly, within unique substance categories (e.g., marijuana use), rather than broad categories (e.g., "soft drugs" in general).

The pharmacological and behavioral effects of substance use vary by substance (Carpenter and Dobkin, 2010), suggesting that a one-size-fits-all approach may be inappropriate. Surveying the literature describing the pharmacological and behavioral effects of substance use, Carpenter and Dobkin (2010) compare the effects of alcohol and illicit drugs. They conclude that the effects of alcohol are most similar to those of cocaine. Both increase aggression and irritability and decrease self-control. The authors discuss how amphetamines increase aggression and produce a paranoid psychotic state. Marijuana and opiate use are negatively associated with aggression/hostility; however, opiates, unlike marijuana, result in elevated levels of aggression during withdrawal. Given that the pharmacological and behavioral effects produced by substance use are likely to vary, the pathway by which use affects behavior will also likely differ.

D.2.d. Varying measures of the frequency and level of substance use

A limited number of studies analyze the frequency and level of substance use. In many cases, researchers rely on a binary measure of substance use (e.g., DeSimone, 2002; MacDonald and Pundey, 2000). The effect of any substance use on employment, however, is likely to be distinct from the level or frequency of substance use. For instance, the impact of marijuana use on employment may be minimal for recreational marijuana users. For daily users, however, the impact of marijuana use may be more pronounced. Binary measures of

[3] In the literature, substance use is often defined broadly as "illicit drug use" (e.g., Zarkin, 1998; French et al., 2001) or categorized as "hard" and "soft" drug use (e.g., MacDonald and Pundey, 2000).

drug use glaze over these nuances. Fortunately, some researchers measure the frequency of drug use, incorporating measures of "chronic" use (e.g., French et al., 2001; Buchmuellar and Zuvekas, 1998). Evidence suggests that chronic use impacts employment (e.g., French et al., 2001; Buchmuellar and Zuvekas, 1998).

D.2.e. Lagged effect of substance use

While studies have examined the lagged effect of substance use on employment, the lag period examined is often short. Although researchers examine the impact of lagged substance use over a single year period (Gills and Michaels, 1992; MacDonald and Pundey, 2000), alcohol and drug consumption may have a more long-term effect. Little evidence of an association between the age of onset of substance use and the probability of having a full time job has been found in the literature (van Ours, 2006), and more research is needed in the area, given the dearth of studies investigating the durational effects of substance use on employment.

D.2.f. Gender roles

Gender plays a crucial role in shaping both substance use and employment behaviors. The National Epidemiologic Survey on Alcohol and Related Conditions finds that 13.8% of men relative to 7.1% of women fit the DSM-IV criteria for any drug abuse disorder (Back et al., 2007). Men and women deviate in expected career trajectories as well. Not only are women less likely to participate in the labor force, but also a wage differential still exists between gender groups with similar levels of human capital attainment (Frank and Bernanke, 2004, pg. 334).

Because individual substance use and employment behavior themselves are different for men and women, the impact of substance use on employment is also likely to vary along

gender lines. Although French and colleagues (2001) find that chronic use negatively affects employment for both sexes, the magnitude of these effects differs across gender groups. DeSimone (2002) finds that cocaine and marijuana have a negative impact on the probability of employment for males but not for females. Kaestner's (1994a) work further demonstrates that cocaine and marijuana affect men and women differently. Together, these results reinforce the importance of incorporating gender into analysis of substance use and employment.

D.2.g. Human capital

Human capital refers to the skills that an individual acquires through training, often through education. An employee brings human capital to a firm via her skill set; thus, human capital attainment directly influences current employment opportunities. Although substance use may affect labor market outcomes directly, alcohol and drug use may indirectly affect employment behaviors by altering the level of human capital attainment, especially in younger populations.

The majority of work finds evidence that drug use negatively impacts human capital attainment (van Ours and Williams, 2010; Chatterji, 2006; Yamada et al., 1996; Register and Williams, 1992). The impact of substance use on education may depend on how substance use is defined. For instance, Register and Williams (1992) find that frequency of drug use has an impact, while any use does not.

D.2.h. Mental health

Given the co-morbidity of mental illness and substance abuse, mental health status is likely to have an important, if not concurrent, effect on behavior. Poor mental health has been found to have a negative effect on employment and earnings (Frank and McGuire, 2000, pg.

898). In addition to substance use, mental health status may influence the desire and ability to hold a job or even affect hiring decisions. These factors may also influence job satisfaction, which, in turn, could alter productivity and potentially impact one's physical health (Fischer and Sousa-Poza, 2009).

Some researchers have highlighted the impact of depression on employment. Lerner and colleagues (2004) investigate this relationship with patient data collected from Massachusetts's physician offices. The authors identify three groups of individuals: a treatment group consisting of individuals with depression, a control group consisting of healthy individuals, and an initial control group consisting of individuals with rheumatoid arthritis. At six months, individuals with depression experienced more new unemployment and more job turnover, conditional on employment, compared to both control groups. The results, therefore, provide some evidence of a positive association between unemployment and depression.

Some researchers have attempted to define the depression-employment relationship in the context of the drift hypothesis. Based on work by Goldberg and Morrison (1963), the drift hypothesis relates mental illness to social class and suggests that mental illness leads to a decline in one's social standing. Goldberg and Morrison's sample population draws from men aged 25 – 34, suggesting that this hypothesis is relevant to younger populations.

Reviewing the literature on depression and underemployment, Dooley and colleagues (2000) find some evidence supporting the drift hypothesis. Both Hamilton and colleagues (1993) and Dooley and colleagues (1994) find lagged depression to be positively related to unemployment. Dooley and colleagues (2000) argue that this lagged and positive association may be evidence of the drift hypothesis.

Although the current literature demonstrates an association between depression and employment, the directionality of this link is unclear. Dooley and colleagues (2000) also emphasize that the relationship between depression and employment may run in the both directions. While depression could result in poor employment outcomes, poor employment outcomes could lead to a depressed state. For individuals in their mid-20s, the authors suggest the uncertainty about employment and the establishment of one's career is a source of anxiety and depression (American Psychiatric Association, 1994). Depression may also result from unemployment or the threat of unemployment (Kessler, Turner, and House 1988). Dooley and colleagues (2000) find further evidence of this link. Taken together, this body of work suggests that a depression variable in an employment model is likely endogenous, since reverse causality is likely present, especially for younger age groups.

Treatment may help improve both health and employment outcomes for depressed individuals. Drawing on a random sample of patients from managed care practices, several studies find that patients with appropriate care experienced both lower rates of depressive disorder and higher levels of employment compared to patients who did not receive care at six months (e.g. Schoenbaum et al., 2002; Miranda et al., 2004).[4] The relative efficaciousness of treatment on employment may vary by race. While nonminority patients with appropriate care benefited from lower rates of unemployment, Miranda and colleagues (2004) find that

[4] Schoenbaum and colleagues (2002) define appropriate treatment as follows: "Appropriate treatment in the first six-months of follow-up was measured by survey items that assessed whether the respondent had four or more specialty counseling visits or used antidepressant medication for any amount of time or above the minimum dosage recommended in the Agency for Health Care Policy and Research (AHCPR) practice guidelines (Depression Guidelines Panel 1993b), adapted to include newer antidepressant medications" (pg. 1149).

this effect does not translate to minority patients with depressive disorders who undergo similar treatment.

D.3. Other factors influencing substance use and related behaviors

An individual's collective environment contextualizes the relative benefit of substance use. Factors like taxes and regulations, prices, substance type, level/frequency of substance use, personal factors, and peer effects are important in shaping the relationship between substance use and related behaviors.

The cost of using substances may strongly influence the level of use. Local, state, and federal policy helps shape the total cost of substance use. Changes in policy have a direct impact on the price and accessibility of alcohol and drugs themselves (DiNardo and Lemieux, 2001; Saffer and Chaloupka, 1999; Manning et al., 1995; Grossman, Chaloupka, Saffer, and Laixuthai, 1994; Becker, Grossman, and Murphy, 1991).

Taxes and other regulatory policies seeking to change the price of substances may have unanticipated consequences. When confronted with a change in alcohol or marijuana policy, individuals may trade one risky behavior for another. DiNardo and Lemieux (2001) find that an increase in the legal drinking age is associated with a decrease in alcohol consumption and an increase in the marijuana use among young people.

Price is another factor affecting consumption behavior. The full price of alcohol for underage drinkers equals the monetary price plus the indirect cost (Grossman, Chaloupka, Saffer, and Laixuthai, 1994). The monetary price is simply the amount an individual must pay to purchase the good. For example, the monetary price of alcohol might equal the market price of the beer, wine, or hard liquor purchased. The indirect cost, the authors argue, is more

difficult to assess. The indirect cost includes the cost of punishment if caught, time and travel costs associated with obtaining the alcohol, as well as opportunity costs.

The frequency and level of substance use may mediate the impact of price on behavior. Manning and colleagues (1995) find that policy changes altering the full price of alcohol are more effective in changing the drinking habits of more moderate drinkers relative to light or heavy drinkers. In the same way, consumers may be more or less responsive to changes in the price of one substance relative to another. Saffer and Chaloupka (1999) find evidence that cocaine users are far less responsive to price than marijuana users.

Personal factors such as race, gender, and family background shape behavior as well. Cooper and colleagues' (1994) work suggests that alcohol and marijuana use are more highly associated with sexual activity for Caucasian adolescents relative to African American adolescents.

In addition to personal factors, peer behavior also impacts substance use decisions, as well as other behaviors. Family expectations, the availability of role models, and peer behavior help mold codes of conduct, influencing many personal choices, including decisions regarding substance use. More broadly, labor force participation rates, occupational choice, education rates, crime rates, and substance use rates within a community are important factors guiding an individual's behavior (Akerlof, 1997). In this regard, own substance use generates externalities, which influence the uptake of substance use by peers. Likewise, peer substance use generates externalities that influence one's own substance use. As a result, substance use is both an individual and social decision.

This literature leads to the following conclusion: an important association exits across the three domains of mental health, substance use, and peer substance use behavior. This

association is complex, involving a variety of personal, social, and community-level contextual factors.

E. Economic models relating to substance use behavior – rational addiction versus social interaction

Much of the recent economics research examines substance use, including cigarettes, alcohol, or drug use, in the context of the rational addiction framework (Becker and Murphy, 1988). Rational addiction theory hypothesizes that substance abuse is the result of a utility maximization process, in which current substance use enhances the value of future substance abuse. The specifics of the rational addiction model are discussed in more detail below. Although rational addiction has been and continues to be important in understanding addictive behaviors, it is not always applicable. The rational addiction model may be inappropriate for adolescents and young adults, since they may not fit the clinical criteria for addiction. Younger individuals may also be more responsive to changes in price compared to older age groups. Akerlof's model of social interaction provides an alternative, albeit not entirely separate, means of investigating patterns of substance use.

E.1. Rationality: A traditional economic perspective on substance abuse

Across many fields, researchers grapple with the rationality of abuse. Becker and Murphy's 1988 model is widely used by economists to understand addictive behaviors. In their rational addiction model, Becker and Murphy argue that a substance abuse results from a rational decision-making process. They define rationality as consistent utility maximization over time. An individual addicted to a substance employs forward-looking maximization in determining the optimal amount of a substance to consume. In this framework, a substance is

considered addictive if past consumption raises current consumption. Thus, utility in the current period depends on past consumption of the addictive good.

Becker and Murphy predict that changes in past prices affect current consumption by changing the current stock of consumption capital. Changes in future consumption affect current consumption by changing the current full prices of a substance and by altering the future stock of consumption capital and, thus, future consumption.

How can future consumption affect past consumption? Suppose a state legislature passes a tax increase on all alcoholic beverages, which will be become law in exactly one year. The higher price may result in the consumer choosing to drink fewer alcoholic beverages in a year's time. Since future consumption will be lower under the new tax regime, the consumer may choose to consume less today. In this manner, future consumption affects current consumption.

This discussion leads directly into the concept of adjacent complementarity. Adjacent complementarity implies that past, current, and future consumption are complements. Along these lines, adjacent complementarily between periods fuels addiction. Adjacent complementarity implies that greater current consumption of a good raises its future consumption. For example, an anticipated price increase should decrease consumption of addictive goods in the current period. This research suggests that a consumer rationally chooses to consume a substance in a given period based on her consumption capital and her knowledge of past, present, and anticipated future prices.

Researchers outside the field of economics often disagree with the claim that substance abuse is rational. Indeed, in a larger sense, rationality reflects something more than consistent and predictable responsiveness to price changes. Becker and Murphy's model fails

to incorporate the less tangible and often unobserved components that shape decision-making. For instance, the decision to consume and continue to consume a substance may be largely shaped by an individual's environment. If one's peers are addicted to a substance, an individual's decision to become involved in substance use may be in some part a response to this environment. In the Becker and Murphy framework. He or she acts independently. Becker and Murphy do not account for the externality created by the group influence on the individual action. What is more, this rationality, or lack thereof, is further mediated by an individual's temporary and permanent mental state, both of which can be shaped by substance use itself.

E.1.a. Rational addiction model and substance use among young adults

The rational addiction model may not be applicable when analyzing the substance use behaviors of young adults. First, adolescents and young adults may not fit the clinical criteria for addiction, making the rational addiction model inappropriate. Second, young people are relatively more responsive to changes in the price of various substances, further suggesting that the rational addiction model is the incorrect approach. In part, responsiveness to price likely reflects more binding price constraints among younger individuals relative to older groups. I discuss both issues below. Genetic predisposition to addiction and dependence also counter the idea of rational addiction – among adolescents, inheritable traits and environmental characteristics impact not only initiation into substance use but also the frequency of and ability to moderate substance use (Rhee et al., 2003).

Because the rational addiction model relies on the premise of addiction, the framework may not be applicable to young adults if this age group does not manifest addictive symptoms. In many cases, some physicians argue, adolescents and young adults do

not suffer from problems of clinical addiction; rather, young people suffer from a problem of over consumption in a single sitting, such as binge drinking (Vaillant, 1995, pg. 309 - 310). The onset of addiction is gradual. Abuse and addiction stem from the social or otherwise recreational use of a substance over many years. For instance, heavy social drinking, the first stage of alcoholism, "can continue asymptomatically for a lifetime" (Vaillant, 1995, pg. 309 - 310).

In addition, the rational addiction theory may not hold for adolescents and young adults because this age group is highly sensitive to price or may have a harder time accurately predicting prices (the rational addiction model assumes that (1) consumption behavior is dependent on consumption capital not just price and (2) individuals can assess the relationship among past, current, and future prices). Young people are generally more responsive to price changes than older individuals. In fact, Lewit and Coate (1982) conclude that price sensitivity and age are inversely related for cigarettes. Even Grossman and Chaloupka (1998) find that cocaine demand by young adults is very price sensitive. Using a sample of high school seniors surveyed as part of the Monitoring the Future Program between the years 1976 to 1985, the authors estimate a short-run price elasticity of demand equal to -.96 and a long-run unconditional elasticity of demand of -1.35. Research using data from the National Household Survey of Drug Abuse estimates that the price elasticity of demand for cocaine within the general adult population ranges from -.55 to -.36 (Saffer and Chaloupka, 1995). The relative price responsiveness of adolescents and young adults suggests that alcohol and drugs may be normal market goods for this age group, since younger populations do not demonstrate the same inertia in consumption patterns as adult populations.

Consumption capital, a term used by Becker and Murphy, captures an individual's substance use history. An individual's consumption stock increases over the consecutive periods in which she engages in substance use. As with any activity, individuals "learn by doing." The more an individual engages in substance use, the higher her stock of consumption capital. Consumption capital can be thought of as dependency – higher consumption capital is associated with higher degree of dependency. As a result, an individual derives increasingly more utility from the consumption of alcohol and/or drugs the longer she engages in substance use (and increasingly more disutility when alcohol and/or drugs are not consumed).

Older individuals have often abused substances for a longer period. In the context of the rational addiction framework, older individuals, having abused a substance over many consecutive periods, amass a substantial amount of consumption capital. Comparatively, younger individuals, having abused a substance over fewer periods, have not had sufficient time to amass enough consumption capital to become addicted to a substance.

Grossman, Chaloupka, Saffer, and Laixuthai (1994) summarize several reasons why adolescents and young adults are more price responsive compared to adults. First, the authors cite Rachal and colleagues (1980), who emphasize the importance of peer effects. Grossman and colleagues write, "a rise in price would curtail youth consumption directly and indirectly through its impact on peer consumption" (pg. 351). Akerlof's social interaction theory, discussed below, reiterates the importance of peer effects. Second, adolescents and young adults spend a relatively larger fraction of their disposable income on alcohol. Finally, the authors reference Becker, Grossman, and Murphy (1991), arguing that young individuals discount the future consequences more heavily. On average, older individuals may give more

weight to the future ramifications of a current activity. While this argument has obvious implications for the volatility of price response, it also has implications for the relative risk aversion of young people – on average, one might expect a young person to be more willing to undertake risky behaviors than an older individual.

This literature leads to two conclusions: (1) adolescents and young adults may be too young to exhibit addictive behaviors manifested in adults; and (2) adolescents and young adults are relatively more responsive to prices than older ages groups. As result, the rational addiction framework may be an inappropriate theoretical foundation for understanding the substance use behavior of young people. For adolescents and young adults, substance use may be best modeled as a normal good.

E.2. Social interaction theory: the importance of individual, family, and community-level factors in mediating individual behavior

Becker and Murphy (1988) developed a theoretical economic model of substance use, emphasizing that potential users intrinsically, if unknowingly, measure the costs and benefits when deciding whether to consume a substance. Many important factors, however, are left unaccounted for in the authors' model. One such factor is peer behavior.

Peer behavior may influence substance use by reinforcing certain behaviors. Substance use is potentially "contagious," in that an individual may be more likely to indulge in substance use if her peers engage in such behavior. Social interaction theory formulizes peer effects in an economic context.

Akerlof writes, "Social interaction theory can explain why social decisions – such as the demand for education, the practice of discrimination, the decision to marry, divorce, and bear children and the decision whether or not to commit a crime – are not simple choices based on primary individual considerations," (pg. 1012). Akerlof's model of social

interaction is a blend of economics, sociology, and the Newtonian theory of gravity. The expected value of trade between two individuals depends on the difference in agents' initial positions. As a result, individuals are more likely to interact with those who are close to them on the social spectrum. For instance, individuals born into affluent communities are expected to interact more closely with and adopt similar behaviors of other individuals from their affluent community.

Following Akerlof's logic, the individual and social decision-making processes shape an individual's behavior. Externalities arise when individuals react positively or negatively to their place on the social spectrum. In a positive reaction, individuals engage in conformist behavior and adopt the accepted and expected behavior of their peers. Such behavior may perpetuate behavioral trends within a community. While some of these patterns can be positive, like going to college, others can be negative, like the use and sale of drugs. Alternatively, an individual can engage in what Akerlof refers to as status-seeking behavior and try to distance himself from his peers. Again, the ramifications of such a move can be both positive and negative.

The models of social conformity and status-seeking are quite different, each modeling two separate behaviors. For instance, each model has different implications for individual steady states. The conformist model results in multiple steady states. Trapped in behavioral patterns, a coordinated breakout is needed to transform the actions of both the individual and the community. These differences arise from the varying construction of the indirect utility function.

I examine the net effect of the conformist and status-seeking models in my conceptual framework. While both examine reactionary behaviors, an individual reacts positively to peer

behavior in the conformist case (i.e., she conforms to peer behavior) and negatively to peer behavior in the status-seeking case (i.e. she attempts to separate herself from her peers).

Very few researchers have examined the impact of peer substance use on individual risky behaviors. Pertold (2010) examines the impact of peer alcohol use on risky sexual behaviors. His sample population is drawn from the European School Survey of Alcohol and Other Drugs (ESPAD) and includes Czech teens roughly 18 years of age. Pertold's findings vary by gender. Female alcohol consumption affects male propensity towards risky sexual behavior but male drinking does not influence female propensity towards risky sexual behavior. Although unmentioned, the results allude to a scenario in which predatory male peers take advantage of intoxicated females. Pertold's work builds on Waddell (2010), who also found evidence of cross-gender peer effects of alcohol on risky sexual behavior.

F. Conclusions

This summary finds several key gaps in the literature. Existing work largely neglects the role of mental health and substance use. The peer effects of substance use also remain under-investigated within this framework. I fill these gaps by controlling for both individual and peer substance use as well as depression in all models. Researchers often ignore the endogeneity of substance use and mental health. Instrumental variable techniques help account for the endogeneity of these measures.

In my analyses, I draw from a broad definition of substance use and sexual behavior. I examine the impact of alcohol, marijuana, cocaine, and methamphetamines on the outcomes of interest. The durational effects of substance use on employment remain under-examined; this work fills this gap by analyzing the impact of substance use on employment of young adults multiple years after the event. I also investigate a unique spectrum of sexual

outcomes, including number of sexual partners and sexually transmitted diseases. Finally, I investigate the link between substance use, depression, and rape victimization, a topic that also remains largely unexamined.

III. CONCEPTUAL FRAMEWORK

In this section, I develop a theoretical model for understanding the interaction of risk-taking, substance use, and employment. I base the conceptual model on a utility maximization framework. The choice variables in this analysis include the amount of market goods, alcohol/drugs, and hours of leisure consumed. In theory, each individual maximizes her utility, which is a function of the aforementioned factors, subject to several constraints. First, an individual is constrained by her budget – her spending on market goods, alcohol, and drugs can only be as large as her income. Second, an individual is constrained by the cost of risk-taking – if the perceived cost of a risk is larger than her reservation cost, an individual will choose not to engage in the risky behavior.

The neo-classical model of the labor-leisure tradeoff stipulates that an individual derives utility from both leisure and market goods. For our purposes, however, an individual's utility is also influenced by the consumption of drugs and alcohol. Thus, the choice variables for this analysis are substance use (S), the peer effects on substance use ($\alpha(S-\bar{S})$), market goods (X), and hours of leisure (L). Social distance is defined by the function $\alpha(\cdot)$, which characterizes the degree to which individual conformity to the peer behavior increases or decreases utility. The difference between own and average peer substance use acts as choice variable since (1) an individual chooses the amount of substance she consumes and (2) an individual chooses her peers. Risk-taking (R) is a function of

substance use. Individual behavior is conditioned on predisposing characteristics (ε) and depression status (D).

A. Background on conceptual framework

Markowitz (2000) and Akerlof (1997) help shape this model. Markowitz argues that two possible alternatives link substance use and violent behavior. In the first case, alcohol has a physiological effect on an individual that produces violent outcomes. Thus, the alcohol consumption causes the behavior (Markowitz [2000] investigates the impact of alcohol consumption on spousal abuse). In the second case, Markowitz argues that an individual uses alcohol consumption as an excuse for a behavior (e.g., spousal abuse). In this scenario, alcohol consumption and the negative outcome are merely associated. I generalize Markowitz's work to examine the relationship of substance use and risk-taking within the context of a neo-classical model of the labor-leisure tradeoff. I also expand her framework to incorporate depression and peer behaviors.

As implied by the above discussion, Markowitz (2000) assumes each case is entirely separate; however, the two scenarios may not be distinct. An individual may, in part, consume alcohol or drugs, knowing full well that she is going to engage in risky behavior. Once consumed, however, these substances may propagate and extenuate riskiness. Arguably, alcohol and drug use themselves qualify as risks, further substantiating the association of substance use with risk-taking. As a result, risky behavior is plausibly a function of substance use.

Possible negative consequences, by definition, accompany any risk. An individual accounts for the costs associated with riskiness, at least to some degree, even if risk-taking is

spurred by substance use. Thus, the constraint includes the costs associated with taking a risk.

In developing my conceptual model, I also draw on Akerlof's discussion of social networks and the effect these social networks on the individual decision-making process. I apply Akerlof's utility framework for status-seekers and conformists to help determine the impact of peer substance use on utility. Status-seekers consist of those individuals who try to distance themselves from their peers in social space. The opposite is true for conformists.

A.1. Status-seekers

A status-seeker attempts to distance herself from the behavior of her peers. In Aklerof's model, a status-seeking individual chooses the level of substance use that maximizes her indirect utility function. In maximizing her utility, a status-seeking individual attempts to gain status by distinguishing herself from her peers through markedly higher or lower levels of substance use. *d* in equation (1) reflects an individual's taste for non-conformity and is important because it measures the significance of peer behavior relative to the individual valuation of substance use. Therefore, if *d* is large relative to the combined effects of *a*, *b*, and *c*, peer effects will dominate personal factors in determining utility for individuals that seek to distance themselves from their peers, both positively or negatively.

(1) $$U = -d(\bar{S} - S) - aS^2 + bS + c$$

S denotes substance use – that is, the level of alcohol or drug consumed. $\bar{S}$ signifies the average level of substance use within an individual's peer group. $\bar{S}$ is available in Add Health (Data section discusses its measurement). The $-aS^2 + bS + c$ component of equation (1) refers to the intrinsic value of substance use, meaning the personal valuation of substance use that is unaffected by one's social network. If an individual derives zero utility from peer

substance use, the utility function will reduce to $U = -aS^2 + bS + c$. By design, the intrinsic value of substance use is assumed to be quadratic and is defined by the parameters a, b, and c.

Based on the definition of the utility function, an individual's utility falls as either $\bar{S}$increases or S falls. Based on Akerlof's specification of the utility function, the optimal level of substance use is $(b+d)/2a$, which is found by setting the first derivative of (1) equal to zero. Desire for nonconformity generates externalities that enter through the $(d/2a)$ term.

For a status-seeking individual (assuming $d > 0$), Akerlof's work implies that utility will increase as an individual's substance use (S) deviates from the average level of substance use within that individual's social network ($\bar{S}$). If, on average, her social network engages in heavy substance use, a status-seeking individual's utility will increase as she decreases her alcohol or drug consumption.

Akerlof only presents one model for status seekers, the indirect utility function detailed in equation (1), suggesting that Akerlof means the equation to be symmetric (the same for positive and negative status seekers, depending on the sign of d).

Observable characteristics may also shape peer group. For instance, race, gender, and age likely influence the peers with whom an individual chooses to associate. As a result, these observable factors enter into my empirical models as control variables (see Data and Methods sections).

A.2. Conformists

A conformist derives utility from emulating the behavior of her peers. Akerlof's indirect utility function for conformist individuals is quite different. A conformist individual

chooses the amount of S that maximizes the following indirect utility (again, d can be positive or negative).

(2) $U = -d \mid S - \bar{S} \mid -aS^2 + bS + c$

Although Akerlof refers to equation (2) as the "twin model of conformity," equation (2), by construction, differs from equation (1) in that equation (2) uses the absolute value of the deviation of individual behavior from mean peer behavior. Akerlof writes that conformists' "tendency to mimic the status quo can result in the underproduction or overproduction of [S]." This suggests that multiple equilibria are possible – people will over use or under use alcohol and drugs in an attempt to mirror the substance behavior of their peers. Mathematically, this effect is modeled through the use of the absolute value.

Utility increases as the distance between S and $\bar{S}$ shrinks. The first order conditions from the conformist utility function result in multiple equilibria. An individual faces several optimal values of substance use. She will choose to consume an amount of a substance within the range of *(b-d)/2d and (b+d)/2d*. This range is heavily influenced by d, the degree to which an individual desires to conform to the behavior of her peers (high values of d are associated with a smaller range). Desire for conformity generates externalities, resulting in the overproduction or underproduction of S.

Akerlof's framework demonstrates not only how substance use can contribute to utility, but also how peers mediate this contribution. For the conformist individual, utility will increase as individual substance use (S) approaches the average level of consumption within her social network ($\bar{S}$), assuming $d > 0$.

A.3. Endogeneity or exogeneity of peer group

An individual's peer group is exogenous to the extent that the community she lives in, the school that she attends, and other factors shaping grouping are out of her control. Parents, not adolescents, make these decisions. On the other hand, an individual may select herself into a peer group that mirrors her preferences. In this case, an individual who enjoys frequent drug use may choose to consort with other drug users, suggesting that peer behavior is endogenous. It is difficult to assess an individual's intentions. Regardless, Akerlof's work underlines the importance of controlling for peer behavior. Observable, individual characteristics may also shape one's peer group. As a result, these factors must be controlled for in empirical models.

A.4 Additional factors shaping sexual and employment behaviors

Environmental factors, social factors, and individual factors, such as relative risk aversion, likely shape sexual partnering and employment as well as substance use choices. Identifying substance use separately from sexual partnering requires the independent identification of substance use from these other factors. As result, the coefficient on substance use in naive models likely reflects the impact of these factors on sexual behavior.

B. Utility function

The utility function is defined in equation (3).

(3) $U(S, d(S - \bar{S}), R, X, L \mid \varepsilon, D)$

Personal characteristics (ε) determine an individual's predisposition towards alcohol and drug use as well as risk-taking. Personal factors will also affect employment behaviors as well as the types of goods individuals purchase with the income earned from this employment or given to them by their parents. Similar to other personal characteristics,

individual tendency towards depression (*D*) mediates employment and consumption decisions; depressed individuals are potentially more or less likely to use alcohol or drugs, take risks, and be employed. Peer behavior ($\bar{S}$) and peer effects on substance use ($\alpha(S-\bar{S})$) reflect the spirit of Akerlof's social interaction theory. If one is interested in the durational effects of substance use, one might make the assumption that *S* measures the cumulative impact of past, present, and future substance use at time *t*. Drawing on Becker and Murphy's 1988 rational addiction model (see Background and Significance section), adjacent complementarity implies that past, current, and future consumption are complements, so this assumption makes intuitive sense.

C. Budget constraint

In order to pay for the consumption of alcohol, drugs, and other market goods, an individual must work or rely on some form of household earnings. Individuals may also resort to black market sources of income, such as theft. Given the relative youth of the population of interest, I allow an individual to receive financial assistance from his or her family. Thus, the model incorporates for household endowments (*HE*): monetary gifts from family members to an adolescent or young adult. These gifts account for allowance, monthly financial transfers, or other pecuniary assistance.

Individuals are also allowed to work in the model. By the fourth wave, individuals in the Add Health data range in age from 24 to 32, by which time most individuals will have matriculated into the labor force if they desire to do so. The number of hours an individual works is equal to the total number of hours available (*T*) minus the number hours he or she spends enjoying leisure (*L*). Assuming she works at a wage of *W*, an individual earns a labor income of *W (T-L).* Therefore, a given individual faces an inherent tradeoff between leisure

(*L)* and the consumption of drugs and alcohol (*S*) as well as other market goods (*X*). The price of other consumption goods (P_x) is normalized to one dollar.

(4) $$HE + W(T - L) = Ps \cdot S + X$$

Although not addressed in this conceptual framework, illegal income, such as the sale of illicit substances or stolen goods, could possibly supplement an individual's finances. Within the context of Becker (1968), one might assume that a person's motivation and probability of engaging in illegal activities, and thus acquiring "black market wages," is in part dependent on her level of risk aversion, her education, legal ramifications, and the opportunity cost of other foregone employment. Illegal income is not available in the data, and therefore is not included in any empirical models.

An individual also maximizes utility, subject to the cost of the risky behavior. In accordance with Becker (1968), an individual assesses the cost of a behavior. For a given risk, an individual faces a cost *C* with probability α. This cost could include punishment from parents, legal ramifications, retribution from peers, jail time, or other sanctions. Imprisonment or other legal sanctions may in turn affect an individual's labor earnings in the next period.

An individual will only engage in risky behavior if the estimated cost (αC) is below some reservation cost (*Co*). Once this cost is too high, she will choose not to participate in a particular risky behavior. When the cost is too high, $\alpha C > Co$. Because the cost of risky behavior exceeds the reservation costs, an individual will not to engage in the risky behavior. Because the individual is not engaging in risky behavior, she does not incur the cost associated with taking the risk. Therefore, when $\alpha C > Co$, costs associated with punishment (*J*) will be equal to zero.

Alcohol or drug consumption may alter the decision-making process by lowering the individual value of *Co* relative to a sober state, suggesting that cost is a function of substance use. In this model, however, I make two simplifying assumptions. First, the level of substance use and the decision to engage in risky behavior are made simultaneously at the outset of the period. Second, current substance use decisions (*S* in period *t*) reflect the impact of prior substance use decisions (Substance use, *S*, is designated in equation [3]. *S* in *t* reflects information acquired in periods 1 though *t* – 1). Equation (5) describes the constraint an individual faces with respect to the cost associated with risky behavior.

(5) *If* $\alpha C < C_o$, $J = R*\alpha C$
If $\alpha C > C_o$, $J = 0$, *since* $R = 0$

D. First order conditions

Utility maximization yields the following first order conditions:

(6) $$\frac{\partial U}{\partial S} + \frac{\partial U}{\partial R}\frac{\partial R}{\partial S} = \lambda_1 P + \lambda_2 \frac{\partial \alpha}{\partial S} C \cdot R$$

(7) $$\frac{\partial U}{\partial X} = \lambda_1$$

(8) $$\frac{\partial U}{\partial L} = \lambda_1 W$$

The first order conditions imply that λ_1, the marginal utility of income and wages, and λ_2, the shadow price of risky behavior, are positive. By assumption, composite market goods and leisure are normal, thus $\partial U/\partial X$ and $\partial U/\partial L$ are positive. Given that wages are also positive, λ_1 and λ_2 must be greater than zero.

Looking at equation (6), $\partial U/\partial S$, $\partial U/\partial R$, *Ps, C*, and *R* are positive. $\partial U/\partial R$ is also assumed to be positive - individuals would never take risks if the utility they derived from risk-taking were negative. The second partial derivative of the utility function with respect to

risk is likely negative, implying that that an individual derives increasingly less utility from risk-taking as they engage in more risky behavior. $\partial\alpha/\partial S$ is positive, since alcohol or drug use may increase the probably of punishment as a result of risk-taking. Taken together, these factors imply that $\partial R/\partial S$ must also be positive.

Substance use is a function of household endowments, wage rate, the relative price of drugs or alcohol, and the cost of risky behavior. Substance use is conditioned on personal factors, tendency toward depression, and peer substance use.

(9) $S = S(HE, w, P_s, C \mid \varepsilon, D)$

From the first order conditions and economic, depression, and substance use literature, one can sign the factors determining substance use in equation (9). Alcohol and drugs are assumed to be normal goods. The model predicts that income and substance use are positively related, since alcohol and drugs are assumed to be normal goods. Interestingly, this assumption is also upheld in the Becker and Murphy framework, since they argue that a permanent increase in price will decrease substance use. Thus, the standard law of demand is expected to hold (i.e., as price increases, quantity demanded falls). Given the co-morbidity between substance use and depression, the relationship between substance use and depression is assumed to be positive. The effect of peer substance use on utility is more complex and can veer in either direction, as discussed in the context of Akerlof's status-seeking and conformists models.

Since substance use is a function of risky behaviors, we know that risky behaviors must be a function of the same factors discussed above.

(10) $R = R(HE, w, P_s, C \mid \varepsilon, D)$

Again applying the first order conditions and economic, depression, and substance use literature, one can sign the factors determining substance use in equation (10). Risky behavior, income, and predisposed factors are expected to be positively associated. I predict that risky behavior and price are negatively related, although this hypothesis can be tested in the empirical models. The relationship between risk-taking and depression is less clear. Although the conceptual model posits that depression contributes to risk-taking, the direction of the effect remains to be seen and will be estimated empirically.

Finally, a labor supply function can be created. Assuming E is a binary indicator of employment, which equals one if an individual works any hours, employment can be defined as follows.

(11) $E = E(HE, w, P_s, C \mid \varepsilon, D)$

Employment decisions are interlocked with consumption and budgetary conditions. The same factors that encourage an individual to consume a particular amount of drugs, alcohol, and other market goods are going to shape employment decisions, since an individual can only consume as much as her wage earnings and household endowments allows.

Several major testable implications result from the first order conditions. First, the conceptual model predicts that risk-taking and substance use are positively associated. This hypothesis will be tested in the empirical models in Aims One and Two. Second, the conceptual model predicts employment to be a function of depression, but this relationship cannot be signed. The direction of this effect will be estimated in Aim Three. Finally, the conceptual model predicts that employment and risk-taking are both functions of depression and peer substance use; however, the direction of these relationships remains indiscernible. The empirical models in Aims One and Two estimate the direction of the relationships

between depression and peer substance and sexual behaviors; Aim Three helps determine whether depression and peer substance and employment outcomes are positively or negatively associated.

IV. DATA

I begin this section by outlining the data source, the National Longitudinal Study of Adolescent Health. Next, I describe all dependent and explanatory variables. While the dependent variables vary across aims, the explanatory variables are largely constant across all aims. I end the section by discussing missing values and how they are treated.

Tables 1 and 2, which can be found in the appendix, detail aim-specific attributes related to each analysis, including information on the dependent variable, waves used, sample, and empirical methodology. These tables may be helpful references while reading through the remainder of the document. I present descriptive statistics for each aim (Tables 5 – 7) at the end of this section.

A. Data source

A.1. Data source – overview of the National Longitudinal Study of Adolescent Health

The Carolina Population Center administers the National Longitudinal Study of Adolescent Health, or Add Health. Add Health is one of the largest nationally representative surveys of adolescent behavior, unique in its assessment of individual health-related behavior as well as environmental factors.

The first wave of the study, conducted between 1994 and 1995, surveyed individuals in grades seven through twelve. Three follow up waves were administered. The second wave took place in 1996, the third wave took place between 2001 and 2002, and the fourth wave took place between 2007 and 2009. By the fourth wave, respondents ranged in age between 24 and 32.

While baseline interviews were conducted in schools, the Add Health data consist of self-reported information collected at home as well as at school. To the extent that individuals cannot remember or misrepresent their substance use, sexual, and health history, the estimates could be biased.

Looking at the validity of student self-reported data, Wilson and Zietz (2004) find that data relating to sensitive subjects are relatively more prone to bias. The substances use measures are likely the most sensitive to systematic inflation or deflation of values, due to the illicit nature of most substances and the negative social mores attached to substance use. Drug use may be commendable or stigmatizing within a young person's peer group, and students may misrepresent their level of substance use accordingly.

The sample population for the Add Health study was drawn from a pool of 26,666 U.S. high schools. The inclusion criterion required that schools span through at least the eleventh grade and have had a minimum of 30 students. Inclusion further necessitated that high schools be affiliated with a middle school. Having a feeder school connected to the high school allowed for a greater variation in age as well as continuity of the high school population in the second wave of the study. Over 70% of eligible schools participated. 134 public, private, and parochial schools chose to participate in the final sample.

A.2. Data source - sample design and accommodation for design effects

The primary sampling unit for data collection consisted of high schools. Student rosters were created for participating schools. Within schools, roughly 17 students who completed the in-school portion of the survey were randomly drawn from each grade-sex group. Certain groups were oversampled. Add Health included oversampling of certain ethnic groups. For instance, African American students with a college educated parent and Chinese, Cuban, and Puerto Rican students were oversampled. Add Health also oversampled disabled students and non-related adolescents respondents who shared a residence. The data are stratified by region, urbanicity, school size, type, racial composition, and grade span.

Because not all groups of individuals were sampled with equal probability, performing an analysis without taking into account the survey design and weights would result in biased estimates (Chantala and Tabor, 1999). To account for these design effects and obtain nationally representative and unbiased estimates, the administers of Add Health provide sampling weights (Chantala and Tabor, 1999). I accommodate for design effects by performing a design-based analysis using the Add Health weights and STATA'a built-in complex survey capabilities (i.e., the *svy* commands). This is the approach suggested by Chantala and Tabor (1999). In addition to following the recommendations discussed by Chantala and Tabor (1999), I follow the weight guidelines outlined in the latest wave of Add Health, which can be found in the Add Health codebook (see "Guidelines for choosing the correct Sampling Weight for Analyzing Add Health Data").

A.3. Data source - response rate

The response rates for Waves One, Two, Three, and Four were 78.9%, 88.2%, 77.4%, and 80.3%, respectively[5]. In Wave One, 20,745 adolescents and 17,700 parents completed the in-home survey. In Wave Two, follow-up interviews were conducted with individuals who were in seventh to eleventh grade as of Wave One. Individuals who were in the twelfth grade as of Wave One were not re-interviewed unless they were part of the genetic sample, which consisted of monozygotic twins, dizygotic twins, full siblings, half siblings, and unrelated siblings who were raised in the same household. Wave Two recruited sixty-five new adolescents for the genetic sample. While the genetic data provide interesting information, the sample size is quite small and, as a result, is not used in this analysis. Wave Two excludes individuals only included in the disabled sample in Wave One. In Wave Two, 14,738 adolescents and young adults completed the in-home survey. Waves Three and Four re-interviewed respondents from the first wave of the study. The sample population sizes for these waves were 15,170 and 15,701 participants, respectively.

B. Measures

First, I discuss the dependent variables, which are unique to each analysis. Second, I describe the key explanatory variables and other important control variables, which are common across all analyses. Finally, I identify missing values (Table 4) and present descriptive statistics (Tables 5 – 7). Tables 1 and 2 detail aim-specific attributes related to each analysis, including information on the dependent variable, waves used, sample, and empirical methodology.

---

[5] While Waves Two through Four draw from the Wave One population, new individuals are also recruited into the sample.

B.1. Dependent variables

The four dependent variables studied in this dissertation are (1) an individual's reported total number of sexual partners, (2) whether or not an individual reported any sexually transmitted disease in the last year, (3) whether or not a woman reported ever being raped at the time of the interview, and (4) reported labor market outcomes.

B.1.a. Total number of sexual partners

The key dependent variable in the Aim One analysis is total number of reported sexual partners. This study draws on Waves One and Two. I draw observations from Waves One and Two and control for time. I cluster the standard errors at the individual level.

Add Health data contain information on sexual history, including romantic, nonromantic, and total number of partners. The Add Health survey asks in Wave One, "With how many people, in total, including romantic relationship partners, have you ever had a sexual relationship?" Wave Two follows-up on this question, asking respondents who many sexual partners they have has since they were last interviewed.

Age is one factor limiting sexual encounters – a large portion of the sample population is under the age of 15. Most individuals under the age of 15 are sexually inactive because they are still quite young. To address this issue, I limit the sample population to those ages 15 and up. Over 30% of the sample population in Waves One and Two are aged 15 and up and are sexually active.

B.1.b. Sexually transmitted diseases status in the last year

The second component of Aim One tests the impact of substance use on risky behavior using an alternative definition of sexual risk-taking – whether or not an individual reports being diagnosed with a sexually transmitted disease in the last year. The sexually

transmitted disease analysis is longitudinal, drawing on Waves One and Two. A binary measure assesses whether or not an individual reports any sexually transmitted disease within a one-year period. The binary measure equals one if an individual has reported having at least one STD in the last twelve months.

Wave One inquires if an individual has been diagnosed with an STD less than a year ago. Wave Two asks individuals if they have been diagnosed with an STD in approximately the last year. Both the STD measures capture whether or not an individual has been diagnosed with an STD in approximately the last year.[6]

B.1.c. Ever being raped as report at the time of interview

This study is longitudinal, drawing on Waves One, Two, and Four (Wave Three does not ask about an individual's history of rape). The dependent variable in the analysis is a binary measure of whether an individual reports ever being raped at the time of the interview. In gauging the prevalence of victimization, Waves One and Two ask women, "Were you ever physically forced to have sexual intercourse against your will?" Men are asked, "Did you ever physically force someone to have sexual intercourse against her will?" By design, perpetuity falls squarely on the shoulder of men in these waves. While other scenarios are possible and certainly do occur, in a majority of cases, males rape females (CDC, 2007). As a result, this analysis does not test the occurrence of rape victimization among males. Questions relating to rape in Wave Four are asked of both men and women, thereby relaxing the assumptions made in Waves One and Two. Because of the design of the survey, I limit longitudinal analyses regarding victimization to females.

[6] Wave One and Two interviews accorded approximately one year apart. Therefore, the Wave Two question asking whether an individual had an STD since the last interview is roughly equivalent to the Wave One question asking whether an individual has had an STD in the last year.

B.1.d. Labor market outcomes

This study will be cross-sectional, with data drawn from Wave Four and with lagged substance use measures drawn from Wave Three.[7] Two dependent variables are included – a binary indicator of employment status and a continuous measure of wage rate. I consider an individual to be employed if they are working at least ten hours a week at the time of the interview.[8] Conditional on employment, the wage rate variable reflects an individual's hourly wage. I construct the hourly wage measure from information on personal earnings and number of hours worked.

Individuals are asked how much income they received from personal earnings before taxes. Some individuals do not know the exact value of their personal income but could approximate their income within a certain range of values. For these individuals, 421 in all, the mean value of the income range is used as a measure of personal earnings. For instance, if an individual does not remember the exact amount of his or her personal earnings, but does report that his or her personal income falls between $25,000 and $29,000, I approximate his or her personal earnings to be $27499.50.

Individuals are also asked about the number of hours per week they normally work at their current job. The hourly wage measure is constructed by dividing the personal earnings by the number of hours worked per week multiplied by the number of weeks in the year. In cases where respondents work more than one job, I use the total number of hours worked at all jobs in the aforementioned calculation.

[7] Although substance use is lagged, the variable is still treated as endogenous.

[8] Previous research has relied on a 10-hour threshold in defining employment (e.g., Norton and Han, 2008).

B.2. Key explanatory variables

The primary explanatory variables for each analysis are substance use, depressive symptoms, and peer substance use. In instrumental variables analyses, substance use and/or depressive symptoms will be dependent variables in first stage estimation. The substance use, depressive symptoms, and peer substance use measures are largely consistent across each analysis.

B.2.a. Substance use

The data gauge the respondent's consumption of a variety of substances. This work will include measures of alcohol, marijuana, cocaine, and methamphetamines. Given that the pharmacological and behavioral effects of substance use vary by substance, the pathway by which substance use affects behavior will also likely be different. As a result, models are run separately by substance.

Depending on the aim, the level of drug use is gauged by (a) any drug use and (b) any drug use in the last year. Since I am interested in drug use and not drug abuse, a clinical definition of drug abuse is not used.

The type of substances studied in each wave will depend on the continuity of the substance measures across the different waves used in the analysis. Methamphetamines data are only available in Waves Three and Four. While Waves One through Three ask about cocaine and other substance use in the last 30 days, Wave Four only asks respondents if they have ever used these drugs. Measures of substance use in the last year are only available for Waves Three and Four. Tables 1 and 2 detail how substance use is defined in each aim.

I measure alcohol abuse with a binge-drinking indicator. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) defined Binge drinking is follows:

> A "binge" is a pattern of drinking alcohol that brings blood alcohol concentration (BAC) to 0.08-gram percent or above. For the typical adult, this pattern corresponds to consuming 5 or more drinks (male), or 4 or more drinks (female), in about 2 hours (Definition approved by the NIAAA on 02/05/2004).

In Waves One through Three, Add Health asks respondents about the number of days in the last twelve months they have consumed five or more drinks in a row. Wave Four uses separate binge-drinking criteria for men and women. The survey asks men about the number of days in which they have had five or more drinks, while asking women about the number of days in which they have had four or more alcoholic beverages. Although women in Wave Four are more likely to fit the binge drinking criterion, since four—not five—alcoholic beverages consumed in a single sitting qualifies as binge drinking, the binge drinking measure is largely consistent across waves. While the indicator changes between Waves Three and Four, both definitions function as a basic gauge of binge drinking behavior based on the NIAAA definition.

B.2.b. Depression

Both Waves One and Two assess symptoms of depression using a nineteen-question series referred to as the Feelings Scale. A benefit of using the Feelings Scale rather than a self-reported depression diagnosis by a health professional is that the prevalence of mental health conditions is far greater than the rate of actual diagnosis (Shapiro et al., 1985; Regier et al., 1990). Nine out of nineteen questions from the original Feelings Scale in Waves One and Two are available in later waves. The nineteen-question Feelings Scale from Waves One and Two is outlined in Table 3. The nine questions that are available in all waves are also indicated in the table.

The Feelings Scale is very similar to the 20-item Center for Epidemiologic Studies Depression Scale (or CES-D) (Radloff, 1977; Goodman and Capitman, 2000; Eleden and

Reeve, 2007). Both the CES-D and the Feeling Scales inquire about an individual's emotional wellbeing during the past seven days. In fact, sixteen of the nineteen questions on the Add Health Feeling Scale are identical to questions asked on the CES-D.

Two of remaining three questions from the Add Health depression measure are very similar to corresponding questions on the CES-D scale. The CES-D asks respondents how many times during the last seven days "[they] felt that everything [they] did was an effort." Add Health asks respondents how often "it was hard to get started doing things." Similarly, the CES-D measure asks respondents about the frequency with which "[they] felt that [they] were too tired to do things," while Add Health asks respondents how often they found "[they] could not get going.'"

Two questions from the CES-D are not included in Add Health and one question from Add Health is not on the CES-D. Unlike the CES-D, the Feeling Scale does not inquire about restless sleep or crying spells. The Feeling Scale asks about the frequency with which an individual "felt life was not worth living," which is not incorporated in the CES-D scale.

The response items for the two scales are also comparable. Both the CES-D and the Feelings Scale ask respondents to rate their feelings according a four-point scale. On the CES-D, the available response items are *rarely or none of the time*, *some of or a little of the time*, *occasionally or a moderate amount of the time*, and *most or all of the time*. On the Add Health Feelings Scale, the response items include *never (zero*), *sometimes (one)*, *a lot of the time (two)*, or *most to all of the time (three)*.

A majority of the questions on the Feeling Scale gauge the frequency with which individuals experience emotion disturbance. For example, the Feeling Scale inquires about thoughts of sadness, loneliness, depression, and fearfulness. Four of the nineteen questions,

however, assess positive emotions, inquiring about the frequency with which individuals have felt happy, hopeful about the future, enjoyed life, and felt that they were as good as other people. In these cases, I invert the four-point scale so that higher values coincide with higher leaves of emotional disturbance (i.e., for these four questions, zero indicates most of the time or all the time and three indicates rarely or never).

I generate a single measure of depressive symptoms by creating percentile scores based on responses to the Feeling Scale, following the work of Evenhouse and Riley (2005), who also worked with Add Health. I create percentile scores by summing individual depression scores, normalizing the score by age and gender, and, finally, calculating a percentile for the normalized sum.

Different mental health assessment tools are used in Waves Three and Four. Both the "Social Psychology and Mental Health" and "Illnesses, Medications, and Physical Disabilities" sections gauge mental health. The "Social Psychology and Mental Health" sections in Waves Three and Four are similar to the Feelings Scale. Although the format of these sections is not identical to the Feelings Scale in Waves One and Two, there is significant overlap. Nine of the nineteen questions from the original Feelings Scale are repeated in Social Psychology and Mental Health section in Waves Three and Four. Items repeated in Waves Three and Four are indicated in Table 3.

For aims relying only on Waves One and Two, the full version of the Feeling Scale will be used. Aims using Waves Three or Four only use the nine questions included on both the "Social Psychology and Mental Health" section and Feelings Scale. This abbreviated CES-D measure has been validated among elderly populations (Fonda and Herzog, 2001). The correlation between shorter versions of the CES-D was found to have a positive

correlation with the full 20-question version (Furukawa et al., 1997). In both Waves One and Two, the correlation between depression percentile variables created for the 19-item scale and the abbreviated 9-item scale is above 0.9.

B.2.c. Peer effects

Across the four waves of Add Health study, four questions inquire about the substance use behavior of an individual's three closest friends: (1) Of your three best friends, how many smoke at least one cigarette a day?; (2) Of your three best friends, how many drink alcohol at least once a month?; (3) Of your three best friends, how many binge drink at least once a month?; and (4) Of your three best friends, how many use marijuana at least once a month?

Some peer substance use questions are only asked in certain waves. As a result, certain peer substance use measures can only be applied to certain analyses, depending on which waves are used. The peer substance use measures used in each analysis are listed in Tables 1 and 2.

B.3. Other controls

Control variables capture other individual and family characteristics. These variables are important because they explicitly control for factors that are constant under the *ceteris paribus* assumption of the model. The control variables are constant across all three aims. In addition to the variables mentioned, other controls include measures of age, gender, and race. Individuals under the age of 15 are not included in any analyses. I exclude this group since a majority of individuals under the age of 15 are not sexually active and most cannot legally work under current federal law.

Body mass index

Research surrounding sexual behavior in young people using the Add Health data has incorporated measures of body mass index (BMI) (Halpern et al., 2006). The BMI variable is constructed using self-reported height and weight.[9] I include BMI in all analyses and treat the variable as exogenous.

Interviewer- assessment of physical attractiveness

The interviewee's attractiveness likely has an effect on the number of sexual partners he or she "attracts." Physical attractiveness might also be associated with personal confidence, which, in turn, may be related with sexual behavior and employment. Investigating emerging sexual partners among young adults, Halpern and colleagues (2006) control for attractiveness. In their analysis, attractiveness is statistically significant. Holding BMI constant, the authors find that physical attractiveness is negatively associated with a probability of virginity.

In the interviewer's remarks sections for each wave, the interviewer is asked to rate the relative attractiveness of the interviewee on a scale of one (*very unattractive*) to five (*attractive*). For ease of interpretation, I consolidate these five outcomes into three dummy variables, which reflect whether someone is of below average, average, or above average attractiveness. Halpern and colleagues (2006) also create an amalgamated attractiveness measure.

The inter-rater reliability of this variable cannot be calculated, since only one interviewer rates attractiveness during the interview. While the measure is subjective, the attractiveness dummies capture unique and perhaps unobserved personal traits that conceivably affect both employment and sexual experiences.

[9] Body mass index (BMI) is calculated as follows: BMI = mass (lb) * 703) / height $(in)^2$.

Education

Five dummy variables are created to assess the highest level of education completed – less than a high school degree, high school degree, a four year college degree, or a post-baccalaureate degree. By Wave Four, a majority of respondents have a four-year college degree. I treat education level as exogenous; I discuss the implications of this decision in the limitations section.

Student status

Waves Three and Four continue to track individuals' educational attainment. A student dummy equals one if an individual reports being currently enrolled in school. The student status variable captures something different from the educational attainment variables. Current students' daily experience is distinct from non-students; and, as a result, current students might face different social, sexual, and employment environments.

Marriage and children

Two dummy variables indicate if an individual has ever been married and if an individual has any children. In Waves One and Two, Add Health limits questions regarding marriage to those respondents who are aged 15 and up. My conceptual model discusses "predisposing characteristics" (ε) that prompt or inhibit risk-taking, substance use, and employment. Marriage and children are assumed to fall into this category.

Current Smoker

Tobacco use likely correlates with other forms of substance use and potentially serves as a "gateway" substance. As such, smoking likely influences the other, often more immediately dangerous, forms of substance use studied. Alexander and colleagues (2001) and Goodman and Whitaker (2002) research smoking in adolescent populations,

investigating the effects of being a "current smoker." They define current smokers as those individuals who smoked at least one cigarette in the last 30 days. I define current smoking in the same manner. I treat current smoking as exogenous.

C. Missing data and sample sizes

Table 4 outlines the initial and final sample populations and details the number of missing values by variable. The values are non-sequential, each cell indicating the number of missing values associated with a variable. Similar to Halpern and colleagues (2006), I exclude observations from the analysis if they have missing values for the variables of interest. The bottom three rows of Table 4 list the initial number of observations, the final number of observations without weights, and the final number of observations with weights for each aim, respectively.

As the table demonstrates, the number of observations dropped due to missing values is not substantial. Of all the aims, Aim Three suffers the greatest reduction in sample size because of missing values. This loss is largely driven by the lagged substance use measures drawn from Wave Three (all other variables for the analysis are drawn from Wave Four). Not all individuals interviewed in Wave Three are interviewed in Wave Four, and not all individuals interviewed in Wave Four are interviewed in Wave Three, resulting in missing values for the lagged substance use measures.

More generally, of all the explanatory variables, substance use measures seem the most prone to missing values. Substance use is largely a prohibited act, with the expectation of alcohol consumption among those 21 and older. Student self-reported data on sensitive subjects are more prone to bias (Wilson and Zietz, 2004). As discussed earlier in this section, in some cases, instrumentation helps mitigate bias related to misreporting.

I ran specifications using ordinary least squares models that assess whether inclusion in the sample population is related to basic demographic characteristics (age, gender, and indicators of race/ethnicity). Right-hand side variables were insignificant, indicating no evidence of selection bias from the baseline survey.

In addition to missing observations, I exclude individuals who were molested as children from the rape analysis. Respondents are quite young in Waves One and Two of the survey, suggesting that the incidence of rape may have resulted from molestation by a family member. In these situations, substance use is unconnected with victimization. To avoid misleading results, I exclude individuals from the analysis if they report having been molested by a primary caregiver as a child. In Wave Four, 288 female respondents indicated that a primary caregiver molested them. I exclude these individuals from the rape analysis.

C.1. Total number of sexual partners analysis and sexually transmitted disease analysis

The Aim One analysis is drawn from Waves One and Two. The initial sample population is comprised of 24,601 individuals. The sample size reduces to 20,322 after excluding those observations with incomplete information.

C.2.Rape analysis

The Aim Two analysis is drawn from Waves One, Two, and Four; the question about rape was not asked in Wave Three. The initial sample population in Wave One is comprised of 19,410 individuals. The sample size reduces to 15,658 after excluding those observations with incomplete information.

C.3. Labor market outcomes analysis

The Aim Three analysis of labor market outcomes are drawn from Wave Four and the substance use measures are drawn from Wave Three. Inclusion requires that an observation to have complete information in both of these waves. The initial sample population in Wave Four is comprised of 15,701 individuals. After dropping the remaining observations with incomplete information, the sample consists of 10,971 individuals.

D. Summary statistics by aim

D.1. Aim One summary statistics

Table 5 presents Aim One summary statistics for the total population and by gender. The average person in the population is roughly seventeen years old. The population is approximately equally divided between males and females. The average person has two sexual partners. A large number of individuals has zero partners (65%) and a smaller portion has one partner (7%) or two or more sexual partners (28%). Males are more likely to be sexually active and more likely to report two or more sexual partners.

Individual-reported STD status in the last year is low, approximately 2% for the total population. Females more commonly report a sexually transmitted disease diagnosis compared to males. This result may reflect the fact that sexually transmitted diseases are often less harmful and less noticeable for males (CDC, 2010).

The incidence of individual substance use varies by substance and gender. Males report higher levels of use in all categories studied. Thirty four percent of the sample are current smokers. For the population as a whole, approximately 31% of individuals report binge drinking in the last month, 34% experience any lifetime marijuana use, and 3%

experience any lifetime cocaine use, making marijuana and tobacco the most commonly used substances within the total population.

In terms of peer substance use, marijuana use is the least common form of reported peer substance use. Alcohol is the most common form of peer substance use, followed by cigarettes, and lastly marijuana.

D.2. Aim Two summary statistics

Table 6 presents Aim Two summary statistics for the population. Roughly 37% of the observations come from Wave One, 32% from Wave Two, and 31% from Wave Four. Women range between 15 and 35 years of age; the average woman is 20 years old. Twelve percent of the population reports ever being raped.

The incidence of individual substance use varies by substance in the population. Thirty four percent of women report binge drinking in the last month, 23% report any lifetime marijuana use, and 7% report any lifetime cocaine use. Most women, who are drawn from Waves One, Two, and Four, have at least one close friend who drinks at least once per month. A majority of individuals have not graduated from high school; a little under a quarter of the population has earned a four-year college degree. Roughly 17% of women are or have been married and 19% report having any biological children.

D.3. Aim Three summary statistics

Table 7 presents Aim Three summary statistics for the population. Eighty two percent of the population is employed. Conditional on being employed, the average person earns

approximately $17 per hour.[10] Wage rate spans a large range of values, between $0 and $961.[11]

While binge drinking and marijuana use are pervasive, methamphetamine use is relatively uncommon within the population. Fifty five percent of individuals report binge drinking in the last month, 34% of individuals report using marijuana in the last year, and 3% of individuals report using methamphetamines in the last year. Thirty six percent of the population smokes. Forty five percent of individuals report that three (out of three) of their best friends drink alcohol at least once per month.

The population is equally divided between men and women. The average person is 29 years in age. Fifty percent of individuals report having ever married and 42% of individuals report having any biological children. Thirteen percent of the population identifies themselves as African American and 12% identifies as Hispanic.

Educational attainment varies within the population. A majority of individuals report earning a high school (24%) or a four-year college degree (61%). A smaller proportion of individuals report earning less than a high school degree or earning a graduate degree. Fifteen percent of individuals report being current students.

---

[10] In May 2007, the average hourly wage rate was $19.56. Given that the Add Health population is younger than the general U.S. population, it is not surprising that the average wage rate is slightly less. (See Bureau of Labor Statistics: http://www.bls.gov/opub/mlr/2009/06/art2full.pdf.)

[11] Only 1.78% of employed individuals report a wage of zero. Based on calculation of wage rate, these individuals may work a very large number of hours for little reimbursement (See Aim Three Methods section). This group may include individuals in apprenticeships and internships, which lead to or serve as a necessary step towards full employment. Norton and Han (2008) treat zero wage earners as employed as well (the authors count students as employed). The 99th percentile wage rate is $67, far less than $961. Assuming that an individual works 40 hours per week and takes two weeks of vacation a year, individuals with a $67 wage rate earn roughly $134,000 annually. Assuming that an individual works 60 hours per week and takes two weeks of vacation a year, individuals with a $67 wage rate earn roughly $201,000 annually.

V. INSTRUMENTS

The efficaciousness of the instrumental variable approach relies largely on the validity of the instruments. Suitable instruments must be both strong (i.e., strong predictors of the endogenous variable) and validly excluded from the second stage equation (i.e., the variation in the instrumental variables must be independent from the variation in the dependent variable). I assess the instrument strength using an F-test on the set of instrumental variables in the first stage equation. The exclusion restriction tests vary depending on the type of instrumental variable approach used. I use two instrumental variable approaches in this analysis - two-stage least squares (2SLS) and two-stage residual inclusion (2SRI). The distinction between the 2SLS and 2SRI and the associated exclusion tests under each regime are discussed in more detail below. Given that the data are self-reported, misreporting of information may result in measurement error. The instrumental variable approaches mitigate bias arising from endogeneity and measurement error when the classical error-in-variance assumption holds (Wooldridge, 2006, pg. 530)[12].

[12] Take the following example. Assume the classical errors-in-variance assumption holds (i.e., measurement error is uncorrelated with unobserved explanatory variables (Wooldridge, 2006, pg. 322)). Suppose both x and z measure some latent variable x*. Further suppose that z is correlated with x but uncorrelated with the error term, meaning z can serve as an instrument for x. If z is uncorrelated with the error term, then, by necessity, z is uncorrelated with the measurement error (Wooldridge, 2006, pg. 530).

A. Instrumental variables

Many factors influence the congruence of substance use, peer substance use, and behavioral health, making the quantitative study of substance use and associated behaviors challenging. Researchers often turn to instrumental variable approaches to help mitigate endogeneity bias and establish causality. In order for instruments to be valid, they must be appropriately incorporated into empirical models. There are limitations to the instrumental variable approach. In some cases, an instrumental variable approach is inappropriate or infeasible (DeSimone, 2010; French and Popovici, 2010). Reflecting on the various models used to analyze the relationship between risky sexual behavior and binge drinking in particular, DeSimone (2010) proposes that a person-level fixed effects approach may provide the strongest estimates (pg. 6). Another drawback of 2SLS and 2SRI is that they inflate confidence intervals, which results in losses in precision.

Given the relative benefit of instrumental variable methodology in mitigating bias and establishing causality, instrumental variables are incorporated into all models. Nevertheless, for comparison, I provide estimates from models without instruments and fixed effects, models with only fixed effects, and models with fixed effects and instruments.

Tables 8 – 10 outline the instrumental variables used in all analyses, summary statistics for these variables, and the first stage equation (exogenous variables utilized in the first stage equations are discussed in detail in the data section). For all aims, the first stage model is the same. Instrument strength is tested with an F-test. The exclusion restrictions are tested using a LaGrange Multiplier (LM) test. Instrumental variables are sufficiently strong (F-stat > 10) and are validly excluded.

The LM test jointly tests whether all instruments are validly excluded. The null hypothesis states that instruments are valid. Rejection of the null hypothesis indicates that *some* of the instruments are invalid. The LM test is conducted by regressing the residuals from the second stage equations on the set of instruments, the exogenous variables, and a constant. The resulting R-squared statistic is multiplied by the number of observations to form the LM statistic.

A.1. Instruments for depression

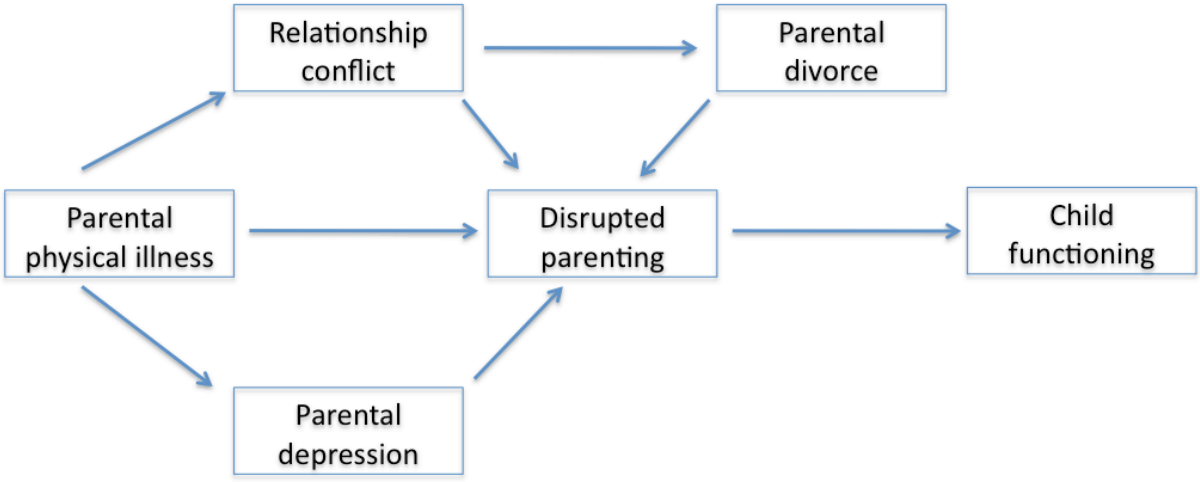
The two instrumental variables for depression gauge the happiness and physical health of the parental figure at the time of the first interview. Wave One of Add Health includes a parental in-home survey, which was preferably completed by the residential maternal guardian. Data collected from the survey indicate whether or not the parental figure is "generally happy" and whether the parental figure is in above average physical health.

In a study including three generations of respondents from a family unit (grandchildren, parents, and grandparents), Wickramaratne and Weissman (1998) find that parental depression is associated with the onset of major depressive disorder during childhood and early adulthood. Parental mental health, including parental history of major depression, generalized anxiety disorder, and antisocial disorder, has been used as an instrument for own psychiatric disorders in the literature (Ettner, Frank, and Kessler, 1997).

Although less obvious, parental physical health also plays a role in childhood depression. Armistead, Klein, and Forehand (1995)'s conceptual model outlines the impact of parental illness on child function. Based on this framework, outlined in Figure 1, the impact of a parental figure's physical health on child function and cognitive outcomes is, to

some degree, similar to the impact of parental depression, since physical illness is often associated with parental depression.

Fig 1. Armistead, Klein, and Forehand's schematic model of how parental illness influences child functioning



Therefore, I use both parental depression and physical health status as instruments. The instruments are sufficiently strong predictors of depression. I assume that parental indicators are to be validly excluded, since they do not reflect choice-level behaviors of an individual – an individual is born or adopted into a family and has little control over his or her parent's health or mental health. While genetic and environmental factors certainly play a role, for those on the margin, children with healthy parents may experience fewer depressive symptoms. Indeed, the parental health instruments may in part capture genetic and environmental factors.

A.2. Instruments for substance use

I utilize a variety of substance use instruments, including county-level crime rates and tract and state-level characteristics. I describe each set of instruments below. Each set of instruments meets the criteria for valid instruments. French and Popovici (2011) provide a

helpful review of existing substance use instruments, which I draw on in the discussion below.

Crime rates

County-level crime statistics have been used as an instrument for alcohol (Averett et al., 2004; Sen, 2002; Rees et al., 2001) and illicit drug use (Johansson et al., 2007; Wolaver, 2002) in research about adolescent populations. In this work, crime rates are measured for the respondents' county, specifically the rate of total violent crime arrests per 100,000 individuals in population. These crime statistics come from an individual's Crime Reporting Area, from the Uniform Crime Report. I assume the functional form of these instruments to be linear.

Tract-level characteristics

Wave Three of Add Health includes information related to the tract-level density of alcohol outlets, as well as other important characteristics about the census-tract in which an individual lives. Tract-level instrumental variables include the number of alcohol outlets, the proportion of the population moved in the past year, and the total square kilometer of the tract.

In previous works, researchers have instrumented for substance use using measures relating to the relative prevalence of alcohol within an individual's community. Within adolescent and young adult populations, previous instruments include whether or not an individual resides in a dry county (Kenkel and Ribar, 1994; Chatterji, 2006) and state sales of ethanol and alcohol (Sen, 2002).

State-level characteristics

Add Health contextual data has information on state-level cigarette prices and whether or not an individual currently attends a school in a state that requires alcohol and drug education, which varies across time and states. In previous work examining substance use in adolescent and young adult populations, measures relating to the cost of cigarettes have served as instruments for both cigarette (Dee, 1999) and alcohol use (Sen, 2002; Bray, 2005). In these works, researchers gauge the cost of cigarette use based on tax-level rather than overall cost. Rees and colleagues (2001) and Sen (2002) use the alcohol and drug education variable from Add Health as an instrument for substance use.

A.3. Implications for effects estimated with instrument variables

Instrumental variable estimation approaches help resolve these issues related to reverse causality and selection bias. Instrumental variable estimates reflect the local average treatment effect, which represents the effect of the instrumental variables on the marginal individual.

McClellan, McNeil, and Newhouse (1994) pioneered the use of instrumental variable approaches in public health. In their study examining the impact of catheterization on one-day mortality rates of acute myocardial infraction in elderly patients, the authors use distance from a catheterization facility as an instrument for catheterization. While relatively severe cases are more likely to receive invasive procedures and relatively less severe cases are not, the mode of treatment is unclear for the marginal patient (i.e., individuals on the borderline). Instrumental variable approaches help explain outcomes for this marginal group with the local average treatment effect – the marginal patient who is on the cusp of needing

catheterization is more likely to receive catheterization if they live closer to a catheterization facility.

The marginal interpretation can be applied here. Individuals on the margin are "at-risk" for becoming substance users or becoming depressed. For these individuals, the instrumental variables may induce or deter substance use behaviors and rouse beneficial or adverse depression outcomes. For instance, I use average cigarette prices, arrests per violent crime, and arrests per crime as instruments for any lifetime cocaine use. For those individuals on the margin, living in areas with low cigarette prices and higher crime rates may promote cocaine use. Each set of substance use and depression instruments helps explain the uptake of substance use and observed depression outcomes for the marginal individual, but do not explain the increase in substance use from other causes.

VI. AIM ONE: IMPACT OF SUBSTANCE USE ON RISKY SEXUAL BEHAVIOR – METHODS, RESULTS, AND DISCUSSION

Aim One analyzes the impact of substance use, depression, and peer substance use on risky sexual behavior. I define sexual risk-taking in two ways: (1) the number of reported sexual partners and (2) whether or not an individual reports having had a sexually transmitted disease in the last year. Aim One relies on three different substance use variables: any binge drinking in the last month, any lifetime marijuana use, and any lifetime cocaine use. The effect of substance use on risky sexual behaviors is analyzed separately by substance. The data are drawn from respondents aged 15 and up in Waves One and Two.

A. Methods

A.1 Methods - total number of sexual partners

The dependent variable counts an individual's reported number of sexual partners to date. Unlike a continuous variable, a count variable takes on only integer values. The resulting expected value function is nonlinear, making ordinary least squares (OLS) inappropriate. The distribution of count data is non-normal, becoming more so if the count data take on a smaller range of integer values. The non-normality of count data reinforces the inappropriateness of an OLS approach. Nonlinear least squares (NLS) is also unsuitable as it does not account for heteroskedasticity, which is always present in count data (Wooldridge, 2006, pg. 604-605). Unlike OLS and NLS, count models place no probability mass on

negative integer values and account for discrete dependent variables (Cameron and Trivedi, 2009, pg. 553). The starting point for most count data analyses is the Poisson model. After running the Poisson regressions in Aim One, the Vuong test demonstrated the presence of overdispersion of the dependent variable, suggesting that a negative binomial or zero-inflated negative binomial model better fits the data. Unlike Poisson models, the negative binomial and zero-inflated negative binomial models allow for overdispersion (i.e., they allow for the variance to exceed the mean). I next ran a series of negative binomial models. The results from the Vuong (1989) likelihood ratio test showed that the zero-inflated negative binomial model is preferrable to a negative binomial model.

Conceptually, the zero-inflated negative binomial model is useful when modeling a dependent variable with large number of zeros, particularly if the zeros can be grouped into two distinct categories. In the case of this analysis, one group of individuals has a zero probability of having a positive number of sexual partners. Under any circumstance, these individuals will have no sexual partners. The second group of zeros has some probability of having one or more sexual partners. Although the dependent variable is observed to be zero, these observations have some chance of having a positive number of sexual partners.

In the context of this work, ZINB makes theoretical sense since approximately 65% of individuals report zero sexual partners (the data for the analysis is drawn from individuals over the age of 15 in Waves One and Two). Furthermore, the zeros can be grouped into two categories. Some individuals will never, under any circumstances, have any sexual partners, which implies that this group will always have a zero probability of having any sexual partners. For example, some adolescents who take abstinence pledges may fall into this category. On the other hands, some adolescents with zero sexual partners may still have a

positive probability of having any partners. Of course, almost all individuals will have partners at some point in their lifetimes.

Since many individuals in the data set report no sexual partners, the zero-inflated negative binomial model may be preferable to the negative binomial framework. Modeling the zero generating process separately from the rest of the outcomes, the zero-inflated negative binomial model does a better job fitting the variance of the data, compared to the negative binomial model. I choose to use a zero-inflated negative binomial (ZINB) approach to model the lifetime number of sexual partners, since ZINB models account for both overdispersion as well as the large number of zeros in the dependent variable. While fixed effects are possible, the distributional assumptions in a fixed effects-ZINB setting are very restrictive and difficult to test. I choose not to incorporate fixed effects in the analysis.

The ZINB density function has two components, a binary density function and a count density function. The binary density function can take on a value of one or zero, and I modeled this process using a logit approach.

Two stage residual inclusion (2SRI) accounts for the likely endogeneity of both the depression and the substance use variables when the outcome model is nonlinear. 2SRI is the nonlinear equivalent of two stage least squares (2SLS), which is used for linear outcome models. Similar to 2SLS, 2SRI is a two-step estimation strategy. Unlike 2SLS, 2SRI provides consistent estimates in nonlinear settings (Terza, Basu, and Rathouz, 2008).

Equations (12) and (13) describe the first stage equations, where the endogenous variables are regressed on the instrumental and exogenous variables using OLS. The strength of these instruments is tested using an F-test.

(12) $$D = a_o + a_1 W_1 + a_2 P + \gamma_1 X + \nu_1$$

(13) $$S = \beta_o + \beta_1 W_2 + \beta_2 P + \gamma_2 X + \nu_2$$

S, *D*, and *P* refer to substance use, depressive symptoms, and peer substance use measures, respectively. *X* is a vector of other coefficients. The impact of peer substance use on sexual behavior depends both on the type of peer and individual substance use being examined.

In the second stage, the dependent variable, the total number of sexual partners, is regressed on the predicted residuals from the first stage equations, the endogenous variable, and the vector of exogenous variables. The second stage equation is described in equation (14).

(14) $$N = \delta_o + \delta_1 \hat{\nu}_1 + \delta_2 \hat{\nu}_2 + \delta_3 S + \delta_4 D + \delta_5 P + \gamma_3 X + \varepsilon$$

N, $\hat{\nu}_1$, $\hat{\nu}_2$, and ε refer to the total number of sexual partners, the predicted residual from equation (12) and (13), and the error term, respectively. Equation (14) is estimated using a ZINB model. Because the instruments are unrelated to the number of sexual partners, the predicted residuals effectively control for the endogenous variable that is correlated with the residuals in the equations, adjusting for the endogenous elements of *S* and *D* variables. A Wald test on $\hat{\nu}_1$ and $\hat{\nu}_2$ in the second stage equation tests the exogeneity of the instruments. The null hypothesis of exogeneity holds that $\hat{\nu}_1$ and $\hat{\nu}_2$ should be jointly equal to 0, demonstrating that the substance use and depression measures are unrelated with ε.

I calculate the average marginal effects on the predicted probability of having any, one, and four or more sexual partners for the population as a whole. Average marginal effects are calculated from coefficients from the 2SRI ZINB analysis. In order to account for the multi-stage process, I use the bootstrap method to calculate standard errors (500 repetitions).

Given that the range of the sexual partners variable is quite large, I run sensitivity tests to assess the influence of individuals with 100 or more partners on the results. The range

of the sexual partner variable spans from zero to 900; however, the right tail in this distribution is comprised of very few individuals – only 26 individuals reported having 100 or more partners (all male).

In sensitivity analyses, I exclude individuals who reported more than 100 partners from the analysis. The exclusion of these individuals from the sample population does not meaningfully change the results; coefficient estimates were well within one another's corresponding standard error using 95% confidence level.

A.2 Methods – sexually transmitted disease analysis

I model individual-reported STD status using a linear probability model (LPM) with individual fixed effects. Individual-reported STD status is measured by whether an individual reports having been diagnosed with an STD in the last year. While LPMs are encumbered by several deficits compared to their nonlinear counterparts,[13] they have several major strengths over nonlinear models. Importantly, unlike nonlinear models, an LPM approach allows for incorporation of fixed effects without restricting the composition of the sample population or the types of predictions that can be made.[14] In addition, fixed effects and instrumental variables cannot be simultaneously applied in logit and probit models. Thus, an instrumental variable model with fixed effects is better modeled using a linear approach. By modeling STD status with LPMs, I can compare results across instrumental variable and non-instrumental variable models.

[13] There are two pitfalls of LPM models: the predicted probabilities can fall outside the [0,1] range and LPMs are heteroskedastic.

[14] A fixed effect logit, often referred to as Chamberlain's fixed effect model, excludes observations if their outcome is constant over time. As a result, all individuals who never reported a STD would be dropped from the analysis, as would individuals who reported having a STD in every wave of the survey. The predicted probabilities and associated marginal effects from a Chamberlain's fixed effect model must be conditioned on a set number of positive outcomes (e.g. having reported having an STD only once).

I take several steps to mitigate the potential drawbacks of LPMs. Robust and clustered standard errors help address issues of heteroskedasticity. Interactions terms estimate the varying marginal effect of one covariate conditional on the value of another. In the final models, only 3% to 6% of observations have predicted probabilities that fall outside the [0,1] interval.

To account for both time invariant heterogeneity and omitted variable bias in the substance use and depression variables, the STD analysis includes instrumental variables, through two-stage least squares (2SLS), and school-level fixed effects (in both stages of the model). I account for the multi-stage process when computing the standard errors. I adjust the variance-covariance matrix by applying the correct mean squared error (Baltagi, 2002).

The linear probability model is specified as follows:

(15) $$R = a_o + a_1 S + a_2 D + a_3 P + \gamma X + \varepsilon$$

R, S, D, and P refer to STD status, substance use, depressive symptoms, and peer substance use measures, respectively. X is a vector of other coefficients. Substance use (S) and depression (D) are potentially correlated with the error term (ε). In the first stage, depression and substance use are regressed on the exogenous right hand side variables and respective sets of instruments. In the second stage, the dependent variable is regressed on all exogenous explanatory variables as well as the predicted values of the endogenous variables from the first stage regressions. I test the exclusion restriction with an LM test. The strength of the instruments is tested using an F-test on the vector of instruments used in the first stage. The relative fit of the model is captured through the R-squared statistic as well as the overall F-statistic for the model.

I use fixed effects at the school level. School-level fixed effects capture characteristics that do not vary within schools, but vary across schools. Unobservable school-level features like neighborhood safely, social mores, codes of conduct, patterns of activity, academic achievement, or other commonalities, are often shaped and reflected by one's school community. Social interaction theory, a fundamental basis for this work, reinforces the importance of these school-level factors (Akerlof, 1997). Moreover, this theory suggests that characteristics of the group, as reflected by one's school, are indicative of individual behavior. As a result, individual unobserved heterogeneity across schools is likely present and not fully captured by the peer substance use variables, since these variables only measure an individual's *estimates* of the substance use behavior of their three *closest friends*. The school community effect is likely broader, less tangible, and more nuanced.

School-level fixed effects have been previously applied when examining individual outcomes in the context social networks. For instance, Morgan and Sorensen (1999) used school-level fixed effects when examining student mathematical achievement in the context of parental networks.

I primarily choose to use school rather than individual-level fixed effects since school-level fixed effects reinforce the importance of peer behavior. Although both school- and individual-level fixed effects potentially reflect peer behavior, school-level fixed effects account for school-level heterogeneity. School-level fixed effects reflect community level behaviors, including time-invariant communalities among all school peers, including those individuals not identified as close friends. An additional benefit of school-level fixed effects is that time-invariant individual-level characteristics can be included in the STD analysis. Table 5 lists Aim One Descriptive statistics.

B. Results

B.1. Results from number of romantic partners analysis

Tables 11 through 13 compare results from the naïve zero-inflated negative binomial (ZINB) models that assume the exogeneity of substance use and depression with ZINB models with two-stage residual inclusion (2SRI). The tables also present the average marginal effects of right-hand side variables on the probability of having any sexual partners, one sexual partner, and two or more sexual partners.

Two-stage residual inclusion inflates confidence intervals, which results in lower precision. As seen in Tables 11 – 13, 2SRI estimates in some cases switch in sign from positive to negative and are often insignificant. This may result in part form the 2SRI estimation process. This may also result from the fundamental fact that there is no direct causal relationship between substance use, depression, and the number of sexual partners.

B.1.a. Substance use

The results presented in Tables 11 through 13 suggest that binge drinking, marijuana use, and cocaine use do not have a statistically significant effect on the number of sexual partners, after controlling for the endogeneity of substance use. Examining the average marginal effect of substance use on the predicted probability of having at least one, exactly one, and two or more sexual partners, the standard errors are quite large, and the 95% confidence intervals include zero.

B.1.c. Depression

While the correlation of depression and the number of sexual partners appears to be positive, the average marginal effect of depression on the probability of having any, one, or

two or more sexual partners is insignificant at conventional levels in the instrument variable models (Tables 11 through 13).

B.1.d. Peer substance use

Peer alcohol use

The average marginal effects presented in Tables 12 and 13 indicate that peer alcohol use is associated with an increase in the probability of having any sexual partners, one sexual partner, or two or more sexual partners. Both the marijuana and cocaine use 2SRI ZINB models predict that the absence of alcohol use among peers increases the predicted probability of having any sexual partners by 3.6 – 8.6 percentage points across models ($p < 0.01 – p < 0.05$) and decreases the predicted probability of having exactly one partner by 1.9 – 2.7 percentage points ($p < 0.01 – p < 0.05$). The 2SRI ZINB marijuana use model also finds that, holding other factors constant, individuals whose three closest friends *do not* drink at least once per month experience, on average, a 5.9 percentage-point decrease in the predicted probability of having two or more sexual partners, compared to individuals who have at least one friend who drinks regularly ($p < 0.05$).

Peer cigarette use

The average marginal effects presented in Tables 11 and 12 show that peer smoking is associated with an increase in the probability of having any sexual partners as well as the probability of having two or more sexual partners. Holding other factors constant, the 2SRI ZINB marijuana use models predict that individuals whose three closest friends *do not* smoke at least one cigarette per month experience, on average, a 4.9 percentage-point increase in the predicted probability of having no sexual partners ($p < .05$), compared to individuals who have at least one friend who smokes regularly. Similarly, the 2SRI ZINB binge drinking and

marijuana use models find that individuals whose three closest friends *do not* smoke at least one cigarette per month experience, on average, a 3.2 – 4.4 percentage-point decrease in the predicted probability of have two or more sexual partners ($p < .01 – p < .05$).

Peer marijuana use

Similar to peer alcohol and cigarette use, the average marginal effects presented in Tables 11 through 13 demonstrate an association between peer marijuana use and own sexual partners. Holding all else equal, the 2SRI ZINB binge drinking and cocaine use models predict that individuals whose three closest friends *do not* use marijuana at least once per month experience, on average, 10 to 16 percentage-point increase in the predicted probability of having zero sexual partners ($p < .01 – p < .05$). The 2SRI ZINB binge drinking model predicts that individuals whose three closest friends *do not* use marijuana at least once per month experience, on average, a 2.8 percentage-point decrease in the predicted probability of having exactly one sexual partner ($p < .05$). The 2SRI ZINB marijuana use and cocaine use models predict that individuals whose three closest friends *do not* use marijuana at least once per month experience, on average, a 12 – 16 percentage-point decrease in the predicted probability of having two or more sexual partners ($p < .01$). Compared to peer alcohol and cigarette use, the magnitude of the average marginal effects for peer marijuana use on own sexual partnering are relatively large.

B.2. Results from sexually transmitted disease analysis

I examine the relationship between substance use and risky sexual behaviors using an alterative but equally relevant measure of sexual risk-taking – whether or not an individual reports having been diagnosed with an STD in the last year. I run linear probability models with both school-level fixed effects and two-stage least squares.

Although substance use appears to be positively correlated with an STD diagnosis, the results presented in Table 14 show no evidence that either binge drinking or marijuana use has a statistically significant effect on reporting an STD diagnosis. While depression is positively associated with having a STD diagnosis in the naïve models, depression becomes insignificant once controlling for the endogeneity of this measure.

Finally, the peer substance use measures are also insignificant. Compared to the results from the number of sexual partners analysis, the impact of peer substance use on whether or not an individual reports having had an STD in the last year are less pronounced.

C. Discussion

Substance use and depression are insignificant after controlling for endogeneity. The results from the instrumental variable models indicate that the observed positive association in the naive models, which assume the exogeneity of substance use and depression, is *not causal*.

While the relationship is not causal, the results do not definitively demonstrate substance use, depression, and risky sexual behavior are unassociated. Instrumental variable estimates reflect the local average treatment effect, which represents the effect of the instrumental variables on the marginal individual. With the current set of instruments and among the Add Health population, the marginal substance user and the marginal depressed individual from the first stage equation do not experience different outcomes; however, among a different population and using different instruments, an association between substance use, depression, and risky sexual behaviors may exist.

For example, state cigarette taxes, an instrument for substance use, are designed to reduce smoking. While these taxes are aimed at cigarette use and not substance use in

general, cigarette taxes could feasibly generate positive externalities in terms of inhibiting a variety of substance use behaviors. While this work finds that cigarettes do not have these protective effects for adolescents and young adults in the Add Health population, cigarette taxes may have protective effects for marginal individuals in other populations.

While the literature is far from consistent, my findings support the work of Grossman, Kaestner, and Markowitz (2004) and Grossman and Markowitz (2005), who argue that a causal relationship between substance use and risky sexual behavior is unlikely. They suggest that Jessor and Jessor's (1977) problem behavior theory may be a key piece of the puzzle. This analysis supports the argument that unobservable time-variant and invariant personal- and community-level factors jointly shape substance use, depression, and sexual behavior.

Social environment is likely an such a factor, which shapes observed risky sexual behavior – the peer variables and school-level fixed effects included in the this analysis emphasize the importance of the social and community effects. Indeed, the results from the sexual partner analysis suggest that peer influences are strong and diverse. For instance, while all forms of peer substance use significantly impact the individual number sexual partners, the magnitude of the average marginal effect of peer marijuana use far exceeds that of alcohol and cigarettes.

While previous work has found a relationship between depression and sexual behavior, in this analysis, depression is insignificant. I believe that the distinction between previous work and my own is largely caused by differences in the treatment of endogenous variables and the integration of substance use into the analysis. Unlike much of the existing literature, I control for the endogeneity of depression. I also control for both depression and substance use, since

the impact of these two outcomes on sexual behavior is likely co-occurring (e.g., NIMH, CSAT).

Of course, the results do not imply that adolescents should not be screened for substance abuse and depression. These conditions in themselves present important public health challenges. Substance use, depression, and issues surrounding reproductive health can be problematic, if not dangerous, for adolescents and young adults, especially if they are slow to seek care (Monroe, 2005).

As Akerlof's (1997) social interactive theory hypothesized and as the peer substance use variables in my analysis demonstrate, one's social group and community are important predictors of own behavior. In developing a treatment strategy aimed at reducing adverse sexual outcomes, providers may consider developing a treatment plan that accounts not only for individual physiological conditions but also treats the person as a function of his or her environment.

Perhaps health care providers cannot impact a patient's social environment, but they can use information on social environment to make inferences about a patient's health and health risks (e.g., risk of having a large number of sexual partners). Treatment outcomes and preventative care may be enhanced if providers take into account a patient's social environment. How many of a patient's peers regularly drink or use drugs like marijuana or cocaine? Are a patient's peers sexually active? If so, do they use condoms or birth control? (Peer sexual behavior provides a fascinating opportunity for expanding this research.) The answers to these questions not only reflect on the health behavior of peers but also that of the patient.

The results of this study also lend credence to the school- and community-based health clinic movement, popularized in the 1990s (Monroe, 2005). If an individual's reproductive health is, at least in part, a product of her peers and community, it makes sense to treat the individual in the context of her social environment. Moreover, school- and community-based health clinics offer providers the opportunity to observe commonalities in reproductive health issues, as well as social behaviors that support or inhibit these observed health outcomes. As a result, providers working in school- and community-based clinics may be able to devise more comprehensive treatment plans and preventative care strategies.

This work makes several contributions to the existing literature. I capture social network effects through peer behavior variables and school-level fixed effects. As mentioned, substance use often co-occurs with mental health conditions. Previous analyses often omit indicators of mental health, and, therefore, are likely suffer from omitted variable bias. I control for both substance use and depression, defining substance use in a variety of ways. I also apply new and more appropriate empirical modeling techniques. Ordinary least squares is inappropriate for count dependent variables. For the number of sexual partners analysis, I use a count data model.

VII. AIM TWO: IMPACT OF SUBSTANCE USE ON THE PROBABILITY OF HAVING EVER BEEN RAPED – METHODS, RESULTS, AND DISCUSSION

Aim Two analyzes the impact of substance use and peer substance use on the probability of rape victimization. The data are drawn from women in Waves One, Two, and Four. Aim Two relies on three different substance use variables: any binge drinking in the last month, any lifetime marijuana use, and any lifetime cocaine use. The effect of substance use on risky sexual behaviors is analyzed separately by substance.

A. Methods

I model rape victimization with linear probability models (LPM)[15]. The rape variable is a binary measure, equaling one if an individual reports having ever been raped at the time of the interview. Similar to the STD analysis in Aim One, I incorporate both school-level fixed effects and two-stage least squares (2SLS) to account for endogeneity in the depression and substance use variables. The modeling procedure in Aim Two is virtually identical to the STD analysis. For more information about the LPM model, fixed effects, and 2SLS, please refer to the methods section of Aim One in the previous chapter.

[15] Ever raped can also be modeled using a hazard model. Hazard models examine the time that passes until an event occurs. While the approach makes sense since "ever raped" is a terminal event, the right hand side variables do not necessarily explain the timing of rape. As a result, I choose not to use a hazard model approach.

I tested the sensitivity of my Aim Two analysis by dropping individuals with previous reported rape. This resulted in a loss of 353 observations, or just over 2% of the sample. Dropping these individuals did not significantly change the outcome. Importantly, depression, substance use, and peer measures are still insignificant in the fixed effects-instrumental variable models.

B. Results

Aim Two tests the impact of substance use on the probability of having ever been raped using three categories of models: (1) linear probability models; (2) linear probability models with school-level fixed effects; and (3) linear probability models with both school fixed effects and instrumental variables using two-stage least squares. Tables 15 through 17 detail findings from each analysis.

Substance use has an insignificant impact on the probability of having ever been raped, once accounting for the endogeneity of binge drinking, marijuana use, and cocaine. In fact, the binge drinking measure is always insignificant, regardless of fixed-effect or endogeneity corrections. Although positive and significant in naïve models that assume the exogeneity of depression, this measure is insignificant once accounting for endogeneity through instrument variables analysis.

Similar to own substance use, peer substance is largely insignificant. The remaining exogenous explanatory variables are also insignificant. In fact, the are R-squared for the analyses are quite low, ranging from .06 to .10, suggesting that the right-hand side variables, which include indicators of age, race/ethnicity, and other demographic and behavioral variables, are not strong predictors of rape.

C. Discussion

Few researchers have examined the impact of substance use and depression on rape; this work represents a novel approach to this topic. The results from the naïve models, which assume the exogeneity of the substance use and depression variables, demonstrate that drug use (i.e., marijuana and cocaine use) and depression are positively correlated with the rape victimization. While substance use and depression are likely endogenous in all aims, the directionality of the relationships between substance use, depression, and rape is especially unclear. Do substance use and depression lead to rape or does the relationship run in the opposite direction?

Rape victims often experience depression following the event, suggesting that rape may lead to depression (e.g., Frank and Stewart, 1984). Alcohol and drugs are also a form of self-medication (Khantzian, 1985; Khantzian 1997). Instrumental variables help clarify whether alcohol and drugs actually lead to rape. The results from the linear probability analyses with fixed effects and two-stage least squares estimation provide a more accurate depiction of the impact of substance use and depression. The Aim Two findings suggest that substance use and depression *do not lead to* rape victimization.

Fagan (1993) hypothesized that substance use not only increases the probability of victimizing an individual but also increases the probability of being victimized. The results from this analysis, however, suggest that substance use may not be a causal factor. While depression has been shown to be a result of victimization (e.g., Frank and Stewart, 1984), the findings provide no evidence that depression leaves adolescents at greater risk for rape.

Why might the Aim Two findings differ from Fagan's hypothesis? In the rape analysis, I define substance use as in terms of any binge drinking in the last month and any

lifetime marijuana and cocaine use. Perhaps the hypothesized relationship between victimization and substance use holds with higher frequencies of substance use. For instance, researchers may find that a clinical measure of substance abuse is indeed positively associated with rape and sexual assault – this relationship warrants further study. While unavailable in Waves One and Two, Waves Three and Four of Add Health contain information on both a clinical depression diagnosis and whether an individual has been prescribed with antidepressants and anxiety medication. Future analyses could utilize this information to assess the impact of a *severe* depression diagnosis or depression/anxiety *treatment* on rape. This analysis is focused on a younger sample population and a broader measure of depression as gauged by the Feeling Scale.

The sensitivity of the dependent variable may also drive the observed differences between the empirical results and from Fagan's hypothesis. Not only is rape self-reported, but rape is also a highly personal outcome. Many women may not feel comfortable sharing this information; moreover, the actual definition of rape is subjective. While one woman might believe that a certain act qualifies as rape, another women might not. The situation becomes more muddled in the presence of substance use. As a result, some of the individuals who reported no victimization in the Add Health data might have, in fact, been victimized.

Because the data are self-reported and because rape itself is a sensitive topic, rape is subject to measurement error. What are the consequences of measurement error in this analysis? As long as the residual has mean of zero and is uncorrelated with the covariates, ordinary least squares (OLS) will produce unbiased estimates (Wooldridge, 2006, pg 319). Therefore, even if rape is measured with error, it likely does not bias our estimates of the key variables but may result a loss of precision.

The Aim Two results indicate that it is difficult to predict potential rape victims based on *observable* characteristics. Unlike Aim One, social network effects are also insignificant – peer substance use is not predictive of rape victimization. The Aim Two findings suggest that it is difficult for policymakers, law enforcement officials, and healthcare providers to identify at-risk groups. As with many aspects of health care, providers need to find culturally relevant ways to inform and educate young women about rape and sexual assault, since victims come from an array of backgrounds.

VIII. AIM THREE: IMPACT OF SUBSTANCE USE ON LABOR MARKET OUTCOMES – METHODS, RESULTS, AND DISCUSSION

Aim Three analyzes the impact of substance use, depression, and peer substance use on the probability of employment and wage rate conditional on employment. The data are drawn from Wave Four, except for substance use measures, which are drawn from Wave Three. Aim Three relies on three different measures of lagged substance use: any binge drinking in the last month, any marijuana use in the last year, and any methamphetamine use in the last year. The effect of substance use on risky sexual behaviors is analyzed separately by substance.

A. Methods

Aim Three tests the impact of substance use on the probability of being employed and on wage rate conditional on being employed using two categories of models: (1) two-part models and (2) two-part models with two-stage least squares (2SLS). The labor market outcomes of interest include employment status and wage rate, conditional on being employed. The sample population for this analysis is drawn from Wave Four. Lagged substance use enters into the model by using Wave Three measures of substance use.

I use a two-part model to analyze the relationship of substance use and labor market outcomes. Equation (16) models the probability that an individual works, using a linear probability model (LPM). Equation (17) models wage rate conditional on being employed

using ordinary least squares (OLS). I define an individual as employed if she or he is working at least 10 hours a week at the time of the interview. Norton and Han (2008), who used Add Health to examine the impact of obesity on employment, define employment in the same manner.

(16) $$P(Employ) = \alpha_o + \alpha_1 S + \alpha_2 D + \alpha_3 P + \gamma_1 X + \varepsilon_1$$

(17) $$\ln(wage) = \beta_o + \beta_1 S + \beta_2 D + \beta_3 P + \gamma_2 X + \varepsilon_2 \quad if\ employ = 1$$

In equation (16), *Employ* is a binary indicator of whether an individual is employed. *S* is the lagged value of substance use from Wave Three. *D* refers to current depressive symptoms. *X* is a vector containing the remaining control variables, such as demographic factors.

Wage refers to wage rate. Wooldridge's pseudo R-squared suggests that the wage variable should be logged. The remaining right-hand side variables in equation (17) mirror those in equation (16).

Endogeneity with respect to depression and substance use may be problematic. In the case of substance use, reverse causality cannot be present, since only lagged values of substance are used. Nevertheless, lagging values do not eliminate the threat of endogeneity. Two-stage least squares (2SLS) instrumental variable approach helps address the endogeneity in the substance use and depression measures. I account for the multi-stage process when computing the standard errors. I adjust the variance-covariance matrix by applying the correct mean squared error (Baltagi, 2002).

The Breusch-Pagan test examines the model error for heteroskedasticity. The results show that the error term from the wage equation is heteroskedastic. I adjust for heteroskedasticity in both parts of the model using robust standard errors.

B. Results

Tables 18 through 20 detail findings from each analysis. The results suggest that substance use does not affect the observed labor market outcomes (Tables 18 – 20). Regardless of definition, lagged individual substance use is largely insignificant in both the naïve models that assume the exogeneity of substance use, as well as more sophisticated models with instrumental variables.

Depression appears to be negatively associated with labor market outcomes; however, the relationship does not seem to be causal (Tables 18 – 20). While significant in the naïve models, depression becomes insignificant once controlling for endogeneity. Although the effect size of depression increases in the 2RSI analyses (i.e., the magnitude increases), the coefficient is insignificant (i.e., the confidence intervals include zero). As discussed, 2SLS results in a loss of precision. The magnitude of the coefficient on depression highlights the importance of its relative impact on the employment outcomes. Depression may be negatively associated with employment despite the results; however, the 2SLS process may obscure this relationship.

Not surprisingly, the relative impact of peer alcohol use varies across models. In the binge drinking and methamphetamine 2SLS models, peer alcohol use is insignificant (Tables 18 and 20). In the marijuana use 2SLS models, however, peer alcohol use does have a significant and positive impact on employment (Table 19). All else equal, having two out of three best friends drink is associated with a 6.8% increase in the predicted probability of employment, compared to individuals whose friends abstain from alcohol. Similarly, having three out of three best friends drink alcohol is associated with a 7.2% increase in the predicted probability of employment. The results suggest that the impact and relative

significance of peer substance use on own labor outcomes is diverse and depends not only on the definition of own substance, but also on the definition of labor market outcomes.

C. Discussion

Even in the naïve models that do not control for endogeneity, lagged substance use is not significantly associated with labor market outcomes. While depression appears to be negatively correlated with both the probability of employment and wage rate, depression is insignificant in all instrumented models. Therefore, the results imply that lagged substance use and depression *do not cause* the observed employment and wage rate outcomes.

While previous research finds a link between substance use and labor market outcomes (e.g. DeSimone, 2002; French et al., 2001; Buchmuellar and Zuekas, 1998), the literature is far from consistent. van Ours (2006) and MacDonald and Pundey (2000), for instance, find limited evidence of such a relationship.

So how does the Aim Three analysis fit into this body of work? The results suggest that the *durational effects* of substance use on employment are limited. There is limited research with which to compare the results of this work, since the durational effects remain under-examined. The closest comparison that can be made is with van Ours (2006), which concludes that, for both males and females, the evidence linking the age of onset of cannabis and cocaine use with adverse labor market outcomes is weak. Nevertheless, age of onset and lagged substance use are quite different, making the comparison of this work with van Ours (2006) tenuous.

While this work focuses on *lagged* indicators of substance *use*, the current empirical framework may yield different results if substance *use* is redefined to substance *abuse*. Moreover, this analysis focuses on young adults – perhaps older individuals experience the

impact of lagged substance use on labor market outcome differently. These areas offer opportunities for future research.

A large body of research details the relationship between labor market outcomes and mental health (e.g., Handbook of Health Economics), including the effect of depression specifically (e.g., Ettner, Frank, and Kessler, 1997). Why was no evidence of such a relationship found in the Aim Three analysis? Dooley and colleagues (1994) emphasize that the relationship between depression and employment runs in both directions. Does depression lead to underemployment or does underemployment result in depression? The bi-directionality of this link reinforces the need for an instrumental variable approach, since instruments mitigate the impact of reverse causality. Consequently, the results from this work will differ from those analyses that do not control for the endogeneity of depression (while some analyses account for the endogeneity of depression [e.g., Ettner et al., 1997], not all research does). What is more, much of the previous work does not control for the co-occurring effects of substance use and depression, suggesting that some analyses may suffer from omitted variables bias.

The results from this analysis may also differ from existing work, since the Aim Three analysis does not rely on a clinical measure of depression. Perhaps the relationship between labor market outcomes and depression would hold among a population diagnosed with severe depression and other acute psychiatric disorders.

Incorporating a measure of psychiatric disorders into the current framework presents an opportunity for future research. For both men and women, Chatterji and colleagues (2011) find evidence that psychiatric disorders are associated with reductions in labor force participation, although the authors find that women are more susceptible to selection effects.

Both van Ours (2006) and Chatterji and colleagues (2011) emphasize the importance of selection effects. Observable and unobservable personal and community-level factors influence own behavior. Peer behavior is likely one such factor. In Aim Three, the impact of peer alcohol use on labor market outcomes, and its relative significance, depends on a variety of factors, including the definition of individual substance use and labor market outcomes. Unlike the binge drinking and methamphetamine models, peer alcohol use has a *significant* and *positive* impact on the probability of employment (but does not significantly impact wage rate).

Work place functions, like office happy hours, may promote alcohol use. If these types of events are common in one's workplace, drinking, or at least associating with people who drink, may help individuals find, maintain, and advance in the their jobs. Nevertheless, this effect is dependent on employment.

The analysis does find some evidence that individuals whose peers drink are more likely to work. Literature emphasizes the importance of secondary ties in employment. In particular, literature suggests that social networks play an important role in young people finding their first job (Marmaros and Sacerdote, 2002). These social networks likely have diverse effects across subgroups. For instance, research suggests that, among college students, social network effects on employment vary along gender and racial lines (Marmaros and Sacerdote, 2002).

The influence of peer effects is also described in the literature. Akerlof's (1997) social interaction theory describes a conceptual framework in which peer behavior impacts individual decision-making, highlighting the idea that individual substance use and employment behaviors are the result of both an individual and social decision-making

process. Grossman and colleagues (1994), who analyze the impact of price changes on substance use behaviors of youth, cite Rachal and colleagues (1980), stating, "a rise in price would curtail youth consumption directly and indirectly through its impact on peer consumption" (pg. 351). Pertold (2010) and Waddell (2010) further discuss the impact of peer substance use on individual risky sexual behaviors of youth.

Taken together, current literature highlights: (1) the role of secondary ties and the direct social network effects on employment and (2) the direct social network effects of peer substance use on individual substance use. This work illustrates a further connection between social networks substance use behavior and the probability of individual employment. The mechanism linking peer substance use and employment remains unclear. Does peer substance use directly or indirectly impact employment? If peer substance use and employment are indirectly related, does the observed relationship found in this analysis merely shadow the underlying relationship between social networks in general and employment? Future research may consider modeling both peer substance use and peer employment to identify the distinct effect of each variable.

This work indicates that *peer* substance use at least has some impact on employment. Employers and policymakers may want to make use of this relationship to help identify and curb individual substance use behaviors. "Generations, like people, have personalities," Scott Keeter and Paul Taylor from Pew Research write, and "[…] America's newest generation, the Millennials, is in the middle of this coming-of-age phase of its life cycle." By understanding the patterns of substance use and the role networking effects play in promoting both substance use and employment outcomes among Millennials, employers, policymakers, and

insurers may all be able to experience an increase in productivity and profit as well as an improvement in individual health outcomes of youth.

IX. DISCUSSION

I begin this section by considering the results of this work within the broader context of the literature. I next discuss the contributions and limitations of the findings. I end by outlining future research.

Literature relating substance use, depression, and risky sexual behavior is inconclusive in its causal attribution. As such, the findings from this work are both concordant and discordant with the existing research. While Cheeson and colleagues (2000) find evidence that substance use and sexually transmitted disease are positivity associated, Tubman and colleagues (1996) suggest that sexual partnering and psychological wellbeing are negatively associated, the sexual partners and the STD analyses come to the opposite conclusion, suggesting that substance use, depression, and these risky sexual behaviors are not causally related.

Nevertheless, my findings are consistent with some existing literature, which suggests that substance use *does not have a direct causal effect* on sexual risk taking among young people (e.g., Grossman et al., 2004). The link between substance use, depression, and sexual behaviors observed by some researchers may be explained by unobserved individual behavioral and social characteristics.

Similar to the results from the risky sexual behaviors analysis, the results from the rape analysis suggest that substance use and depression *do not lead to* rape. The findings are contrary to Fagan's (1993) hypothesis, which suggests that substance use increases the potential for victimization by violent crime.

One might hypothesize that substance use and depression are positively related with rape victimization. However, issues of reverse causality and temporality make any such analysis difficult. For instance, the psychology literature documents the potential for depression *following* rape (Frank and Stewart, 1984). Moreover, substance use may serve as a form of self-medication following rape. As a result, instrumental variable methods are important in parsing out the directionality of this relationship.

Mirroring the risky sexual behavior and rape analyses, the results from the employment analysis hint that lagged substance use and depression *do not cause* the observed employment and wage rate outcomes. The results suggest that the *durational effects* of substance use on employment are limited, supporting somewhat related work of van Ours (2006), which indicates that age of onset of substance use has limited effects on labor market outcomes. Given the dearth of literature investigating durational effects over many years, more research is needed.

In the large part, peer effects contribute to the novelty of this work. Each aim tests the importance of peer effects. Indeed, the results suggest that peer substance use behavior is an important correlate of sexual partnering and, perhaps, employment. Taken as a whole, the results lend credence to Akerlof (1997)'s social interaction theory – peer behaviors do shape individual choices.

A. Study contributions

This work makes several major contributions. Many studies do not concurrently control for substance use and depression, which could result in omitted variable bias since the two outcomes are often linked. I fill in this gap by controlling for both substance use and depression in all analyses. Instrument variable techniques help account for the endogeneity of

these measures. I also draw from a broad definition of substance use and sexual behavior. While previous research examines the onset of depression following rape, few have investigated the impact of depression on victimization. Finally, the durational effects of substance use on employment remain under-examined; this work fills this gap by analyzing the impact of substance use on employment of young adults multiple years after the event.

B. Limitations

The major limitations revolve around two issues: (1) potentially endogeneity of right-hand side variables, and (2) the generalizability of the results. In addition, this work does not address the role of substance abuse or depression *treatment*.

Although this work models two main constructs (depression and substance use) as endogenous, there are clearly other measures that could be considered endogenous in these models as well. Body mass index and smoking are potentially endogenous, given that these variables are possibly correlated with both the outcome of interest and the error term.

Treating an endogenous variable as exogenous may produce biased estimates. The direction of this bias depends on the relationship of the variable with the error term and the dependent variable.

While it is impossible to instrument for everything, I have applied instruments to the variables most susceptible to endogeneity (i.e., substance use and depression). Moreover, the extensive set of explanatory variables likely mitigates the potential for omitted variable bias.

Instrumental variable estimates reflect the effect of the instrumental variables on the marginal individual. With the current set of instruments and among the Add Health population, the marginal substance user and the marginal depressed individual do not experience different outcomes. Therefore, the results do not definitively demonstrate that substance use, depression, and risky sexual behavior are unassociated. With a different set of instruments or within a different population, significant relationships may be found.

Add Health is a nationally representative survey of adolescents and young adults. While the findings can be generalized to these groups, the results *do not* reflect the behavior of adults or youths in custodial care (e.g., hospitals and juvenile detention centers). Young people who are institutionalized in behavioral health and detention facilities likely experience stronger depression and substance use symptoms (e.g., Cuellar, Markowitz, and Libby, 2004). As a result, the impact of depression and substance use on sexual risk-taking, rape, and employment is likely distinct across groups in and out of custodial care. The relationship between substance use, depression, and the dependent variables of interest (i.e., risky sexual behavior, rape, and employment) among institutionalized adolescents and young adults remains unexamined in the current framework; therefore, I cannot comment on the existence and directionality of these associations. Nevertheless, the impact of substance use and

depression on these outcomes among young adults with a history of severe mental illness or delinquent behaviors presents an excellent opportunity for future research.

Finally, while this work examines the impact of substance use and depression on the outcomes of interest, the effect of treatment remains unexamined. The link between *treatment* for substance abuse and depression and the outcomes of interest remains unexamined in the current framework and warrants further study. While substance use and depression likely *do not cause* the observed changes in outcomes interest, the impact of treatment may be altogether different. Indeed, individuals receiving treatment are likely distinct from individuals who engage in substance use but never undergo substance abuse treatment, or depressed individuals who manifest symptoms but never seek counseling or medication. Moreover, treatment may alter the way an individual approaches everyday behaviors, including employment and sexual activity. Therefore, the impact of treatment on the outcomes of interest is likely distinct from the effect of substance use or being depressed.

C. Future research

This work presents several opportunities for future research. Future analyses may consider examining the association of substance use and depression with sexual and labor market behaviors among more severely mentally ill populations or using an alternative mental health diagnosis. While the association of substance use and depression with sexual and labor market outcomes is weak within this general adolescent and young adult population, perhaps the relationship is stronger for youths in custodial care (e.g., inpatient mental health facilities and juvenile detention centers).

Future research may also consider using the same framework to investigate these relationships in older populations. Older individuals exhibit different sexual, employment,

and substance use behaviors, suggesting that the relationships among these outcomes may be drastically different.

Lastly, peer effects, which are included in all analyses, contribute to the novelty of this work. Future research may consider using alternative measures of peer behavior, such as indicators of sexual behavior, depression, and employment.

Table 1. Overview of Aim One research question, methodology, and sample population

Research Question	Does substance use have a causal impact on the report number of sexual partners?	Does substance use have a causal impact on reported one-year sexually transmitted disease status?
Dependent variable	Number of individuals with whom an individual reports being sexually involved	Binary measure of whether an individual reports a sexually transmitted disease diagnosis in the last 12 months
Analytical Model	Zero-inflated negative binomial model	Linear probability model with school-level fixed effects
Substance use measures – binge drinking [b]	Any binge drinking in the last month	Any binge drinking in the last month
Substance use measures – drug use [b]	Any lifetime marijuana use Any lifetime cocaine use	Any lifetime marijuana use
Depression measures	Percentile score from the Feeling Scale	Percentile score from the Feeling Scale
Peer substance use measures	Three measures are included. These measures are based on three questions from Waves One and Two of the Add Health survey Of your three best friends, how many (1) Smoke at least one cigarette a day? (2) Drink alcohol at least once a month? (3) Use marijuana at least once a month?	Gauged by a series of dummy variables that measure peer substance use. Add Health asks, "Of your three best friends, how many drink alcohol at least once peer month?"
Waves	One and two	One and two
Weights and survey correction?	Yes	Yes
Number of observations [c]	20,322	20,322

a. The wave one measure of the number of lifetime sexual partners will be defined as the total number of partners an individual has ever had. Wave two asks about the number of partners since the last wave. Therefore, the cumulative wave two measure of total number of partners ever will be constructed by adding the wave one partners measure to the wave two partners measure.
b. Models will be run separately by each unique measure of the types and frequency of substance use.
c. Individuals under the age of 15 are not included in any analyses. I exclude this group since a majority of individuals under the age of 15 are not sexually active and most cannot legally work under current federal law.

Table 2. Overview of Aims Two and Three research questions, methodologies, and sample populations

Aim	Two	Three
Research Question	Does substance use have a causal impact on the probability of ever being a victim of rape?	Does substance use have a durational effect on later labor market outcomes?
Dependent variable	An indicator of whether or not an individual has ever been a victim of rape	Employment status and wage rate
Analytical Model	Linear probability model with school-level fixed effects	Two part model
Substance use measures – binge drinking [a]	Any binge drinking in the last month	Any binge drinking in the last month
Substance use measures – drug use [a]	Any lifetime marijuana use Any lifetime cocaine use	Any marijuana use in the last year Any methamphetamine use in the last year
Depression measures [b]	Percentile score from the Feeling Scale	Percentile score from the Feeling Scale
Peer substance use measures	A series of dummy variables that measure peer substance. I developed the measure from a question in Add Health that asks, "Of your three best friends, how many drink alcohol at least once a month?"	A series of dummy variables that measure peer substance. I developed the measure from a question in Add Health that asks, "Of your three best friends, how many drink alcohol at least once a month?"
Waves	One, two, and four	Four (lagged substance use measures are drawn from Wave Three)
Weights and survey correction?	Yes	Yes
Number of observations [c]	15,658	9,091

a. Models will be run separately by each unique measure of the types and frequency of substance use.
b. 9-item abridged depression is used in Aims 2 and 3.
c. Individuals under the age of 15 are not included in any analyses. I exclude this group since a majority of individuals under the age of 15 are not sexually active and most cannot legally work under current federal law.

Table 3. Depression measures: 19-item Add Health Feeling Scale

Respondents are asked how often was each of the following things true during the past seven days. An individual response can range from "never or rarely" to "most of the time" or "all of the time."

You were bothered by things that usually don't bother you. *

You didn't feel like eating, or your appetite was poor.

You felt that you could not shake off the blues, even with help from your family and your friends. *

You felt that you were just as good as other people. *

You had trouble keeping your mind on what you were doing. *

You felt depressed. *

You felt that you were too tired to do things. *

You felt hopeful about the future.

You thought your life had been a failure.

You felt fearful.

You were happy.

You talked less than usual.

You felt lonely.

People were unfriendly to you.

You enjoyed life. *

You felt sad. *

You felt that people disliked you. *

It was hard to get started doing things.

You felt life was not worth living.

* Indicates that these questions were also included in questions included in all waves

Table 4. Number of missing of observations by variable and aim

Variable	Aim 1 Risky sexual behavior	Aim 2 Ever raped at time of the interview	Aim 3 Labor market outcome
Total number of partners	656	--	--
Sexually transmitted disease status	91	--	--
Rape	--	653	--
Worked at least 10 hours per week	--	--	310
Wage	--	--	0
Any binge drinking in the last month	319	151	--
Any lifetime marijuana use	550	217	--
Any lifetime cocaine use	747	509	--
Any binge drinking in the last month (lagged)	--	--	2,910
Any cocaine use in the last year (lagged)	--	--	2,856
Any methamphetamines use (lagged)	--	--	2,827
Peer alcohol use	466	1,588	3,143
Peer cigarette use	438	--	--
Peer marijuana use	456	--	--
Current smoker	345	173	137
Attractiveness	94	27	28
Body Mass Index	553	509	221
Current student	6	8	4
Married	20	31	19
Any children	2	112	0
Highest level of education completed			
High school	0	2	4
4-year college	--	2	4
Graduate school	--	2	4
Age	13	7	0
Male	0	--	0
African American	44	85	29
Hispanic	60	111	46
Initial number of observations	*24,601*	*17,682*	*15,701*
Unweighted final number of observation:	*21,946*	*15,101*	*11,678*
Weighted final number of observations:	*20,322*	*14,379*	*10,971*

Notes: (1) The values are non-sequential. Each cell indicates the number of missing values associated with a variable. Observations are dropped if they have any missing values for the variables of interest. (2) The lagged substance use measures in Aim Three are drawn from Wave Three (all other variables are drawn from Wave Four). (3) -- indicates that variable not included in specified analysis

Table 5. Aim one descriptive statistics

Variables	Weighted			Unweighted
	Mean (Female)	Mean (Male)	Mean (Total)	Mean (Total)
Dependent variable				
Total number of sexual partners	1.16 (0.069) [0, 987]	2.60 (0.19) [0, 102]	1.89 (0.11) [0, 987]	1.96 (13.38) [0, 987]
Proportion with 0, 1, and 2 or more sexual partners				
0 partners	70%	74%	65%	
1 partner	6%	5%	7%	
2 or more partners	24%	21%	28%	
Any STD diagnosis in the last year	0.029 (0.0034)	0.0096 (0.0014)	0.019 (0.0020)	0.021 (0.14)
Endogenous explanatory variables				
Binge drinking	0.28 (0.011)	0.34 (0.016)	0.31 (0.012)	0.29 (0.45)
Marijuana use	0.32 (0.014)	0.35 (0.014)	0.34 (0.013)	0.34 (0.47)
Cocaine use	0.023 (0.0027)	0.034 (0.0044)	0.028 (0.0028)	0.028 (0.16)
Exogenous explanatory variables				
None of best friends drink alcohol at least once per month	0.38 (0.012)	0.39 (0.014)	0.38 (0.012)	0.40 (0.49)
None of best friends smoke cigarettes at least once per month	0.49 (0.017)	0.49 (0.012)	0.49 (0.012)	0.52 (0.50)
None of best friends use marijuana at least once per month	0.60 (0.013)	0.60 (0.013)	0.60 (0.011)	0.61 (0.49)
Current smoker	0.34 (0.017)	0.33 (0.012)	0.33 (0.013)	0.30 (0.46)
Interviewer-rated attractiveness				
Below average	0.044 (0.0037)	0.063 (0.0058)	0.054 (0.0036)	0.05 (0.23)
Average	0.39 (0.010)	0.51 (0.011)	0.45 (0.0089)	0.44 (0.50)
Above average	0.56 (0.011)	0.42 (0.011)	0.49 (0.0089)	0.50 (0.50)
Body mass index	22.7 (0.12) [12.86, 63.49]	23.1 (0.11) [13.25, 54.79]	22.9 (0.090) [12.86, 63.49]	22.93 (4.58) [12.86, 63.49]

Table 5. Aim one descriptive statistics

Variables	Weighted			Unweighted
	Mean (Female)	Mean (Male)	Mean (Total)	Mean (Total)
Current student	0.95 (0.0045)	0.95 (0.0045)	0.95 (0.0039)	0.95 (0.21)
Ever married	0.010 (0.0022)	0.0026 (0.00078)	0.0064 (0.0013)	0.01 (0.081)
Any children	0.038 (0.0045)	0.0057 (0.0013)	0.022 (0.0024)	0.02 (0.15)
Completed high school	0.032 (0.0041)	0.029 (0.0042)	0.030 (0.0039)	0.03 (0.18)
Age	16.6 (0.063) [15, 22]	16.7 (0.069) [15, 22]	16.7 (0.064) [15, 22]	16.82 (1.32) [15, 22]
African American	0.17 (0.024)	0.15 (0.021)	0.16 (0.022)	0.22 (0.41)
Hispanic	0.11 (0.019)	0.12 (0.019)	0.12 (0.018)	0.17 (0.38)
Wave one indicator	0.37 (0.0098)	0.35 (0.0089)	0.36 (0.0089)	0.37 (0.48)
Number of Observations	10,391	9,931	20,322	21,946

Standard errors in parentheses. Range of continuous variables in brackets.
Note: The Aim One number of partners analysis is a cross-sectional analysis drawing from individuals in Waves One and Two who are ages 15 and older.

Table 6. Aim two descriptive statistics

Variables	Mean (Weighted)	Mean (Unweighted)
Dependent variable		
Rape	0.12 (0.0054)	0.11 (0.32)
Endogenous explanatory variables		
Binge drinking	0.34 (0.011)	0.31 (0.46)
Marijuana use	0.39 (0.013)	0.37 (0.48)
Cocaine use	0.069 (0.0043)	0.06 (0.24)
Exogenous explanatory variables		
Number of best friends who drink alcohol at least once per month		
Zero of three	0.34 (0.0095)	0.36 (0.48)
One of three	0.23 (0.0052)	0.24 (0.42)
Two of three	0.18 (0.0045)	0.17 (0.38)
Three of three	0.25 (0.0094)	0.23 (0.42)
Current smoker	0.33 (0.014)	0.29 (0.46)
Interviewer-rated attractiveness		
Below average	0.056 (0.0035)	0.056 (0.23)
Average	0.41 (0.0089)	0.40 (0.49)
Above average	0.54 (0.0095)	0.55 (0.50)
Body mass index	24.4 (0.13) [3.66, 72.17]	24.38 (6.14) [3.66, 72.17]
Current student	0.72 (0.0067)	0.73 (0.44)
Ever married	0.17 (0.0054)	0.16 (0.37)
Any children	0.19 (0.0071)	0.18 (0.39)

Table 6. Aim two descriptive statistics

Variables	Mean (Weighted)	Mean (Unweighted)
Highest level of education completed		
Less than high school	0.68 (0.0077)	0.69 (0.46)
High school	0.085 (0.0040)	0.085 (0.28)
4-year college	0.20 (0.0069)	0.18 (0.39)
Graduate school	0.033 (0.0029)	0.036 (0.19)
Age	20.4 (0.056) [15, 34]	20.37 (5.55) [15, 34]
African American	0.16 (0.022)	0.24 (0.43)
Hispanic	0.12 (0.019)	0.16 (0.36)
Wave two indicator	0.32 (0.0035)	0.32 (0.47)
Wave four indicator	0.31 (0.0088)	0.30 (0.46)
Number of Observations	14,379	15,101

Standard errors in parentheses. Range of continuous variables in brackets.
Note: The Aim Two analysis draws on women in Waves One, Two, and Four who are ages 15 and older and incorporates school-level fixed effects.

Table 7. Aim Three Descriptive Statistics

Variables	Mean (Weighted)	Mean (Unweighted)
Dependent variable		
Employed	0.82 0.01	0.82 (0.38)
Wage if employed > 0 (in thousands of dollars)	17.36 (0.39) [0, 961.53]	18.25 (25.45) [0, 961.53]
Endogenous explanatory variables		
Binge drinking in the last month (lagged from wave three)	0.55 (0.014)	0.49 (0.50)
Marijuana use in the last year (lagged from wave three)	0.34 (0.0099)	0.32 (0.47)
Methamphetamine use in the last year (lagged from wave three)	0.03 (0.0025)	0.03 (0.17)
Depression percentile	0.43 (0.0066) [0, 1]	0.45 (0.30) [0, 1]
Exogenous explanatory variables		
Number of best friends who drink alcohol at least once per month		
Zero of three	0.21 (0.010)	0.24 (0.43)
One of three	0.18 (0.0066)	0.19 (0.39)
Two of three	0.16 (0.0058)	0.16 (0.37)
Three of three	0.45 (0.014)	0.40 (0.49)
Current smoker	0.36 (0.0096)	0.35 (0.48)
Interviewer-rated attractiveness		
Below average	0.062 (0.0043)	0.07 (0.26)
Average	0.46 (0.010)	0.47 (0.50)
Above average	0.48 (0.011)	0.46 (0.50)
Body mass index	28.23 (0.16) [6.45, 72.17]	28.36 (6.96) [6.45, 72.17]

Table 7. Aim Three Descriptive Statistics

Variables	Mean (Weighted)	Mean (Unweighted)
Current student	0.15 (0.0056)	0.17 (0.37)
Ever married	0.50 (0.014)	0.50 (0.50)
Any children	0.42 (0.015)	0.45 (0.50)
Highest level of education completed		
Less than high school	0.06 (0.0056)	0.07 (0.25)
High school	0.24 (0.012)	0.25 (0.43)
4-year college	0.61 (0.012)	0.59 (0.49)
Graduate school	0.09 (0.0077)	0.09 (0.29)
Age	28.91 (0.12) [25, 34]	29.06 (1.75) [25, 34]
Male	0.51 (0.0078)	0.45 (0.50)
African American	0.13 (0.017)	0.15 (0.36)
Hispanic	0.12 (0.018)	0.21 (0.40)
Number of Observations	9.056	11,678

Standard errors in parentheses. Range of continuous variables in brackets.
Note: The Aim Three analysis draws on individuals from Wave Four except for the lagged substance use measures, which are drawn from Wave Three.

Table 8. Tests of instrument strength and exclusion restriction (first-stage equations are the same across all aims)

		Instrument strength in first stage	*Second stage exclusion restrictions* [1]				
Endogenous Variable	**Instruments**	**F-Stat**	**Aim 1a**	**Aim 1b**	**Aim 2**	**Aim 3 – employ**	**Aim 3 – wage**
Binge drinking	-Tract-level number of alcohol outlets -Average cigarette price -Currently attending a high school in a state that requires alcohol and drug education -County-level arrests per violent crime	F(4, 43,682) = 12.02 Prob>F = 0.00	P = 0.7986	P = 0.9999	P = 0.6351	P = 1.0000	P = 0.9999
Any marijuana use (ever)	-Tract-level number of alcohol outlets -State-level average cigarette price -Currently attending a high school in a state that requires alcohol and drug education	F(3, 42,370) = 33.33 Prob>F = 0.00	P = 0.9974	P = 0.9884	P = 0.9028	--	--
Any cocaine use (ever)	-State-level average cigarette price -County-level arrests per violent crime -County-level arrests per crime	F(3, 42,729) = 35.74 Prob>F = 0.00	P = 0.9989	P = 0.1849	P = 0.9999	--	--
Any marijuana use (year)	-Tract-level the proportion moved in the past years -Tract-level total area per square kilometer -County-level average cigarette price	F(3, 11,987) = 2.74 Prob>F = 0.00	--	--	--	P = 0.9996	P = 0.4447
Any meth-amphetamine use (year)	-County-level arrests per violent crime -County-level arrests per violent crime -County-level arrests per juvenile violent crime -County-level arrests per juvenile crime	F(3, 12014) = 13.34 Prob>F = 0.00	--	--	--	P = 0.8641	P = 0.9768
Depression score percentile	-Is parental figure generally happy -Is parental figure in above average physical health	F(2, 44250) = 54.11 Prob>F = 0.00	P = 0.9962	P = 0.8456	P = 0.9999	P = 0.9527	P = 0.9999

Note: The instrumental variables used are the same across all aims. For example, in Aims One, Two, and Three, the instruments for binge drinking include: the tract-level number of alcohol outlets, the average cigarette price, whether the individual is currently attending a high school in a state that requires alcohol and drug education, and the county-level arrests per violent crime.

Table 9. Weighted descriptive statistics for instrumental variables (first-stage equations are the same across all aims)

	Mean	Linearized Standard errors	95% confidence interval	
State-level average cigarette price	278.42	5.25	268.03	288.81
Student in state that requires alcohol and drug education	0.31	0.02	0.27	0.35
Proportion moved in the past five years (Tract-level densities)	0.67	0.05	0.58	0.77
Total area per square kilometer (Tract-level densities)	2246.86	645.40	969.83	3523.89
Total number of alcohol outlets (Tract-level densities)	105.75	22.25	61.74	149.77
Total arrests per 100,000 (County-level)	3179.61	151.36	2880.12	3479.11
Violent crime arrests per 100,000 (County-level)	634.09	71.78	492.06	776.12
Juvenile violent crime arrests per 100,000 (County-level)	56.19	6.02	44.26	68.11
Parental figure generally happy	0.85	0.01	0.83	0.86
Parental figure in above average general health	0.24	0.01	0.22	0.26

Table 10.First-stage regressions by substance-type and depression percentile

	(1) Any binge drinking in the last month	(2) Any lifetime marijuana use	(3) Any lifetime cocaine use	(4) Any marijuana use in the last year	(5) Any metham-phetamine use in the last year	(6) Depressions score percentile
Instrumental variables						
Total number of alcohol outlets (Tract-level densities)	0.00000045	-0.000044**				
	(0.0000081)	(0.0000086)				
State-level average cigarette price	0.000031	0.000048*	0.000030*	0.00041**		
	(0.000022)	(0.000023)	(0.000012)	(0.00010)		
State requires alcohol and drug education when in high school	-0.0032	-0.055**				
	(0.0062)	(0.0066)				
Violent crime arrests per 100,000	-0.0000092**		0.000033**		0.000028**	
	(0.0000013)		(0.0000035)		(0.0000055)	
Juvenile violent crime arrests per 100,000			-0.00040**		-0.00039**	
			(0.000040)		(0.000069)	
Proportion moved in the past five years (Tract-level densities)				0.078**		
				(0.023)		
Total area per square kilometer (Tract-level densities)				-0.0000093**		
				(0.0000020)		

Table 10.First-stage regressions by substance-type and depression percentile

	(1) Any binge drinking in the last month	(2) Any lifetime marijuana use	(3) Any lifetime cocaine use	(4) Any marijuana use in the last year	(5) Any metham-phetamine use in the last year	(6) Depressions score percentile
Juvenile crime arrests per 100,000					0.000083**	
					(0.000018)	
Parental figure generally happy						-0.015**
						(0.0048)
Parental figure in above average general health						0.025**
						(0.0042)
Exogenous variables						
One of an individual's three best friends drinks alcohol at least once per month	0.15**	0.12**	0.015**	0.12**	0.015**	0.033**
	(0.0056)	(0.0061)	(0.0032)	(0.011)	(0.0037)	(0.0040)
Two of an individual's three best friends drinks alcohol at least once per month	0.27**	0.20**	0.023**	0.18**	0.0100*	0.046**
	(0.0065)	(0.0068)	(0.0039)	(0.012)	(0.0039)	(0.0044)
Three of an individual's three best friends drinks alcohol at least once per month	0.41**	0.30**	0.073**	0.31**	0.022**	0.022**
	(0.0055)	(0.0058)	(0.0037)	(0.0097)	(0.0034)	(0.0038)

Table 10.First-stage regressions by substance-type and depression percentile

	(1) Any binge drinking in the last month	(2) Any lifetime marijuana use	(3) Any lifetime cocaine use	(4) Any marijuana use in the last year	(5) Any metham-phetamine use in the last year	(6) Depressions score percentile
Current smoker	0.20**	0.33**	0.12**	0.21**	0.032**	0.061**
	(0.0049)	(0.0049)	(0.0038)	(0.0092)	(0.0038)	(0.0032)
Very unattractive/ unattractive	-0.042**	-0.022*	-0.0078	-0.019	0.0011	0.058**
	(0.0086)	(0.0088)	(0.0058)	(0.016)	(0.0060)	(0.0059)
Average attractiveness	-0.022**	-0.013**	-0.0023	-0.012	0.0046	0.035**
	(0.0042)	(0.0044)	(0.0029)	(0.0082)	(0.0031)	(0.0029)
Current student	0.0049	-0.026**	-0.018**	-0.0018	0.011*	-0.016**
	(0.0059)	(0.0062)	(0.0043)	(0.011)	(0.0042)	(0.0040)
Ever married	-0.058**	-0.068**	-0.045**	-0.080**	-0.015**	-0.040**
	(0.0070)	(0.0072)	(0.0053)	(0.0093)	(0.0035)	(0.0046)
Children	-0.073**	0.036**	-0.00050	-0.033**	-0.00024	0.014**
	(0.0069)	(0.0071)	(0.0051)	(0.0093)	(0.0034)	(0.0045)
High school	0.051**	-0.040**	-0.041**	-0.013	0.0069	-0.047**
	(0.0079)	(0.0083)	(0.0065)	(0.018)	(0.0078)	(0.0054)
4-year college	0.090**	-0.032**	-0.054**	-0.013	-0.0074	-0.100**
	(0.010)	(0.010)	(0.0081)	(0.017)	(0.0073)	(0.0066)
Graduate/professional degree	0.10**	-0.097**	-0.12**	-0.056**	-0.018*	-0.14**
	(0.016)	(0.017)	(0.011)	(0.021)	(0.0077)	(0.010)
Age	0.068**	0.055**	0.017**	-0.080	0.017	-0.00045
	(0.0069)	(0.0072)	(0.0046)	(0.067)	(0.026)	(0.0048)
Age squared	-0.0015**	-0.0012**	-0.00041**	0.0011	-0.00031	0.000021
	(0.00015)	(0.00015)	(0.00010)	(0.0012)	(0.00045)	(0.00010)

Table 10.First-stage regressions by substance-type and depression percentile

	(1) Any binge drinking in the last month	(2) Any lifetime marijuana use	(3) Any lifetime cocaine use	(4) Any marijuana use in the last year	(5) Any metham-phetamine use in the last year	(6) Depressions score percentile
Male	0.039*	-0.030	-0.097**	-0.079	-0.017	0.0055
	(0.018)	(0.019)	(0.012)	(0.13)	(0.051)	(0.013)
Male*age	0.0014	0.0032**	0.0054**	0.0038	0.00100	-0.0010
	(0.00084)	(0.00086)	(0.00059)	(0.0044)	(0.0017)	(0.00056)
African American	-0.14**	0.031**	-0.070**	-0.012	-0.022**	0.041**
	(0.0051)	(0.0055)	(0.0028)	(0.0100)	(0.0029)	(0.0036)
Hispanic	0.0087	0.030**	0.0072	-0.028*	-0.0000018	0.044**
	(0.0060)	(0.0060)	(0.0042)	(0.011)	(0.0044)	(0.0039)
Wave one indicator	-0.070**	-0.20**	-0.19**	.	.	-0.072**
	(0.019)	(0.020)	(0.014)	.	.	(0.013)
Wave two indicator	-0.089**	-0.17**	-0.21**	.	.	-0.075**
	(0.018)	(0.019)	(0.013)	.	.	(0.012)
Wave three indicator	-0.050**	-0.099**	-0.099**	.	.	-0.042**
	(0.012)	(0.012)	(0.0096)	.	.	(0.0080)
Constant	-0.56**	-0.24**	0.027	1.41	-0.23	0.49**
	(0.083)	(0.086)	(0.052)	(0.97)	(0.38)	(0.059)
Observations	43708	42395	42754	12009	12036	44274
R-squared	0.27	0.25	0.12	0.17	0.030	0.040

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$

Note: Regressions (1) through (6) show the results from the first stage regressions. The first stage regressions are common across all analyses (i.e., regardless of the aim, the first stage equations are the same for identical definitions of substance use).

Table 11. Impact of any binge drinking in the last month on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 sexual partners)	Pr(1 sexual partners)	Pr(2 or more sexual partners)
Binge drinking	-1.30**	0.055	4.31	1.55	0.28	-0.13	-0.14
	(0.18)	(0.072)	(4.22)	(2.24)	(0.39)	(0.12)	(0.32)
Depression percentile	-1.21**	-0.13	-1.79	0.41	-0.19	0.036	0.16
	(0.21)	(0.13)	(2.86)	(1.95)	(0.24)	(0.090)	(0.21)
None of an individual's three best friends drinks alcohol	0.24	0.22	1.32	0.51	0.083	-0.042	-0.041
	(0.14)	(0.14)	(0.80)	(0.45)	(0.076)	(0.022)	(0.063)
None of an individual's three best friends smokes	0.18	-0.25*	0.18	-0.22	0.032	-0.00026	-0.032*
	(0.14)	(0.11)	(0.23)	(0.15)	(0.018)	(0.0074)	(0.015)
None of an individual's three best use marijuana	0.51**	-0.18*	1.14*	0.022	0.10*	-0.028*	-0.074
	(0.11)	(0.087)	(0.50)	(0.31)	(0.047)	(0.014)	(0.039)
Current smoker	-1.30**	0.28*	-2.83*	-0.17	-0.24*	0.072*	0.17
	(0.18)	(0.11)	(1.21)	(0.62)	(0.11)	(0.033)	(0.093)
Very unattractive/unattractive	0.38	0.0013	0.60*	0.032	0.052*	-0.015	-0.037
	(0.21)	(0.18)	(0.29)	(0.24)	(0.025)	(0.0096)	(0.023)
Average attractiveness	0.0060	-0.020	0.071	-0.019	0.0079	-0.0014	-0.0065
	(0.10)	(0.096)	(0.14)	(0.11)	(0.011)	(0.0047)	(0.010)
Body mass index	0.0097	0.026	0.0077	0.025	-0.0011	-0.00066	0.0018
	(0.012)	(0.016)	(0.013)	(0.016)	(0.0011)	(0.00052)	(0.0012)

Table 11. Impact of any binge drinking in the last month on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 sexual partners)	Pr(1 sexual partners)	Pr(2 or more sexual partners)
Current student	1.27**	-0.19*	1.13**	-0.19	0.12**	-0.024**	-0.093**
	(0.27)	(0.085)	(0.30)	(0.12)	(0.027)	(0.0084)	(0.021)
Ever married	-1.63	-0.37	-1.63	-0.36	-0.12	0.047	0.075
	(1.88)	(0.24)	(2.14)	(0.26)	(0.19)	(0.054)	(0.14)
Any children	-2.94*	0.037	-2.49	0.094	-0.23	0.059	0.17*
	(1.43)	(0.13)	(1.39)	(0.19)	(0.12)	(0.035)	(0.087)
High school graduate	0.31	-0.24	-0.017	-0.27	0.018	0.0056	-0.024
	(0.40)	(0.19)	(0.51)	(0.21)	(0.041)	(0.015)	(0.030)
Age	-2.42**	0.44	-2.78**	0.40	-0.28**	0.060*	0.22**
	(0.89)	(0.65)	(0.95)	(0.63)	(0.087)	(0.029)	(0.073)
Age squared	0.056*	-0.010	0.064*	-0.0100	0.0066**	-0.0014	-0.0052*
	(0.027)	(0.019)	(0.028)	(0.018)	(0.0025)	(0.00084)	(0.0021)
Male	-3.95**	-0.27	-0.66	0.49	-0.096	0.0067	0.089
	(1.45)	(1.09)	(3.03)	(1.61)	(0.28)	(0.084)	(0.23)
Male*age	0.20*	0.045	-0.015	-0.0064	-0.00090	0.00049	0.00041
	(0.088)	(0.064)	(0.19)	(0.100)	(0.018)	(0.0053)	(0.015)
African American	-1.64**	0.49**	-1.08*	0.59*	-0.14**	0.015	0.13**
	(0.18)	(0.14)	(0.50)	(0.24)	(0.045)	(0.014)	(0.035)
Hispanic	0.51*	0.40	0.48	0.35	0.018	-0.018	0.00038
	(0.23)	(0.21)	(0.33)	(0.26)	(0.026)	(0.011)	(0.023)

Table 11. Impact of any binge drinking in the last month on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 sexual partners)	Pr(1 sexual partners)	Pr(2 or more sexual partners)
Wave one indicator	0.13	0.0095	-0.065	-0.052	-0.0020	0.0026	-0.00054
	(0.094)	(0.053)	(0.18)	(0.097)	(0.016)	(0.0053)	(0.013)
Residuals (Binge drinking)			0.59	-0.53			
			(2.86)	(2.02)			
Residuals (Depression)			-5.62	-1.50			
			(4.24)	(2.24)			
Constant	24.6**	-3.95	26.8**	-4.21			
	(7.60)	(5.68)	(8.17)	(5.34)			
Ln(α)	0.80**		0.81**				
	(0.11)		(0.11)				
Observations	15,709	15,709	15,709	15,709	15,709	15,709	15,709

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$

Note: The Aim One number of partners analysis is cross-sectional, drawing on individuals from individuals from Waves One and Two who are ages 15 and older.

Table 12. Impact of any lifetime marijuana use on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 partners)	Pr(1 partner)	Pr(2 or more partners)
Any lifetime marijuana use	-1.02**	0.37**	8.99	-2.11	1.02	-0.20	-0.82
	(0.16)	(0.14)	(7.34)	(2.41)	(0.68)	(0.20)	(0.51)
Depression percentile	-1.11**	-0.16	-0.98	0.68	-0.14	0.012	0.13
	(0.22)	(0.13)	(2.80)	(2.06)	(0.25)	(0.097)	(0.21)
None of an individual's three best friends drinks alcohol	0.48**	0.19	0.96*	0.081	0.086*	-0.027*	-0.059*
	(0.15)	(0.14)	(0.41)	(0.19)	(0.034)	(0.012)	(0.026)
None of an individual's three best friends smokes	0.19	-0.24*	0.33	-0.23	0.049*	-0.0040	-0.044**
	(0.14)	(0.11)	(0.25)	(0.16)	(0.021)	(0.0087)	(0.017)
None of an individual's three best use marijuana	0.29*	-0.071	3.71	-0.82	0.42	-0.082	-0.33*
	(0.12)	(0.093)	(2.49)	(0.82)	(0.23)	(0.069)	(0.17)
Current smoker	-1.18**	0.23	-4.28	0.95	-0.48*	0.095	0.39*
	(0.16)	(0.13)	(2.35)	(0.79)	(0.21)	(0.065)	(0.16)
Very unattractive/unattractive	0.51*	0.063	0.67*	-0.021	0.065*	-0.017	-0.048
	(0.21)	(0.19)	(0.30)	(0.25)	(0.029)	(0.010)	(0.027)
Average attractiveness	0.0081	0.0023	0.053	-0.088	0.012	0.00033	-0.012
	(0.10)	(0.098)	(0.14)	(0.12)	(0.011)	(0.0054)	(0.010)
Body mass index	0.0042	0.024	-0.0051	0.020	-0.0019	-0.00026	0.0022
	(0.012)	(0.015)	(0.012)	(0.015)	(0.0011)	(0.00053)	(0.0012)

Table 12. Impact of any lifetime marijuana use on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 partners)	Pr(1 partner)	Pr(2 or more partners)
Current student	0.95**	-0.17*	1.78**	-0.34	0.20**	-0.041*	-0.15**
	(0.24)	(0.083)	(0.66)	(0.24)	(0.063)	(0.018)	(0.048)
Ever married	-1.44	-0.31	-0.28	-0.58	0.016	0.019	-0.035
	(2.26)	(0.24)	(3.23)	(0.40)	(0.30)	(0.089)	(0.21)
Any children	-3.39	-0.037	-3.71	0.12	-0.36	0.096	0.27
	(2.49)	(0.12)	(2.66)	(0.22)	(0.24)	(0.069)	(0.17)
High school graduate	0.26	-0.21	0.27	-0.17	0.038	-0.0038	-0.034
	(0.35)	(0.17)	(0.42)	(0.20)	(0.034)	(0.014)	(0.025)
Age	-1.83*	0.37	-4.83*	0.97	-0.53**	0.11	0.42**
	(0.89)	(0.66)	(2.17)	(0.98)	(0.19)	(0.063)	(0.15)
Age squared	0.039	-0.0085	0.12*	-0.025	0.014*	-0.0028	-0.011**
	(0.027)	(0.019)	(0.061)	(0.028)	(0.0054)	(0.0018)	(0.0041)
Male	-3.30*	-0.038	-0.90	-1.38	0.015	0.051	-0.066
	(1.46)	(1.15)	(2.34)	(1.27)	(0.22)	(0.073)	(0.18)
Male*age	0.16	0.030	-0.0087	0.11	-0.0092	-0.0020	0.011
	(0.088)	(0.068)	(0.15)	(0.078)	(0.014)	(0.0047)	(0.011)
African American	-1.41**	0.49**	-1.81**	0.58**	-0.22**	0.037**	0.18**
	(0.17)	(0.14)	(0.41)	(0.17)	(0.034)	(0.012)	(0.026)
Hispanic	0.54*	0.41	0.12	0.50	-0.026	-0.013	0.039
	(0.24)	(0.21)	(0.47)	(0.31)	(0.042)	(0.016)	(0.034)

Table 12. Impact of any lifetime marijuana use on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 partners)	Pr(1 partner)	Pr(2 or more partners)
Wave one indicator	0.033	0.013	0.031	-0.0033	0.0032	-0.00076	-0.0025
	(0.082)	(0.050)	(0.097)	(0.052)	(0.0082)	(0.0032)	(0.0061)
Residuals (Marijuana use)			-0.11	-0.76			
			(2.80)	(2.12)			
Residuals (Depression)			-9.96	2.53			
			(7.37)	(2.41)			
Constant	19.9**	-3.60	41.4*	-7.93			
	(7.53)	(5.72)	(15.9)	(7.55)			
Ln(α)	0.75**		0.75**				
	(0.12)		(0.12)				
Observations	16,124	16,124	16,124	16,124	16,124	16,124	16,124

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$

Note: The Aim One number of partners analysis is cross-sectional, drawing on individuals from individuals from Waves One and Two who are ages 15 and older.

Table 13. Impact of any lifetime cocaine use on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 sexual partners)	Pr(1 sexual partners)	Pr(2 or more sexual partners)
Any lifetime cocaine use	-1.42**	0.76**	5.79	7.54	0.022	-0.31	0.29
	(0.37)	(0.13)	(23.7)	(13.2)	(2.17)	(0.78)	(1.69)
Depression percentile	-1.14**	-0.14	-0.87	0.12	-0.095	0.021	0.074
	(0.23)	(0.12)	(2.65)	(1.99)	(0.25)	(0.094)	(0.21)
None of an individual's three best friends drinks alcohol	0.52**	0.22	0.53**	0.23	0.036**	-0.019**	-0.017
	(0.15)	(0.15)	(0.16)	(0.16)	(0.012)	(0.0060)	(0.011)
None of an individual's three best friends smokes	0.20	-0.23*	0.27	-0.16	0.039	-0.0040	-0.035
	(0.14)	(0.11)	(0.28)	(0.18)	(0.025)	(0.0097)	(0.020)
None of an individual's three best use marijuana	0.64**	-0.11	0.91	0.15	0.080	-0.028	-0.052
	(0.11)	(0.084)	(0.84)	(0.46)	(0.076)	(0.027)	(0.059)
Current smoker	-1.39**	0.28*	-1.59*	0.10	-0.16**	0.041	0.12**
	(0.17)	(0.11)	(0.65)	(0.37)	(0.058)	(0.021)	(0.045)
Very unattractive/unattractive	0.43*	-0.065	0.39	-0.11	0.047	-0.0083	-0.038
	(0.20)	(0.18)	(0.24)	(0.23)	(0.024)	(0.0090)	(0.023)
Average attractiveness	-0.0058	-0.063	-0.028	-0.083	0.0034	0.0024	-0.0058
	(0.11)	(0.090)	(0.13)	(0.11)	(0.012)	(0.0050)	(0.010)
Body mass index	0.00033	0.022	0.000039	0.022	-0.0016	-0.00044	0.0020
	(0.012)	(0.016)	(0.012)	(0.015)	(0.0011)	(0.00056)	(0.0012)

Table 13. Impact of any lifetime cocaine use on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 sexual partners)	Pr(1 sexual partners)	Pr(2 or more sexual partners)
Current student	1.09**	-0.19*	1.21**	-0.076	0.13**	-0.031*	-0.095**
	(0.27)	(0.085)	(0.46)	(0.21)	(0.044)	(0.014)	(0.033)
Ever married	-1.28	-0.48	-1.42	-0.53	-0.10	0.049	0.054
	(4.07)	(0.31)	(3.91)	(0.31)	(0.37)	(0.11)	(0.26)
Any children	-2.70	0.022	-2.90	-0.11	-0.28	0.081	0.20
	(1.91)	(0.15)	(1.95)	(0.30)	(0.18)	(0.053)	(0.13)
High school graduate	0.29	-0.25	0.30	-0.24	0.047	-0.0033	-0.044
	(0.36)	(0.19)	(0.44)	(0.22)	(0.038)	(0.015)	(0.027)
Age	-2.25*	0.47	-2.11*	0.59	-0.25**	0.045	0.21**
	(0.90)	(0.64)	(0.99)	(0.66)	(0.097)	(0.033)	(0.079)
Age squared	0.051	-0.012	0.047	-0.016	0.0058*	-0.00094	-0.0048*
	(0.027)	(0.019)	(0.030)	(0.019)	(0.0029)	(0.0010)	(0.0024)
Male	-3.22*	-0.95	-3.07*	-0.80	-0.25	0.099	0.15
	(1.43)	(1.07)	(1.54)	(1.08)	(0.13)	(0.056)	(0.11)
Male*age	0.15	0.080	0.14	0.070	0.0087	-0.0052	-0.0036
	(0.087)	(0.064)	(0.096)	(0.066)	(0.0083)	(0.0035)	(0.0066)
African American	-1.44**	0.54**	-1.38**	0.61**	-0.18**	0.025	0.16**
	(0.18)	(0.13)	(0.37)	(0.22)	(0.034)	(0.013)	(0.026)
Hispanic	0.54*	0.47*	0.39	0.32	0.015	-0.017	0.0021
	(0.24)	(0.22)	(0.52)	(0.31)	(0.050)	(0.017)	(0.041)

Table 13. Impact of any lifetime cocaine use on the number of sexual partners – Results from naive and 2SRI ZINB models

	(1)		(2)		(3)	(4)	(5)
	ZINB		2SRI ZINB		Average marginal effects from 2SRI analysis		
	Any sexual partners	Number sexual partners among those with any	Any sexual partners	Number sexual partners among those with any	Pr(0 sexual partners)	Pr(1 sexual partners)	Pr(2 or more sexual partners)
Wave one indicator	0.041	-0.038	-0.093	-0.16	0.0027	0.0058	-0.0085
	(0.083)	(0.052)	(0.43)	(0.24)	(0.039)	(0.014)	(0.031)
Residuals (Cocaine use)			-7.20	-6.79			
			(23.7)	(13.2)			
Residuals (Depression)			-0.28	-0.26			
			(2.64)	(2.06)			
Constant	23.0**	-4.07	21.5*	-5.47			
	(7.60)	(5.63)	(8.96)	(5.78)			
Ln(α)	0.72**			0.72**			
	(0.13)			(0.12)			
Observations	15,547	15,547	15,547	15,547	15,547	15,547	15,547

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$

Note: The Aim One number of partners analysis is cross-sectional, drawing on individuals from individuals from Waves One and Two who are ages 15 and older.

Table 14. Impact of substance use on the probability any that an individual reports being diagnosed with a STD diagnosis – Results from linear probability models with two-stage least squares

	Probability of any STD diagnosis in the last year		Probability of any STD diagnosis in the last year	
	(1) LPM with FE	(2) LPM with FE and IV	(3) LPM	(4) LPM with FE and IV
Binge drinking in the last month	0.0043	0.15		
	(0.0057)	(6.90)		
Any lifetime marijuana use			0.026**	0.019
			(0.0045)	(2.53)
Depression percentile	0.036**	0.050	0.035**	0.14
	(0.0054)	(1.42)	(0.0052)	(1.44)
One of an individual's three best friends drinks alcohol at least once per month	0.0025	-0.020	0.0019	-0.0015
	(0.0053)	(1.02)	(0.0052)	(0.32)
Two of an individual's three best friends drinks alcohol at least once per month	0.0070	-0.033	0.00087	-0.0021
	(0.0041)	(1.88)	(0.0040)	(0.53)
Three of an individual's three best friends drinks alcohol at least once per month	0.010	-0.051	0.0032	0.0027
	(0.0062)	(2.84)	(0.0047)	(0.78)
Current smoker	0.0081*	-0.021	-0.00045	-0.0029
	(0.0038)	(1.37)	(0.0036)	(0.86)
Very unattractive/unattractive	-0.0052	0.00038	-0.0040	-0.0098
	(0.0054)	(0.30)	(0.0058)	(0.11)
Average attractiveness	-0.00011	0.0029	0.0016	-0.0014
	(0.0037)	(0.15)	(0.0034)	(0.076)
Body mass index	0.00032	0.00030	0.00048	0.00040
	(0.00040)	(0.0050)	(0.00040)	(0.0048)
Current student	-0.046*	-0.048	-0.050**	-0.052
	(0.019)	(0.23)	(0.018)	(0.20)
Ever married	-0.030	-0.021	-0.038	-0.036
	(0.026)	(0.44)	(0.022)	(0.38)
Children	0.068**	0.080	0.064**	0.065
	(0.023)	(0.54)	(0.022)	(0.29)
High school	-0.032	-0.038	-0.030	-0.025
	(0.021)	(0.43)	(0.021)	(0.24)

Table 14. Impact of substance use on the probability any that an individual reports being diagnosed with a STD diagnosis – Results from linear probability models with two-stage least squares

	Probability of any STD diagnosis in the last year		Probability of any STD diagnosis in the last year	
	(1) LPM with FE	(2) LPM with FE and IV	(3) LPM	(4) LPM with FE and IV
Age	-0.032	-0.043	-0.028	-0.025
	(0.052)	(0.73)	(0.050)	(0.64)
Age squared	0.0012	0.0015	0.0011	0.0010
	(0.0016)	(0.020)	(0.0015)	(0.019)
Male	0.15**	0.15	0.14**	0.14
	(0.026)	(0.48)	(0.033)	(0.36)
Male*age	-0.010**	-0.010	-0.0097**	-0.0094
	(0.0016)	(0.021)	(0.0020)	(0.023)
African American	0.019**	0.040	0.016**	0.013
	(0.0062)	(0.91)	(0.0059)	(0.13)
Hispanic	-0.0013	-0.0037	-0.0040	-0.0083
	(0.0045)	(0.12)	(0.0045)	(0.12)
Wave one indicator	-0.012**	-0.015	-0.010**	-0.011
	(0.0037)	(0.14)	(0.0034)	(0.11)
Constant	0.24	0.34	0.20	0.14
	(0.45)	(6.34)	(0.43)	(5.40)
Observations	16770	16770	17054	17054
R-squared	0.035	0.031	0.040	0.031
F	11.1	12.4	12.4	12.4

Standard errors in parentheses. Significant values denotes as follows: * $p < 0.05$, ** $p < 0.01$.
Note: The Aim One STD analysis draws on individuals from individuals in Waves One and Two who are ages 15 and older and incorporates school-level fixed effects.

Table 15. Impact of binge on the probability of ever being raped – Results from linear probability models with two-stage least squares

	(1) LPM	(2) FE	(3) FE and IV
Any binge drinking in the last month	-0.0092	-0.0096	0.40
	(0.014)	(0.015)	(4.87)
Depression	0.14**	0.14**	0.17
	(0.018)	(0.023)	(4.47)
Number of best friends who drink at least once per month			
One of three	0.035**	0.031**	-0.030
	(0.013)	(0.011)	(0.74)
Two of three	0.036*	0.033*	-0.079
	(0.014)	(0.014)	(1.31)
Three of three	0.048**	0.039*	-0.13
	(0.016)	(0.017)	(2.00)
Current smoker	0.10**	0.10**	0.023
	(0.014)	(0.019)	(1.01)
Attractiveness			
Less than average	-0.0032	-0.0078	-0.0044
	(0.021)	(0.021)	(0.25)
Average	-0.00013	-0.0032	-0.011
	(0.011)	(0.011)	(0.25)
Body mass index	-0.00023	-0.00024	-0.000076
	(0.00090)	(0.0011)	(0.011)
Current student	0.0087	0.012	0.013
	(0.022)	(0.025)	(0.30)
Ever married	-0.0074	-0.017	0.0073
	(0.022)	(0.021)	(0.39)
Children	0.052*	0.049*	0.082
	(0.020)	(0.019)	(0.41)
Highest level of education completed			
High school	0.00022	0.00100	-0.018
	(0.028)	(0.025)	(0.39)
College	0.059	0.051	0.016
	(0.033)	(0.031)	(0.77)
Graduate school	0.058	0.056	0.019
	(0.038)	(0.033)	(0.91)

Table 15. Impact of binge on the probability of ever being raped – Results from linear probability models with two-stage least squares

	(1) LPM	(2) FE	(3) FE and IV
Age	0.064**	0.065**	0.036
	(0.013)	(0.014)	(0.34)
Age squared	-0.0013**	-0.0013**	-0.00071
	(0.00030)	(0.00031)	(0.0073)
African American	-0.0029	0.00026	0.055
	(0.011)	(0.018)	(0.73)
Hispanic	-0.041*	-0.034	-0.041
	(0.016)	(0.017)	(0.28)
Wave two indicator	-0.031**	-0.030**	-0.024
	(0.0094)	(0.0076)	(0.10)
Wave four indicator	0.020	0.036	0.0031
	(0.052)	(0.059)	(0.74)
Constant	-0.73**	-0.73**	-0.48
	(0.14)	(0.17)	(4.33)
Observations	11696	11696	11696
R-squared	0.089	0.073	0.061
F	16.7	27.3	17.4

Standard errors in parentheses. Significant values denotes as follows: * $p < 0.05$, ** $p < 0.01$.
Note: The Aim Two analysis draws on women in Waves One, Two, and Four who are ages 15 and older and incorporates school-level fixed effects.

Table 16. Impact of marijuana use on the probability of ever being raped – Results from linear probability models with two-stage least squares

	(1) LPM	(2) FE	(3) FE and IV
Any lifetime marijuana use	0.075**	0.090**	-0.49
	(0.013)	(0.014)	(4.37)
Depression	0.13**	0.13**	0.20
	(0.018)	(0.021)	(4.89)
Number of best friends who drink at least once per month			
One of three	0.026*	0.022*	0.086
	(0.013)	(0.010)	(0.54)
Two of three	0.016	0.012	0.12
	(0.013)	(0.015)	(0.89)
Three of three	0.026	0.016	0.18
	(0.014)	(0.018)	(1.28)
Current smoker	0.077**	0.072**	0.26
	(0.014)	(0.016)	(1.47)
Attractiveness			
Less than average	0.00070	-0.0023	-0.017
	(0.021)	(0.020)	(0.26)
Average	0.0021	-0.00083	0.0063
	(0.010)	(0.010)	(0.25)
Body mass index	-0.000083	-0.00015	-0.000052
	(0.00090)	(0.0011)	(0.013)
Current student	0.010	0.015	-0.00025
	(0.021)	(0.024)	(0.33)
Ever married	0.0035	-0.0068	-0.045
	(0.021)	(0.021)	(0.48)
Children	0.053**	0.048*	0.068
	(0.020)	(0.019)	(0.27)
Highest level of education completed			
High school	0.00031	0.0020	-0.014
	(0.027)	(0.024)	(0.36)
College	0.053	0.044	0.041
	(0.032)	(0.030)	(0.66)
Graduate school	0.053	0.053	0.015
	(0.038)	(0.031)	(0.85)

Table 16. Impact of marijuana use on the probability of ever being raped – Results from linear probability models with two-stage least squares

	(1) LPM	(2) FE	(3) FE and IV
Age	0.063**	0.061**	0.094
	(0.013)	(0.013)	(0.28)
Age squared	-0.0013**	-0.0012**	-0.0020
	(0.00030)	(0.00031)	(0.0061)
African American	-0.0010	-0.0040	0.015
	(0.011)	(0.018)	(0.29)
Hispanic	-0.046**	-0.040*	-0.024
	(0.016)	(0.018)	(0.30)
Wave two indicator	-0.035**	-0.034**	-0.016
	(0.0092)	(0.0073)	(0.19)
Wave four indicator	0.0013	0.0096	0.15
	(0.052)	(0.060)	(1.20)
Constant	-0.74**	-0.71**	-1.03
	(0.14)	(0.16)	(4.04)
Observations	11865	11865	11865
R-squared	0.097	0.086	0.062
F	20.4	28.5	17.0

Standard errors in parentheses. Significant values denotes as follows: * $p < 0.05$, ** $p < 0.01$.
Note: The Aim Two analysis draws on women in Waves One, Two, and Four who are ages 15 and older and incorporates school-level fixed effects.

Table 17. Impact of cocaine use on the probability of ever being raped – Results from linear probability models with two-stage least squares

	(1) LPM	(2) FE	(3) FE and IV
Any lifetime cocaine use	0.16**	0.17**	1.01
	(0.029)	(0.028)	(8.01)
Depression	0.13**	0.13**	0.17
	(0.019)	(0.022)	(4.72)
Number of best friends who drink at least once per month			
One of three	0.032*	0.029**	0.013
	(0.013)	(0.011)	(0.22)
Two of three	0.030*	0.028	0.0025
	(0.014)	(0.015)	(0.33)
Three of three	0.036*	0.026	-0.039
	(0.014)	(0.018)	(0.63)
Current smoker	0.086**	0.084**	-0.017
	(0.014)	(0.016)	(0.90)
Attractiveness			
Less than average	-0.0076	-0.013	-0.0073
	(0.021)	(0.021)	(0.25)
Average	-0.0011	-0.0048	-0.0048
	(0.011)	(0.010)	(0.22)
Body mass index	0.000065	0.00011	-0.000043
	(0.00090)	(0.0011)	(0.012)
Current student	0.0063	0.0095	0.034
	(0.021)	(0.024)	(0.29)
Ever married	0.0014	-0.0072	0.028
	(0.022)	(0.021)	(0.43)
Children	0.055**	0.052**	0.053
	(0.020)	(0.018)	(0.22)
Highest level of education completed			
High school	0.0011	0.0023	0.044
	(0.027)	(0.024)	(0.45)
College	0.063*	0.055	0.11
	(0.031)	(0.030)	(0.67)
Graduate school	0.074*	0.073*	0.18
	(0.037)	(0.032)	(1.09)
Age	0.064**	0.063**	0.046
	(0.013)	(0.013)	(0.18)

Table 17. Impact of cocaine use on the probability of ever being raped – Results from linear probability models with two-stage least squares

	(1) LPM	(2) FE	(3) FE and IV
Age squared	-0.0013**	-0.0013**	-0.00089
	(0.00030)	(0.00030)	(0.0043)
African American	0.0063	0.0074	0.072
	(0.012)	(0.017)	(0.70)
Hispanic	-0.043**	-0.037*	-0.045
	(0.016)	(0.017)	(0.28)
Wave two indicator	-0.029**	-0.028**	-0.014
	(0.0095)	(0.0073)	(0.14)
Wave four indicator	-0.012	-0.0028	-0.16
	(0.052)	(0.058)	(1.62)
Constant	-0.73**	-0.72**	-0.59
	(0.14)	(0.15)	(3.37)
Observations	11631	11631	11631
R-squared	0.10	0.088	0.061
F	18.2	29.1	17.7

Standard errors in parentheses. Significant values denotes as follows: * $p < 0.05$, ** $p < 0.01$.
Note: The Aim Two analysis draws on women in Waves One, Two, and Four who are ages 15 and older and incorporates school-level fixed effects.

Table 18. Impact of lagged binge drinking on employment and wages – Results from two-part model with two-stage least stages

	(1)	(2)	(3)	(4)
Dependent variable	Employed	Wage if employ > 0	Employed	Wage if employ > 0
Model	LPM	OLS	LPM with IV	OLS with IV
Binge drinking in last month (lagged from wave three)	0.024	0.041	0.0092	0.34
	(0.013)	(0.028)	(0.081)	(0.19)
Depression percentile	-0.070**	-0.18**	-0.086	-1.02
	(0.019)	(0.043)	(0.49)	(1.14)
One of an individual's three best friends drinks alcohol	0.011	0.052	0.016	0.028
	(0.020)	(0.040)	(0.030)	(0.075)
Two of an individual's three best friends drinks alcohol	0.046*	0.079*	0.053	0.022
	(0.019)	(0.038)	(0.038)	(0.091)
Three of an individual's three best friends drinks alcohol	0.036*	0.12**	0.047	-0.0023
	(0.017)	(0.036)	(0.042)	(0.10)
Current smoker	-0.037**	-0.16**	-0.032	-0.14
	(0.012)	(0.028)	(0.035)	(0.082)
Very unattractive/unattractive	-0.076**	-0.058	-0.072	-0.0067
	(0.025)	(0.051)	(0.041)	(0.096)
Average attractiveness	-0.045**	-0.088**	-0.044*	-0.059
	(0.011)	(0.026)	(0.022)	(0.056)
Body mass index	0.00034	-0.0052**	0.00027	-0.0047*
	(0.00093)	(0.0018)	(0.0011)	(0.0024)
Current student	-0.100**	-0.094**	-0.10**	-0.11*
	(0.016)	(0.035)	(0.021)	(0.051)
Ever married	-0.00013	0.10**	0.000051	0.073
	(0.012)	(0.025)	(0.025)	(0.058)
Any children	-0.057**	-0.11**	-0.057**	-0.089*
	(0.013)	(0.027)	(0.017)	(0.040)
High school	0.11**	0.21**	0.10*	0.15
	(0.030)	(0.054)	(0.043)	(0.092)
4-year college	0.18**	0.39**	0.18**	0.28*
	(0.028)	(0.053)	(0.062)	(0.13)
Graduate school	0.17**	0.62**	0.17*	0.48**
	(0.033)	(0.060)	(0.081)	(0.18)

Table 18. Impact of lagged binge drinking on employment and wages – Results from two-part model with two-stage least stages

	(1)	(2)	(3)	(4)
Dependent variable	Employed	Wage if employ > 0	Employed	Wage if employ > 0
Model	LPM	OLS	LPM with IV	OLS with IV
Age	0.24*	0.076	0.24	0.040
	(0.11)	(0.22)	(0.13)	(0.30)
Age squared	-0.0043*	-0.00082	-0.0042	-0.00021
	(0.0019)	(0.0039)	(0.0023)	(0.0053)
Male	-0.41*	0.14	-0.39	0.12
	(0.20)	(0.41)	(0.24)	(0.56)
Male*age	0.017*	0.0016	0.017*	0.00077
	(0.0070)	(0.014)	(0.0084)	(0.020)
Hispanic	0.046**	0.060	0.044	0.11
	(0.017)	(0.035)	(0.031)	(0.072)
African American	-0.015	-0.15**	-0.017	-0.068
	(0.018)	(0.042)	(0.031)	(0.081)
Constant	-2.56	0.94	-2.55	1.79
	(1.57)	(3.16)	(1.90)	(4.28)
Observations	8149	6620	8149	6620
r2	0.073	0.11	0.070	0.10
F	16.1	28.8	15.4	27.5

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$.

Note: The Aim Three analysis draws on individuals from Wave Four. Lagged substance use measures are drawn from Wave Three.

Table 19. Impact of lagged marijuana use on employment and wages – Results from two-part model with two-stage least stages

	(1)	(2)	(3)	(4)
Dependent variable	Employed	Wage if employ > 0	Employed	Wage if employ > 0
Model	LPM	OLS	LPM with IV	OLS with IV
Marijuana use in the last year (lagged from wave three)	-0.025*	0.0093	-0.060	0.24
	(0.013)	(0.029)	(0.076)	(0.18)
Depression percentile	-0.066**	-0.18**	-0.083	-1.06
	(0.019)	(0.043)	(0.49)	(1.15)
One of an individual's three best friends drinks alcohol	0.020	0.060	0.026	0.051
	(0.019)	(0.040)	(0.029)	(0.074)
Two of an individual's three best friends drinks alcohol	0.059**	0.091*	0.068*	0.074
	(0.018)	(0.037)	(0.034)	(0.083)
Three of an individual's three best friends drinks alcohol	0.057**	0.13**	0.072*	0.070
	(0.016)	(0.036)	(0.034)	(0.084)
Current smoker	-0.029*	-0.16**	-0.025	-0.13
	(0.013)	(0.028)	(0.035)	(0.083)
Very unattractive/unattractive	-0.078**	-0.059	-0.072	-0.0059
	(0.025)	(0.051)	(0.041)	(0.097)
Average attractiveness	-0.046**	-0.089**	-0.043*	-0.059
	(0.011)	(0.026)	(0.022)	(0.057)
Body mass index	0.00034	-0.0050**	0.00024	-0.0048*
	(0.00093)	(0.0018)	(0.0011)	(0.0024)
Current student	-0.10**	-0.095**	-0.10**	-0.11*
	(0.016)	(0.035)	(0.021)	(0.051)
Ever married	-0.0032	0.10**	-0.0011	0.074
	(0.012)	(0.026)	(0.025)	(0.059)
Any children	-0.058**	-0.11**	-0.058**	-0.094*
	(0.013)	(0.027)	(0.017)	(0.040)
High school	0.11**	0.21**	0.10*	0.16
	(0.030)	(0.053)	(0.043)	(0.092)
4-year college	0.18**	0.39**	0.18**	0.29*
	(0.028)	(0.053)	(0.062)	(0.13)
Graduate school	0.18**	0.62**	0.17*	0.50**
	(0.033)	(0.060)	(0.081)	(0.18)

Table 19. Impact of lagged marijuana use on employment and wages – Results from two-part model with two-stage least stages

	(1)	(2)	(3)	(4)
Dependent variable	Employed	Wage if employ > 0	Employed	Wage if employ > 0
Model	LPM	OLS	LPM with IV	OLS with IV
Age	0.23*	0.069	0.23	0.088
	(0.11)	(0.22)	(0.13)	(0.30)
Age squared	-0.0042*	-0.00071	-0.0042	-0.00088
	(0.0019)	(0.0039)	(0.0023)	(0.0052)
Male	-0.41*	0.14	-0.39	0.15
	(0.20)	(0.41)	(0.24)	(0.56)
Male*age	0.017*	0.0017	0.017*	0.00045
	(0.0070)	(0.014)	(0.0084)	(0.020)
Hispanic	0.044**	0.058	0.041	0.11
	(0.017)	(0.035)	(0.031)	(0.073)
African American	-0.021	-0.16**	-0.021	-0.11
	(0.018)	(0.042)	(0.029)	(0.075)
Constant	-2.51	1.03	-2.45	0.97
	(1.57)	(3.17)	(1.90)	(4.26)
Observations	8149	6620	8149	6620
r2	0.073	0.11	0.070	0.10
F	16.1	29.0	15.5	27.3

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$

Note: The Aim Three analysis draws on individuals Wave Four. Lagged substance use measures are drawn from Wave Three.

Table 20. Impact of lagged methamphetamine use on employment and wages – Results from two-part model with two-stage least stages

	(1)	(2)	(3)	(4)
Dependent variable	Employed	Wage if employ > 0	Employed	Wage if employ > 0
	LPM	OLS	LPM with IV	OLS with IV
Methamphetamines use in last year (lagged from wave three)	-0.021 (0.034)	-0.098 (0.069)	0.39 (0.40)	0.24 (0.18)
Depression percentile	-0.069** (0.018)	-0.18** (0.043)	-0.092 (0.50)	-1.06 (1.13)
One of an individual's three best friends drinks alcohol	0.017 (0.019)	0.064 (0.039)	0.011 (0.028)	0.051 (0.072)
Two of an individual's three best friends drinks alcohol	0.055** (0.018)	0.093* (0.036)	0.051 (0.031)	0.074 (0.081)
Three of an individual's three best friends drinks alcohol	0.049** (0.016)	0.14** (0.033)	0.040 (0.023)	0.070 (0.082)
Current smoker	-0.034** (0.012)	-0.15** (0.028)	-0.037 (0.035)	-0.13 (0.081)
Very unattractive/unattractive	-0.077** (0.025)	-0.061 (0.051)	-0.071 (0.042)	-0.0059 (0.095)
Average attractiveness	-0.046** (0.011)	-0.089** (0.026)	-0.044* (0.022)	-0.059 (0.055)
Body mass index	0.00040 (0.00093)	-0.0051** (0.0018)	0.00025 (0.0011)	-0.0048* (0.0024)
Current student	-0.100** (0.016)	-0.093** (0.035)	-0.10** (0.022)	-0.11* (0.050)
Ever married	-0.0013 (0.012)	0.098** (0.026)	0.00097 (0.026)	0.074 (0.058)
Any children	-0.057** (0.013)	-0.11** (0.027)	-0.056** (0.017)	-0.094* (0.040)
High school	0.11** (0.030)	0.21** (0.053)	0.10* (0.044)	0.16 (0.090)
4-year college	0.18** (0.028)	0.39** (0.052)	0.18** (0.063)	0.29* (0.13)
Graduate school	0.18** (0.033)	0.62** (0.060)	0.17* (0.082)	0.50** (0.18)

Table 20. Impact of lagged methamphetamine use on employment and wages – Results from two-part model with two-stage least stages

	(1)	(2)	(3)	(4)
Dependent variable	Employed	Wage if employ > 0	Employed	Wage if employ > 0
	LPM	OLS	LPM with IV	OLS with IV
Age	0.23*	0.064	0.23	0.088
	(0.11)	(0.22)	(0.13)	(0.29)
Age squared	-0.0042*	-0.00062	-0.0041	-0.00088
	(0.0019)	(0.0039)	(0.0024)	(0.0051)
Male	-0.40*	0.14	-0.38	0.15
	(0.20)	(0.41)	(0.24)	(0.55)
Male*age	0.017*	0.0015	0.016	0.00045
	(0.0070)	(0.014)	(0.0086)	(0.019)
Hispanic	0.045**	0.056	0.046	0.11
	(0.017)	(0.035)	(0.031)	(0.071)
African American	-0.021	-0.16**	-0.0078	-0.11
	(0.018)	(0.042)	(0.031)	(0.074)
Constant	-2.51	1.11	-2.46	0.97
	(1.57)	(3.17)	(1.93)	(4.18)
Observations	8149	6620	8149	6620
r2	0.073	0.11	0.071	0.10
F	15.9	28.9	15.3	27.3

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$

Note: The Aim Three analysis draws on individuals Wave Four. Lagged substance use measures are drawn from Wave Three.

REFERENCES

1. Akerlof, G.A., *Social Distance and Social Decisions.* Econometrica, 1997. **65**(5): p. 1005-1027.

2. Alexander, C., et al., *Peers, Schools, and Adolescent Cigarette Smoking.* Journal of Adolescent Health, 2001. **29**(1): p. 22-30.

3. Allen, et al., *HIV Infection in Intravenous Drug Users Entering Drug Treatment, United States, 1988 - 1989.* American Journal of Public Health, 1992. **82**(4): p. 541 - 546.

4. Andreasson, et al., *Cannabis and Schizophrenia: A Longitudinal Study of Swedish Conscripts.* The Lancet, 1987. **26**: p. 1483 - 1486.

5. Armistead, L., K. Klein, and R. Forehand, *Parental Physical Illness and Child Functioning* Clinical Psychology Review, 1995. **15**(5): p. 405-422.

6. Arseneault, L., M.C. Annon, and J. Witton, *Causal association between cannabis and psychosis: Causal association between cannabis and psychosis: examination of the evidence examination of the evidence.* British Journal of Psychiatry, 2004(184): p. 110 - 117.

7. Association, A.P., *Diagnostic and Statistical Manual of Mental Disorders 4th ed.* 1994, Washington, DC: American Psychiatric Association.

8. Averett, S.L., et al., *Race, Ethnicity, and Gender Differences in the Relationship Between Substance Use and Adolescent Sexual Behavior.* Topics in Economic Analysis & Policy, 2004. **4**(1): p. Art. 22.

9. Back, et al., "Substance Abuse in Women: Does Gender Matter?" *Psychiatric Times,* 2007: 24(1).

10. Baltagi, B.H., *Econometrics, Third Edition*. 2002, New York: Springer-Verlag.

11. Becker, G.S., *Crime and Punishment: An Economic Approach.* Journal of Political Economy, 1968. **76**(2): p. 169-217.

12. Becker, G.S., M. Grossman, and K. Murphy, *Rational Addiction and the Effect of Price on Consumption.* The American Economic Review, 1991. **81**(2): p. 237-241.

13. Becker, G.S. and K.M. Murphy, *A Theory of Rational Addiction.* Journal of Political Economy, 1988. **96**(4): p. 675-700.

14. Bray, J.W., *Alcohol Use, Human Capital, and Wages.* Journal of Labor Economics, 2005. **23**(2): p. 279 - 312.

15. Brooks, T.L., et al., *Association of adolescent risk behaviors with mental health symptoms in high school students*, in *Society of Adolescent Medicine annual meeting*. Sept. 2002: Arlington Virginia, March 2000. p. 240-246.

16. Brown, A., et al., *Depressed mood as a risk factor for unprotected sex in young people.* Australian Psychiatry, Sept. 2006. **14**(3): p. 310-312.

17. Bruneau, J., et al., *High Rates of HIV Infection among Injection Drug Users Participating in Needle Exchange Programs in Montreal: Results of a Cohort Study.* American Journal of Epidemiology 1997. **146**(12): p. 994 - 1002.

18. Buchmueller, T. and S.H. Zuvekas, *Drug Use, Drug Abuse, and Labour Market Outcomes.* Health Economics, 1998. **7**(3): p. 229 – 245.

19. Cameron, A.C. and P.K. Trivedi, *Microeconometrics Using Stata*. 2009, College Station Stata Press. 692.

20. Carpenter, Christopher S. and Dobkin, Carlos, Alcohol Regulation and Crime (March 2010). NBER Working Paper Series, Vol. w15828, pp. -, 2010. Available at SSRN: http://ssrn.com/abstract=1574647

21. Center for Substance Abuse Treatment. *The Epidemiology of Co-Occurring Substance Use and Mental Disorders*. COCE Overview Paper 8. DHHS Publication No. (SMA) 07-4308. Rockville, MD: Substance Abuse and Mental Health Services Administration, and Center for Mental Health Services, 2007.

22. Centers for Disease Control. (2007). *Fact Sheet: Understanding Sexual Violence*. [Available at: http://www.cdc.gov/ncipc/pub-res/images/SV%20Factsheet.pdf].

23. Centers for Disease Control and Prevention. *Sexually Transmitted Disease Surveillance 2009*. Atlanta: U.S. Department of Health and Human Services; 2010.

24. Chantala, K. and J. Tabor, *Strategies to Perform a Design-Based Analysis Using the Add Health Data.* June 1999 (Revised August 2000). National Longitudinal Survey of Adolescent Health, Carolina Population Center, University of North Carolina at Chapel Hill.

25. Chatterji, P., *Illicit drug use and educational attainment.* Health Economics, 2006. **15**(5): p. 489-511.

26. Chatterji, P., M. Alegria, and D. Takeuchi, *Psychiatric disorders and labor market outcomes: Evidence from the National Comorbidity Survey-Replication.* Journal for Health Economics, 2011. **30**(5): p. 858– 868.

27. Chesson, H., P. Harrison, and W. Kassler, *Sex under the influence: The effect of alcohol on sexually transmitted disease rates in the United States.* Journal of Law and Economics, 2000. **43**(1): p. 215 - 238.

28. Chirgwin, K., et al., *HIV Infection, Genital Ulcer Disease, and Crack Cocaine Use among Patients Attending a Clinic for Sexually Transmitted Diseases.* American Journal of Public Health, 1991. **81**(12): p. 1576 - 1579.

29. Cooper, M.L., R.S. Peirce, and R.F. Huselid, *Substance Use and Sexual Risk Taking Among Black Adolescents and White Adolescents.* Health Psychology, 1994. **13**(3): p. 251-262.

30. Cottler, L.B., J.E. Helzer, and J.E. Tipp, *Lifetime patterns of substance use among general population subjects engaging in high risk sexual behaviors: implications for HIV risk.* American Journal of Drug and Alcohol Abuse 1990. **16**(3-4): p. 207-222.

31. Cuellar, A.E., S. Markowitz, and A.M. Libby, *Mental Health and Substance Abuse Treatment and Juvenile Crime.* The Journal of Mental Health Policy and Economics, 2004. **7**: p. 59-68.

32. Dee, T.S., *The complementarity of teen smoking and drinking* Journal of Health Economics, 1999. 18: p. 769-793.

33. DeSimone, J., *Illegal Drug Use and Employment* Journal of Labor Economics, 2002. **20**(4): p. 952-977.

34. DeSimone, Jeffrey S., Binge Drinking and Risky Sex Among College Students (April 2010). NBER Working Paper Series, Vol. w15953, pp. -, 2010. Available at SSRN: http://ssrn.com/abstract=1598069

35. DiNardo, J. and T. Lemieux, *Alcohol, marijuana, and American youth: the unintended consequences of government regulation.* Journal of Health Economics, 2001. **20**(6): p. 991–1010.

36. Dooley, D., R. Catalano, and G. Wilson, *Depression and Unemployment: Panel Findings from the Epidemiologic Catchment Area Study.* American Journal of Community Psychology, 1994. **22**: p. 745-65.

37. Dooley, D., J. Prause, and K.A. Ham-Rowbottom, *Underemployment and Depression: Longitudinal Relationships.* Journal of Health and Social Behavior, December 2000. **41**(4): p. 421-436.

38. Centers for Disease Control. (2007). *Fact Sheet: Understanding Sexual Violence*. [Available at: http://www.cdc.gov/ncipc/pub-res/images/SV%20Factsheet.pdf].

39. Edelen, M.O. and B.B. Reeve, *Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement.* Qual Life Res, 2007. **16**(1): p. 5-18.

40. Elter, J., et al., *Relationship of clinical depression to periodontal treatment outcome.* Journal of Periodontology, April 2002. **73** (4): p. 441 - 449.

41. Ettner, S.I., R.G. Frank, and R.C. Kessler, *The Impact of Psychiatric Disorders on Labor Market Outcomes.* Industrial and Labor Relations Review, 1997. **51**(1): p. 64-81.

42. Evenhouse, E. and S. Reilly, *Improved Estimates of the Benefits of Breastfeeding Using Sibling Comparisons to Reduce Selection Bias.* Health Services Research, 2005. **40**(6p1): p. 1781-1802.

43. Fagan, J., *Interactions among drugs, alcohol, and violence.* Health Affairs, 1993. **12**(4): p. 65 - 79.

44. Fergusson and Horwood, *Early onset cannabis use and psychosocial adjustment in young adults.* Addiction, 1997. **92**(3): p. 279 - 296.

45. Fergusson, D.M., L.J. Horwood, and E.M. Ridder, *Tests of causal linkages between cannabis use and psychotic symptoms.* Addiction, 2005. **100**: p. 354–366.

46. Fischer, J.A.V. and A. Sousa-Poza, *Does Job Satisfaction Improve the Health of Workers? New Evidence using Panel Data and Objective Measures of Health.* Health Economics, 2009. **18**(1): p. 71 – 89.

47. Fonda, S.J. and R. Hezog, *Patterns and Risk Factors of Change in Somatic and Mood Symptoms among Older Adults.* Annals of Epidemiology, 2001. **11**(6): p. 361-368.

48. Fontanet, A.L., et al., *Protection against sexually transmitted diseases by granting sex workers in Thailand the choice of using the male or female condom: results from a randomized controlled trial.* AIDS 1998. **12**(14): p. 1851–1859.

49. Frank, E. and B.D. Stewart, *Depressive Symptoms in Rape Victims: A Revisit.* Journal of Affective Disorders, 1984. **7**(1): p. 77 - 85.

50. Frank, R.G. and T.G. McGuire, *Chapter 16: Economics and mental health*, in *Handbook of Health Economics*, A.J. Culyer and J. Newhouse, Editors. 2000, North Holland: New York, NY. p. 893 - 854.

51. Frank, R.H. and B.S. Bernanke, *Principles of Economics 2nd edition*. 2004, New York, NY: McGraw-Hill.

52. French, M.T., et al., *Chronic illicit drug use, health services utilization and the cost of medical care.* Social Science & Medicine, 2000. **50**(12): p. 1703 - 1713.

53. French, M.T. and I. Popovici, *That instrument is lousy! In search of agreement when using instrumental variables estimation in substance use research.* Health Economics, February 2011. **20**(2): p. 127–146.

54. French, M.T., M.C. Roebuck, and P.K. Alexandre, *Illicit Drug Use, Employment, and Labor Force Participation.* Southern Economics Journal, 2001. **68**(2): p. 349-368.

55. Furukawa, T., et al., *Screening for depression among first visit psychiatric patients: comparison of different scoring methods for the Center for the Epidemiologic Studies Depression Scale using receiver operating characteristic analyses.* Psychiatry Clinical Neuroscience, 1997. **51**: p. 71 - 78.

56. Gill, A.M. and R.J. Michaels, *Does Drug Use Lower Wages?* Industrial and Labor Relations Review, 1992. **45**(3): p. 419-434.

57. Goldberg, E.M. and S.L. Morrison, *Schizophrenia and Social Class.* British Journal of Psychiatry, 1963. **109**: p. 785-802.

58. Goodman, E. and J. Capitman, *Depressive symptoms and cigarette smoking among teens.* Pediatrics, 2000. **106**: p. 748–755.

59. Goodman, E. and R.C. Whitaker, *A Prospective Study of the Role of Depression in the Development and Persistence of Adolescent Obesity.* Pediatrics, 2002. **109**(3): p. 497-504.

60. Grossman, M. and F.J. Chaloupka, *The demand for cocaine by young adults: a rational addiction approach.* Journal of Health Economics, 1998. **17**(4): p. 427 – 474.

61. Grossman, M., et al., *Effects of Alcohol Price Policy on Youth: A Summary of Economic Research.* Journal of Research on Adolescence, 1994. **4**(2): p. 347-364.

62. Grossman, M., R. Kaestner, and S. Markowitz, *Get High and Get Stupid: The Effect of Alcohol and Marijuana Use on Teen Sexual Behavior.* Review of Economics of the Household, 2004. **2**(4): p. 413–441.

63. Grossman, M. and S. Markowitz, *I did what last night?! Adolescent risky sexual behaviors and substance abuse.* Eastern Economic Journal, Summer 2005. **31**(3): p. 383.

64. Gunn, et al., *Syphilis in San Diego County 1983-1992: Crack Cocaine, Prostitution, and the Limitations of Partner Notification.* Sexually Transmitted Diseases, 1995. **22**(1): p. 60 - 66.

65. Halpern, C.T., et al., *Adolescent Predictors of Emerging Adult Sexual Pattern.* Journal of Adolescent Health, 2006. **39**: p. 926.e1–926.e10

66. Hamilton, V.L., et al., *Unemployment, Distress, and Coping: A Panel Study of Autoworkers.* Journal of Personality and Social Psychology 1993. **65**: p. 234-47.

67. Henquet, et al., *Prospective cohort study of cannabis use, predisposition for psychosis, and psychotic symptoms in young people.* BMJ, 2004. **330**(7481): p. 11-14.

68. Jessor, R. and S.L. Jessor, *Problem behavior and psychosocial development: A longitudinal study of youth.* 1977, New York: Academic Press.

69. Johansson, E., et al., *The Association of Alcohol Dependency with Employment Probability: Evidence from the Population Survey 'Health 2000 in Finland'.* Health Economics, 2007. **16**: p. 739-754.

70. Kaestner, R., *The Effect of Illicit Drug Use on the Labor Supply of Young Adults.* The Journal of Human Resources, 1994. **29**(1): p. 126-155.

71. Kandel, D. and K. Yamaguchi, *From Beer to Crack: Developmental Patterns of Drug Involvement.* American Journal of Public Health, 1993. **83**(6): p. 851 - 855.

72. Kenkel, D.S. and D.C. Ribar, *Alcohol consumption and young adults' socioeconomic status.* Brooking Papers on Economic Activity: Microeconomics, 1994. 1: p. 119-161.

73. Kessler, R.C., J.B. Turner, and J.S. House., *Effects of Unemployment on Health in a Community Survey: Main, Modifying, and Mediating Effects.* Journal of Social Issues, 1988. **44**(4): p. 69-85.

74. Back, et al., "Substance Abuse in Women: Does Gender Matter?" *Psychiatric Times,* 2007: 24(1).

75. Keeter S. and P. Taylor, "The Millennials." *Pew Research Center Publications*, December 11, 2009. [Available: http://pewresearch.org/pubs/1437/millennials-profile]

76. Khantzian, E.J., *The Self-Medication Hypothesis of Addictive Disorders: Focus on Heroin and Cocaine Dependence.* American Journal of Psychiatry, 1985. **142**(11): p. 1259-1264.

77. Khantzian, E.J., *The Self-Medication Hypothesis of Substance Use Disorders: A Reconsideration and Recent Applications.* Harvard Rev Psychiatry, 1997. **4**(5): p. 231 - 244.

78. Kirby, D., *Antecedents of Adolescent Initiation of Sex, Contraceptive Use, and Pregnancy.* American Journal of Health Behaviors, 2002. **26**(6): p. 473-485.

79. Kral, L.H., et al., *HIV Seroprevalence among Street-Recruited Injection Drug and Crack Cocaine Users in 16 US Municipalities.* American Journal of Public Health 1998. **88**(1): p. 108 - 113.

80. Lerner, D., et al., *Unemployment, Job Retention, and Productivity Loss Among Employees With Depression.* Psychiatric Services, December 2004. **55**(12): p. 1371–1378.

81. Lewit, E.M. and D. Coate, *The Potential for Using Excise Taxes to Reduce Smoking.* Journal of Health Economics, 1982. **1**(2): p. 121-145.

82. MacDonald, Z. and S. Pudney, *Illicit drug use, unemployment, and occupational attainment.* Journal of Health Economics, 2000. **19**(6): p. 1089 – 1115.

83. Manning, W.D., M.A. Longmore, and P.C. Giordano, *Adolescents' involvement in nonromantic sexual activity.* Social Science Research, 2005. **34**(2): p. 384-407.

84. Manning, W.G., L. Blumberg, and L.H. Moulton, *The demand for alcohol: The differential response to price.* Journal of Health Economics, 1995. **14**(2): p. 123-148.

85. Markowitz, *The Price of Alcohol, Wife Abuse, and Husband Abuse.* Southern Economics Journal, 2000. **67**(2): p. 279 - 303.

86. Markowitz, S. and M. Grossman, *The effects of beer taxes on physical child abuse.* Journal of Health Economics, 2000. **19**(2): p. 271 - 282.

87. Marmaros, D. and B. Sacerdote, *Peer and social networks in job search.* European Economic Review, 2002. **46**(4-5): p. 870 - 879.

88. McClellan, M., B. McNeil, and J. Newhouse, *Does more intensive treatment of acute myocardial infraction in the elderly reduce mortality* JAMA, 1994. **77**: p. 39-64.

89. McCurley and Snyder. (2008). Co-occurrence of Substance Use Behaviors in Youth. NCJ 219239. Washington, DC: U.S. Department of Justice

90. Miranda, J., et al., *Effects of Primary Care Depression Treatment on Minority Patients' Clinical Status and Employment.* Arch Gen Psychiatry, 2004. **61**: p. 827-834.

91. Monroe, J.A., *Morality, Politics, and Health Policy*, in *Policy Challenges in Modern Health Care*, D. Mechanic, et al., Editors. 2005, Rutgers University Press: New Brunswick, N.J. p. 290.

92. Morgan, S.L. and A.B. Sørensen, *Parental Networks, Social Closure, and Mathematics Learning: A Test of Coleman's Social Capital Explanation of School Effects.* American Sociological Review, 1999. **64**(5): p. 661-681.

93. National Institute on Drug Abuse. (Last reviewed: March 2010). NIDA InfoFacts: Inhalants. [Available at http://www.drugabuse.gov/infofacts/inhalants.html]

94. National Institute of Mental Health. (Last reviewed: 2009, March 31). Questions and answers about the National Comorbidity Survey Replication (NCSR) Study. [Available at http://www.nimh.nih.gov/health/topics/statistics/ncsr-study/questions-and-answers-about-the-national-comorbidity-survey-replication-ncsr-study.shtml#q6]

95. National Highway Traffic Safety Administration. (2008). Alcohol-Impaired Driving. DOT HS 811 155. Washington, DC: National Highway Traffic Safety Administration.

96. National Highway Traffic Safety Administration. (2009). Fatalities and Fatality Rates in Alcohol-Impaired-Driving Crashes by State, 2007-2008. DOT HS 811 250. Washington, DC: National Highway Traffic Safety Administration.

97. Norton, E.C. and E. Han, *Genetic Information, Obesity, and Labor Market Outcomes.* Health Economics, 2008. **17**(9).

98. Peck, R.C., et al., *The relationship between blood alcohol concentration (BAC), age, and crash risk.* Journal of Safety Research, 2008. **39**(3): p. 311 – 319.

99. Pertold, F., *Don't Drink and… Avoid Risky Sex of Your Peers: The Influence of Alcohol Consumption of Opposite-Gender Peers on Youth Risky Sexual Behavior* in *ISSN 1211-3298*. October, 2010, CERGE-EI: Prague.

100. Phelps, C.E., *Risks and Perceived Risk of Drunk Driving among Young Drivers.* Journal of Policy Analysis and Management, 1987. **6**(4): p. 708-714.

101. Rachal, J.V., et al., *The extent and nature of adolescent alcohol abuse: The 1974 and 1978 national sample surveys*. 1980, Springfield, VA: U.S. National Technical Information Service.

102. Radloff, L.S., *The CES-D scale: A self-report depression scale for research in the general population.* Applied Psychological Measurement, 1977. **I**(13): p. 385–401.

103. Rees, Argys, and Averett, *New evidence on the relationship between substance use and adolescent sexual behavior.* Journal of Health Economics, 2001. **20**(5): p. 835 - 845.

104. Regier, D.A., et al., *Comorbidity of Mental Disorders With Alcohol and Other Drug Abuse: Results From the Epidemiologic Catchment Area (ECA) Study.* JAMA, 1990. **264**(19): p. 2511 - 2518.

105. Register, C.A. and D.R. Williams, *Labor Market Effects of Marijuana and Cocaine Use Among Young Men* Industrial and Labor Relations Review, 1992. **45**(3): p. 435-448.

106. Rhee, S.H., et al., *Genetic and Environmental Influences on Substance Initiation, Use, and Problem Use in Adolescents.* Archives General Psychiatry, 2003. **60**(12): p. 1256-1264.

107. Rolfs, R.T., M. Goldberg, and R. Sharrar, *Risk Factors for Syphilis: Cocaine Use and Prostitution.* American Journal of Public Health, July 1990. **80**(7): p. 853 - 857.

108. Rosenberg, S.D., et al., *Prevalence of HIV, Hepatitis B, and Hepatitis C in People With Severe Mental Illness.* American Journal of Public Health, 2001. **91**(1): p. 31-37.

109. Saffer, H. and F. Chaloupka, *The Demand for Illicit Drugs.* Economic Inquiry, 1999. **36**(3): p. 401-411.

110. Santelli, J.S., et al., *Multiple Sexual Partners Among U.S. Adolescents And Young Adults.* Family Planning Perspectives, November/December 1998. **30**(6): p. 271-275.

111. Schoenbaum, M., et al., *The Effects of Primary Care Depression Treatment on Patients' Clinical Status and Employment.* Health Services Research, October 2002. **37**(5): p. 1145-1158.

112. Schroeder, S.A., *An Agenda To Combat Substance Abuse.* Health Affairs, 2005. **24**(4): p. 1005 - 1013.

113. Sen, B., *Does alcohol-use increase the risk of sexual intercourse among adolescents? Evidence from the NLSY97.* Journal of Health Economics, 2002. **21**(6): p. 1085–1093.

114. Shapiro, S., et al., *Measuring Need for Mental Health Services in a General Population.* Medical Care, 1985. **23**(9): p. 1033-1043.

115. Substance Abuse and Mental Health Service Administration (SAMHSA, 2006), Office of Applied Studies. (2006). *The NSDUH Report - Serious Psychological Distress and Substance Use among Young Adult Males*. [Available at http://oas.samhsa.gov/2k6/menMH/menMH.htm]

116. Substance Abuse and Mental Health Service Administration (SAMHSA, 2007a), (Office of Applied Studies. (May 3, 2007). *The NSDUH Report - Depression and the Initiation of Alcohol and Other Drug Use among Youths Aged 12 to 17*. [Available at http://oas.samhsa.gov/2k7/newUserDepression/newUserDepression.htm]

117. Substance Abuse and Mental Health Service Administration (SAMHSA, 2007b), Office of Applied Studies. (November 15, 2007). *The NSDUH Report - Depression and the Initiation of Cigarette, Alcohol, and Other Drug Use among Young Adults.* [Available at http://www.oas.samhsa.gov/2k7/newusers/depression.htm]

118. Substance Abuse and Mental Health Services Administration, *Results from the 2010 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-41, HHS Publication No. (SMA) 11-4658. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2011 [Available at: http://www.samhsa.gov/data/NSDUH/2k10Results/Web/HTML/2k10Results.htm#2.1].

119. Terza, J.V., A. Basu, and P.J. Rathouz, *Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling.* Journal of Health Economics, 2008. **27**(531- 543).

120. Tubman, J.G., M. Windle, and R.C. Windle, *Cumulative sexual intercourse patterns among middle adolescents: Problem behavior precursors and concurrent health risk behaviors.* Journal of Adolescent Health, March 1996. **18**(3): p. 182-191.

121. Vaillant, G.E., *The Natural History of Alcoholism: Causes, Patterns, and Paths to Recovery*. 1983, Cambridge, MA: Harvard University Press.

122. van Ours, J.C., *Cannabis, Cocaine, and Jobs.* Journal of Applied Econometrics, 2006. **21**: p. 897 – 917.

123. van Ours, J.C. and J. Williams, *Cannabis Use and Mental Health Problems.* Journal of Applied Econometrics, 2010.

124. Vuong, Q.H., *Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses.* Econometrica, March 1989. **57**(2): p. 307 - 333.

125. Wickramaratne, P.J. and M.M. Weissman, *Onset of psychopathology in offspring by developmental phase and parental depression.* Journal of American Academy of Child and Adolescent Psychiatry, 1998. **37**(9): p. 933-42.

126. Wilson, M.L. and J. Zietz, *Systematic Bias in Student Self-Reported Data.* Journal for Economic Educators, Fall 2004. **4**(4): p. 13-19.

127. Wolaver, A.M., *Effects of heavy drinking in college on study effort, grade point average, and major choice.* Contemporary Economic Policy 2002. **20**: p. 415-428.

128. Wooldridge, J.M., *Introductory Econometrics, 3rd Edition*. 2006, Mason, OH: Thomson Southwestern.

129. Yamada, T., M. Kendix, and T. Yamada, *The Impact of Alcohol Consumption and Marijuana Use on High School Graduation.* Health Economics, 1996. **5**(1): p. 77-92.

130. Zarkin, G.A., et al., The relationship between drug use and labor supply for young men. Labour Economics, 1998. 5(4): p. 385 – 409.