THAT INSTRUMENT IS LOUSY! IN SEARCH OF AGREEMENT WHEN USING INSTRUMENTAL VARIABLES ESTIMATION IN SUBSTANCE USE RESEARCH

Michael T. French[a] and **Ioana Popovici**

[a]Health Economics Research Group, Department of Sociology, Department of Epidemiology and Public Health, and Department of Economics, University of Miami, 5202 University Drive, Merrick Building, Room 121F, P.O. Box 248162, Coral Gables, FL 33124-2030; 305-284-6039 (phone); 305-284-5310 (fax); mfrench@miami.edu

[b]Health Economics Research Group, Department of Sociology, University of Miami, 5665 Ponce de Leon Blvd, Flipse Bldg, First Floor, Room 120, Coral Gables, Florida, 33146-0719; ipopovici@miami.edu

SUMMARY

The primary statistical challenge that must be addressed when using cross-sectional data to estimate the consequences of consuming addictive substances is the likely endogeneity of substance use. While economists are in agreement on the need to consider potential endogeneity bias and the value of instrumental variables estimation, the selection of credible instruments is a topic of heated debate in the field. Rather than attempt to resolve this debate, our paper highlights the diversity of judgments about what constitutes appropriate instruments for substance use based on a comprehensive review of the economics literature since 1990. We then offer recommendations related to the selection of reliable instruments in future studies.

Keywords

instrumental variables (IV); endogeneity; substance use

> There is certainly no absolute standard of beauty. That precisely is what makes its pursuit so interesting.
>
> John Kenneth Galbraith

1. INTRODUCTION

An extensive body of economic research has examined the relationships between addictive substance use (e.g., tobacco, alcohol, marijuana, cocaine) and its potential consequences (e.g., criminal activity, lower earnings, increased health services utilization, lower educational attainment). The methods and results employed in these studies, however, are far from consistent. Lack of research consensus on the relationships between substance use and its various outcomes can be attributed to a number of factors, including sample

---

heterogeneity, analysis method used, dissimilar measures for substance use, and low statistical power.

One important factor in explaining study variability could stem from the treatment of endogeneity of the key regressor (substance use), which is often the main statistical challenge encountered in this type of analysis. Endogeneity can arise when substance use is correlated with important unobserved regressors that are omitted from the model (i.e., omitted variables bias) or when the outcome variable has a causal impact on substance use (i.e., reverse causality). For example, finding a positive correlation between illicit drug use and unemployment does not necessarily mean that drug use causes a higher likelihood of being unemployed. Losing a job could lead to illicit drug consumption. Alternatively, the consumption of drugs and the probability of unemployment may be jointly determined by a common attitude towards risk that is unobserved and therefore omitted from the model. Failing to address endogeneity in either of these examples will lead to biased coefficient estimates (Angrist and Pischke, 2009).

Many studies have acknowledged and attempted to address the likely endogeneity of substance use. Nevertheless, there is a great deal of variability in findings even across studies that employ the same measures and techniques. Instrumental variable (IV) estimation is a powerful tool that, when used correctly, can generate consistent estimates in the presence of endogeneity.[1] Yet IV methods sometimes come with a steep price, particularly if the instruments are weak. The reliability of instrumental variables is an important source of concern and debate, as their acceptability is based on theoretical, intuitive, and statistical criteria. Several methodological papers and books (e.g., Angrist *et al.*, 1996;Angrist and Krueger, 2001;Angrist and Pischke, 2009;Bound *et al.*, 1995;Greene, 2008;Murray, 2006;Rashad and Kaestner, 2004;Staiger and Stock, 1997;Wooldridge, 2002) have discussed at length the critical importance of selecting predictive and valid instruments, the challenges involved, and techniques for minimizing the liabilities of weak instruments.

Briefly, a reliable instrumental variable must meet at least two essential criteria. First, it must be theoretically justified and statistically correlated with (after controlling for all other exogenous regressors) the endogenous variable of interest. Second, it must be exogenous to all other important and unobserved factors (i.e., uncorrelated with the disturbance term in the structural or outcome equation). In our previous example, a suitable instrumental variable would have a theoretical connection to and would be statistically correlated with (preferably highly so) the illicit drug use variable. An instrument would be considered valid if it affected employment status indirectly and solely through its association with illicit drug consumption. Although these two conditions are fairly straightforward and well accepted in the literature, economists disagree over how to operationally assess these criteria.

Rather than redefine standards or attempt to resolve discrepancies, the present study seeks to highlight the diversity of judgment among authors, discussants, and journal referees about what constitutes a credible instrument for measures of substance use. We do not evaluate the key findings of studies using IV methods (i.e., the effects of substance use on the outcomes being analyzed) but rather summarize and critically review the instrumental variable choices made by economists who have published studies on the consequences of substance use. In addition, this article highlights the varying degrees of rigor with which instrument predictive power and validity have been assessed in the literature and presents guidelines related to the

---

[1]Throughout this paper, we use the generic term "instrumental variable estimation" to refer to all methods using identifying variables excludable from the structural equation to address the endogeneity of the substance use regressor. Several of the studies reviewed in this paper employ models (e.g., recursive bivariate probit) that are not strictly IV methods. Nevertheless, the majority of published studies use the IV terminology even when these alternative estimation methods are employed.

selection of appropriate instruments in future analyses. To accomplish this goal, we conducted a comprehensive literature search and identified 60 studies published between January 1990 and March 2009 that have used IV methods to address the likely endogeneity of substance use. We decided not to include studies that were published before 1990 because IV methods have advanced considerably during the past two decades and very few IV studies with substance use as the endogenous regressor were published prior to this year.

We report the range of instruments for three categories of substance use: alcohol, illicit drugs, and tobacco. The discussion of instrumental variables in this paper is limited to cases with only one endogenous explanatory variable. The modeling and statistical tests are more complex in the case of two or more endogenous variables.

To the best of our knowledge, this study is the first comprehensive and critical evaluation of the IV literature for substance use. The summary tables can be used as a quick reference document and status report for researchers in the field, but we strongly caution against using these tables alone to select and defend instruments that may be employed in future substance use studies. As will be explained later in the paper, some of these instruments do not pass the required theoretical and statistical tests. Others have potential utility but should be re-assessed carefully and rigorously based on the particular research objective, sample, and setting.

2. SELECTING INSTRUMENTAL VARIABLES

OLS or other single-equation estimation produces biased coefficient estimates when one or more of the regresssors (substance use in our case) are significantly correlated with the error term. When used correctly, IV estimation can generate consistent parameter estimates if the analyst is able to obtain theoretically sound and statistically reliable instrumental variables. The instruments are used to identify and isolate a part of the variation in the endogenous explanatory variable that is not influenced by the omitted variables (Angrist and Krueger, 2001). A reliable instrumental variable satisfies the following two conditions:

1. It must be significantly correlated with the endogenous regressor once the other exogenous explanatory variables from the structural equation have been netted out. This is often referred to as the "strength" of the instrument.

2. It must be exogenous in the structural equation (i.e., uncorrelated with the error term), which is commonly called the "validity condition."

The first condition is directly testable by regressing the endogenous explanatory variable on the instrumental variable(s) and all other exogenous variables from the structural equation. Analysts should then rely on standard statistical tests to avoid choosing "weak" instruments. The statistical threshold for a strong instrument, however, is not as transparent as one might imagine. At a minimum, the estimated coefficient for the instrument in the reduced-form equation for the endogenous explanatory variable should be statistically different from zero at conventional levels, such as 1% or 5%. When multiple variables are used to instrument for one endogenous regressor, the joint explanatory power of the instruments should be assessed. In this case, joint significance at 1% or 5% levels may not be sufficient, as an F-statistic above 10 is commonly viewed as the threshold (Staiger and Stock, 1997; Stock *et al*., 2002). Instruments that fail to explain a sufficient amount of the variation in the endogenous regressor can generate IV estimates with large standard errors as well as lead to large asymptotic biases (Bollen *et al.*, 1995; Bound *et al*., 1995; Staiger and Stock, 1997).

The second requirement for a reliable instrumental variable is the absence of significant correlation with the error term in the equation of interest (i.e., structural equation with the endogenous substance use regressor). In other words, the instrumental variable should have

no direct effect on the dependent or outcome variable and should not be correlated with important unobserved and omitted factors in the structural equation. This condition cannot be tested or checked directly as it involves a relationship between the instrument and the error term. Hence, analysts often rely on theoretical considerations and intuition as a guide. A few statistical tests are available, however, to indirectly address this excludability condition.

An over-identification test can be conducted when two or more variables are used to instrument for an endogenous regressor. Several over-identification tests (Bollen *et al.*, 1995; Hansen, 1982; Sargan, 1958; Wooldridge, 2002) are readily available in Stata and other statistical packages. The approach developed by Sargan tests the null hypothesis that all instruments are uncorrelated with the error term in the structural equation by regressing the residuals (obtained by estimating the structural model with 2SLS) on all exogenous variables. When the substance use variable is binary, a common approach is to estimate a just-identified model and then include the remaining instruments in the structural equation as additional explanatory variables (Bollen *et al.*, 1995; Wooldridge, 2002). One can then test the null hypothesis that all instruments in the structural equation have zero coefficients. The rejection of this null hypothesis would raise concerns about the validity of at least one of these remaining instruments. Of course, the test of over-identifying restrictions hinges on the assumption that the instrument(s) used to identify the endogenous variable in the just identified model is not correlated with the error term in the structural equation (Murray, 2006).

The conventional tests of over-identifying restrictions are also suspect when heterogeneity is present in "treatment" effects. Angrist *et al.* (1996) and Angrist and Pischke (2009) argue that IV methods can consistently estimate average causal treatment effects for those who change treatment status (i.e., local average treatment effects or LATE). Each valid instrumental variable estimates a unique causal parameter that is specific to the subpopulation of "compliers" for that instrument (i.e., those for which the instrument influences their treatment status). Thus, the test of over-identifying restrictions, which checks the validity of different instruments by determining whether they estimate the same thing, is inappropriate in this case.

An alternative assessment of instrument excludability from the structural equation was proposed by Card (1995), who performed a "refutability test" to check whether college proximity can be regarded as an exogenous determinant of education in earnings equations. Card proposed an alternative instrument for education—the interaction between college proximity and low family income—that permitted the inclusion of the original instrument (college proximity) in the wage equation. He argued that his test supported the assumption that the instrumental variable (college proximity) is an exogenous determinant of schooling, as results showed a small and non-significant direct effect of college proximity on wage and very little change in the estimated return to education. In the same way, the validity of a particular instrumental variable for substance use could be examined when at least one additional instrumental variable is available to consistently estimate the structural equation.

Other studies have proposed employing reduced-form regressions, with the instruments and the exogenous variables as the regressors (with either the outcome of interest or substance use as the dependent variable), to check the intuition behind the identification strategy (Angrist and Krueger, 2001; Murray, 2006). If the sign and/or magnitude of the reduced-form estimates are implausible or counterintuitive, then one should seriously question the identification strategy. Similarly, the absence of statistical significance for the instruments in these reduced-form equations could mean that substance use does not affect the outcome

variable or that the IV estimation is uninformative (Angrist and Krueger, 2001; Murray, 2006).

3. LITERATURE SEARCH CRITERIA

Extensive literature reviews were conducted to locate all published studies that estimated the consequences or effects of addictive substances using IV techniques. We searched the following databases: EconLit, EbscoHost, Ideas for Economists, Web of Science, and Google Scholar. Different combinations of the following keywords were used: instrumental variables, instrument, two-stage least-squares, bivariate probit, alcohol, drinking, cigarettes, tobacco, smoking, illicit drugs, illegal drugs, marijuana, cocaine, heroin, and illicit substances. After an initial list of studies had been compiled, we then reviewed the Reference section of each article to discover any overlooked studies. The initial literature search identified hundreds of studies, the vast majority of which did not meet our criteria.

The final set includes 60 studies that were available in the literature and satisfied all of our criteria pertaining to substance use regressors and IV methods. In some cases, authors failed to reject the null hypothesis of exogeneity of substance use, thereby employing single-equation estimation techniques instead of IV methods.[2] These studies were also included in our final list if the instrumental variables and statistical tests were clearly reported. We limited our search to studies published between January 1990 and March 2009.

4. SUMMARY OF IV STUDIES

This section summarizes the IV choices of the studies reported in the tables, organized by type of substance: alcohol (Table I), illicit drugs (Table II), and tobacco (Table III). The alcohol use table and corresponding discussion are more voluminous than those for illicit drugs and tobacco because IV estimation with alcohol use has been more common in the published literature. In addition to the instrumental variables that appear in the tables and are discussed below, other less common instruments for substance use are listed at the bottom of the tables. Whenever possible, we tried to highlight each instrument's predictive power as well as its validity. Unfortunately, several studies did not provide test results to support their instrument choices. This could reflect the fact that the dangers of weak instruments have only become a point of focus in the literature in the mid 1990s (Bollen *et al.*, 1995;Bound *et al.*, 1995;Staiger and Stock, 1997).

4.1. Instruments for alcohol use

4.1.1. Family characteristics—Three of the most commonly used instrumental variables for alcohol use are measures of family drinking history: indicators of having a parent (mother or father) with an alcohol problem, having other relatives with an alcohol problem, and having resided with an alcoholic relative while growing up. These variables have been used as instruments to estimate the effect of alcohol use on educational achievement (Koch and McGeary, 2005; Renna, 2007; Wolaver, 2002), labor market performance (Kenkel and Ribar, 1994; Mullahy and Sindelar, 1996; Johansson *et al.*, 2007; Terza, 2002), delinquency

---

[2]One should exercise caution when selecting the final estimation technique because the IV estimator is less efficient than simple OLS when substance use is exogenous. Thus, it is useful and recommended first to test whether substance use is endogenous in the structural equation. Hausman (1978) proposed a straightforward endogeneity test that compares the OLS and IV estimates to determine whether the differences are statistically significant. The analyst should first estimate the reduced-form equation for the potentially endogenous variable (*SU*) by regressing it on all exogenous variables from the structural equation as well as the additional IVs. Next, *SU* and the predicted residuals should be added from the reduced-form equation ($\hat{u}$) to the structural equation. A coefficient on $\hat{u}$ significantly different from zero suggests that *SU* is endogenous and that it is best to proceed with IV estimation. It is worth mentioning that the Hausman test has very low power in finite samples when the instruments are weak. Hence, differences between the IV and OLS coefficient estimates should be examined along with the *p*-value of the test statistic.

(French and Maclean, 2006), and the impact of parental substance use on children's behavioral health (Jones *et al.*, 1999).

There are several biological and environmental reasons why authors have selected these measures as instrumental variables for alcohol use. First, a respondent with a problem-drinking parent might be more likely to consume alcohol due to possible genetic transmission of problematic alcohol consumption. This hypothesis is supported by considerable clinical research (Johnson and Pickens, 2001; McGue *et al.*, 2001). Second, living with an adult relative who has a drinking problem (particularly while the respondent is a child) might affect an individual's alcohol use behaviors as an adolescent and later in life. Third, alcohol is more likely to be available in a home where a family member is a problem drinker. Finally, relatives with a drinking problem are often more tolerant of underage alcohol use.

In most cases, analysts find a statistically significant and quantitatively large positive effect of family drinking variables on the respondent's use of alcohol and likelihood of dependence and abuse (Farrell *et al.*, 2003). Nevertheless, questions have been raised about the validity of these instrumental variables, as they might be correlated with aspects of an individual's early environment that could directly affect the dependent variables. For example, being raised in a family with an alcoholic relative could affect a child's educational attainment, labor market success, and delinquency independently of drinking per se. Only a few of the studies reviewed (French and Maclean, 2006; Renna, 2007; Wolaver, 2002) offer empirical evidence that these instruments can be excluded from the structural equation.

Some studies use indicators for whether parents smoked regularly as a proxy for parental drinking habits to control for the endogeneity of alcohol use in occupational attainment equations (MacDonald and Shields, 2001, 2004). Although consensus has not been reached in the debate about whether smoking and drinking are complements, a vast literature (Cameron and Williams, 2001; Decker and Schwartz, 2000; Dee, 1999a) confirms that the two behaviors are related. Similarly, studies that use measures of parental smoking as an instrument for drinking find it to be a strong positive predictor. This instrument's excludability from the structural equation, however, is suspect just as in the case of parental drinking history.

4.1.2. Personal beliefs/characteristics—One of the most popular instruments for alcohol use (as well as other addictive substances) is religiosity. Auld (2005) and Heien (1996) have used it to control for the potential endogeneity of alcohol use in wage equations. Renna (2007), Williams *et al.* (2003), and Wolaver (2002) have used religiosity as an instrument when studying the impact of alcohol use on schooling outcomes. All of these studies find a strong negative and statistically significant association between drinking and religiosity. One potential explanation is that religiosity raises the psychic costs associated with alcohol use. More directly, some religions frown upon the regular use of alcohol and other addictive substances or preach abstinence altogether. Conversely, an individual fighting alcohol addiction may seek help in religion.

Although this instrument is usually a strong predictor of drinking, some authors have expressed concern about its validity. Renna (2007), for example, recognizes that the same underlying mechanisms that motivate individuals with intense religious beliefs to limit their alcohol use might increase their demand for education. Religiosity might also be correlated with unobserved personal characteristics valued by employers that directly affect labor market outcomes (e.g., employment, earnings), rendering it an invalid instrument (Auld, 2005; Heien, 1996). Among those studies that have used religiosity as an instrument for alcohol use, over-identification tests have mixed results.

A few studies (Johansson *et al.*, 2007; MacDonald and Shields, 2001, 2004) have included indicators for chronic illnesses (e.g., diabetes, asthma) in the IV set to estimate the effects of alcohol consumption on labor market outcomes. These authors argue that although chronic diseases will influence drinking, the conditions will not have a direct effect on labor supply or earnings. For example, most diabetics are encouraged to reduce their alcohol consumption, especially of beer and wine, which are high in calories. Results show that diabetes and asthma are negatively correlated with alcohol consumption in these studies. Nevertheless, excludability of chronic disease instruments from the labor market performance structural equations is tenuous (Dooley *et al.*, 1996; Luft, 1975).

To identify the effect of drinking on earnings, some authors have used an indicator variable for whether the individual smoked by age 18 as an instrument (Barrett, 2002; Johansson *et al.*, 2007). Smoking and drinking are complimentary behaviors for many individuals, and studies find a strong empirical relationship between the two. A reasonable objection to the use of smoking as an instrument for alcohol use is related to its excludability, as both smoking and drinking are associated with personality and motivational traits (non-cognitive skills) that could also affect earnings. Thus, a retrospective measure of tobacco use that proxies for the same personal characteristics is problematic in a model of current earnings as well.

4.1.3. State laws, taxes, policies, and prices—Many studies use state-level instrumental variables along with the individual-level instruments presented above. Among these, two of the most popular are the state minimum legal drinking age (MLDA) and the state beer tax. Although currently set at age 21 in the U.S., the MLDA varied considerably across states in the 1970s. The gradual transition to a MLDA of 21 started at the beginning of the 1980s. This policy variable could reflect regional attitudes towards drinking among youth. It is used to proxy for the availability and acceptance of alcohol for adolescents and young adults in studies examining the effect of drinking on schooling (Cook and Moore, 1993; Dee and Evans, 2003; Koch and McGeary, 2005; Renna, 2007, 2008; Yamada *et al.*, 1996), wages (Bray, 2005), sexual activity (Sen, 2002), and behavioral health (Jones *et al.*, 1999). Some studies find a significant effect of the MLDA on drinking (Dee and Evans, 2003) while others show that MLDA is a weak determinant of alcohol use (Renna, 2007, 2008). These mixed results could reflect the fact that MLDA laws are not enforced uniformly in all states. A combination of the policy and its enforcement may therefore offer a better determinant of adolescent drinking.

An alternative instrument used to proxy for some of the personal costs associated with excessive drinking is the blood alcohol concentration (BAC) threshold for driving under the influence of alcohol. Williams *et al.* (2003) used the BAC level to examine the impact of drinking on educational performance while Williams (2005) studied the effect of high-school alcohol use on college drinking.

Because they directly reflect the monetary cost of drinking, state beer taxes are commonly used in the literature to instrument for the consumption of alcohol among adolescents and adults. Several outcomes were analyzed in these studies: schooling (Chatterji, 2006a; Cook and Moore, 1993; Koch and McGeary, 2005; Renna, 2007; Yamada *et al.*, 1996), labor market performance (Bray, 2005; Mullahy and Sindelar, 1996; Terza, 2002;), delinquency (French and Maclean, 2006), sexual activity (Sen, 2002), and health care utilization (Balsa *et al.*, 2008). Although many published studies find that alcohol use is negatively related to the beer tax, some economists have questioned the mechanism for this relationship given that the beer tax represents a small fraction of the full price of beer (Dee, 1999b; Mast *et al.*, 1999). For example, the 2007 excise tax on beer in Florida was only 4.4% of the average

price for a six-pack of Heineken (American Chamber of Commerce Researchers Association [ACCRA], 2007).

Whenever available, authors have used alcohol prices as instrumental variables to identify the effects of alcohol use on wages/earnings (Auld, 2005; Kenkel and Ribar, 1994) and educational attainment (Yamada *et al.*, 1996). Reliable price data, however, are difficult to obtain and state regulatory authorities distort market prices (Cook, 2007; Cook and Moore, 2000; Cook and Peters, 2005). Another limitation in the use of alcohol prices is the heterogeneity in brands, ethanol content, quality, and other factors. Several studies use quasi-standardized data on the price of alcoholic beverages published by ACCRA, the only nationally available data source on alcohol prices (Farrell *et al.*, 2003; Manning *et al.*, 1995). Yet ACCRA has changed the brands they price over the years, which represents a key limitation of these data. Moreover, Young and Bielinska-Kwapisz (2003) find strong evidence that ACCRA price data contains substantial measurement error that could bias the estimated effect of prices on alcohol demand.

Cross-state variation in alcohol taxes or prices could be, as in the case of MLDA, correlated with unobserved state characteristics that influence both alcohol use and the outcome variable. An alternative would be to use within-state variation in taxes as an instrument for drinking. Unfortunately, within-state variation in alcohol taxes is typically minimal (Dee and Evans, 2003), leading to low statistical power for detecting the effect of taxes on alcohol consumption (Dee, 1999b). Consequently, failing to reject the null hypothesis of no effect of taxes on alcohol consumption is not very informative.

Alternative state- or county-level instrumental variables commonly used in the literature include state ethanol/alcohol consumption/sales (French *et al.*, 2008; Mullahy and Sindelar, 1996; Sen, 2002; Terza, 2002), state cigarette tax (Bray, 2005; Mullahy and Sindelar, 1996; Sen, 2002; Terza, 2002), county/state police expenditure per capita, county arrest rate per crime (Averett *et al.*, 2004; Rees *et al.*, 2001; Sen, 2002), and percentage of the state's population living in dry areas (Chatterji, 2006a; Feng *et al.*, 2001; Jones *et al.*, 1999; Kenkel and Ribar, 1994). The assumption that the individual's environment affects his/her drinking behavior motivates the use of instruments such as the percentage of the state's population living in dry areas to examine the impact of drinking on schooling (Chatterji, 2006a) or labor market outcomes (Feng *et al.*, 2001; Kenkel and Ribar, 1994). Based on the assumption that drinking and smoking are complements (Sen, 2002), the state cigarette tax is often used along with the state beer tax to examine the effects of alcohol use on employment or wages (Bray, 2005; Mullahy and Sindelar, 1996; Terza, 2002). The county/state police expenditures per capita and the county arrest rates per crime reflect the state/local resources devoted to law enforcement, which increase the opportunity cost of underage drinking and reduce the likelihood of proprietors selling alcohol to underage consumers. These instruments have been used to examine sexual activity among adolescents (Averett *et al.*, 2004; Rees *et al.*, 2001; Sen, 2002).

When it comes to the validity of state-level instrumental variables, virtually all can be correlated with unobserved state sentiments or attributes that affect both drinking and the specific outcomes (Chatterji, 2006a, 2006b; Chatterji and DeSimone, 2005; Dee, 1999a; Dee and Evans, 2003; Rashad and Kaestner, 2004). States that set relatively high taxes on alcohol or allocate more resources for policy enforcement might also be inclined to implement policies that affect educational, labor market, and criminal activity outcomes. The issue, as well as strategies to control for policy endogeneity, is further discussed in the following section.

4.2. Instruments for illicit drug use

4.2.1. Policies and prices—The monetary price of cocaine is theoretically an important determinant of cocaine consumption. A few studies have used cocaine prices from the U.S. Drug Enforcement Administration's System to Retrieve Information from Drug Evidence (STRIDE) database as an instrumental variable in this context (Chatterji, 2006b; DeSimone, 2002; Grossman and Chaloupka, 1998). However, Horowitz (2001) finds evidence that STRIDE prices are not representative of market prices and, hence, he argues that STRIDE price data are not reliable for economic and policy analysis purposes. Contrary to expectations, tests of instrument strength show that cocaine prices are sometimes only weak predictors of cocaine use (Chatterji, 2006b). Part of the difficulty here is that conventional prices for illicit drugs are not readily available and alternative measures are hard to obtain. Moreover, some authors are concerned about the excludability of this instrument from employment equations, as drug expenditures might represent a large fraction of an individual's income (DeSimone, 2002). The validity of this instrumental variable is less likely to be problematic when studying other outcomes. Chatterji (2006b) used drug prices to address the endogeneity of illicit drug use in educational attainment equations. Grossman and Chaloupka (1998) examined the addictive nature of cocaine use for young adults by using cocaine prices to instrument for past and future use.

An indicator of whether marijuana is decriminalized in the respondent's state of residence is a popular instrument for marijuana use given that reliable information on marijuana prices is often unavailable (Chatterji, 2006b; DeSimone, 2002; Grossman and Chaloupka, 1998; Yamada *et al.*, 1996). As an alternative to prices, marijuana decriminalization is used to capture geographic differences in the indirect cost of illicit drug consumption. Specifically, it has been used to estimate the effect of illicit drug use on schooling (Yamada *et al.*, 1996; Chatterji, 2006b) and employment (DeSimone, 2002). Tests of instrument strength for marijuana decriminalization, however, are mixed. Pacula and colleagues (2003) question the interpretation of these findings and offer a likely explanation for the inconsistency in results found in the literature. The authors argue that "decriminalized" states display considerable heterogeneity in the specifics of these laws, thereby damping any meaningful and unified differences with non-decriminalized states. In addition, this instrumental variable, like several of the state policies reviewed so far, may be endogenous.

Other variables that reflect the opportunity cost of illicit drug use are statutory jail terms or fines for marijuana possession (Chatterji, 2006b; DeSimone, 1998). Both are assumed to limit marijuana consumption by raising the expected full price of marijuana use in terms of lost income and the stigma associated with being arrested. Chatterji (2006b) assumed no direct effect of these instruments on education. DeSimone (1998) used them to control for the likely endogeneity of past/future cocaine use in current use equations.

Two studies (Averett *et al.*, 2004; Rees *et al.*, 2001) use county/state police expenditures per capita and the county arrest rate to instrument for illicit drug consumption in models of adolescent sexual behavior. The intuition behind these instruments is similar to that presented above for alcohol use, and the same policy endogeneity concerns remain.

4.2.2. Household/parental characteristics—Parental supervision and family values are assumed to affect the likelihood of consuming illicit drugs. An indicator for parental presence in the household when the respondent was an adolescent has been used to study the effect of illicit drug use on employment (DeSimone, 2002; Kaestner, 1994a). Nevertheless, if parental presence is correlated with educational attainment, an important determinant of employment, then the identification strategy is flawed. As with alcohol, an alternative family background measure for illicit drug use is an indicator of parental problem drinking (DeSimone, 1998, 2002). It is assumed that preferences for alcohol are genetically

transmitted and that drinking raises the probability of other (illicit) substance use (DeSimone, 2002).

4.2.3. Personal beliefs/characteristics—As with alcohol use, one of the most popular instrumental variables for illicit drug use is religiosity (French *et al.*, 2001; French *et al.*, 2000; Kaestner, 1991, 1994a, 1994b; Register and Williams, 1992; Roebuck *et al.*, 2004; Zavala and French, 2003). The reasons for using this variable to instrument for illicit drug use in studies examining labor market performance, education, health care utilization, and health status are similar to those discussed above with regard to alcohol use. Namely, most religions advocate a lifestyle that is free of addictive substances and illegal activities, and strong religiosity might dissuade individuals from using illicit drugs. As expected, religiosity is consistently negative and strongly associated with illicit drug use. Nevertheless, the potential endogeneity of this instrument is a concern for reasons similar to those mentioned above with regard to alcohol use.

Kaestner (1991, 1994b) and Van Ours (2007) use alternative individual-level instruments like non-wage income, the number of delinquent acts as an adolescent, and the presence or number of dependent individuals in the household to identify the impact of illicit drug use on wages. The validity of these instruments, however, is suspect as they may be correlated with individual earnings.

4.3. Instruments for tobacco use

Relative to alcohol and illicit drug use, using instrumental variables for tobacco use is less common in the literature due to the lower prevalence of studies examining the negative effects of smoking. Most of the health consequences of smoking were well established in the literature before 1990. Economists have therefore focused mainly on estimating the price and income elasticity of demand in various markets (Chaloupka and Grossman, 1996; DeCicca *et al.*, 2002, 2008; Tauras, 2007) and the effectiveness of tobacco control policies (Levy *et al.*, 2004; Ross and Chaloupka, 2004). Moreover, many of the outcomes associated with tobacco use (e.g., low birth weight, respiratory illness, obesity) are different and more diverse relative to those associated with alcohol and illicit drug use (e.g., labor market success, educational attainment, health care utilization).

4.3.1. Policies, taxes, and prices—Natural candidates to instrument for tobacco use are cigarette prices or taxes. These are by far the most popular instrumental variables used in the literature to estimate the effect of smoking on later cannabis use (Beenstock and Rahav, 2002), health (Leigh and Schembri, 2004; Mullahy and Portney, 1990; Rashad, 2006), future smoking (Chaloupka, 1991; Jones and Labeaga, 2003), and birth outcomes (Evans and Ringel, 1999; Lien and Evans, 2005). Use of these variables as instruments is based on the assumption that higher cigarette prices/taxes will discourage smoking. Whenever over-identification tests are presented, these instruments prove to be excludable from the structural equation for a variety of dependent variables. In a study of the determinants of adult obesity, Rashad (2006) supplemented the use of cigarette taxes with another policy instrument, clean indoor air laws, which is assumed to discourage cigarette smoking overall (Chaloupka and Wechsler, 1997; Ross et al., 2005). In addition to cigarette prices, taxes, and clean indoor air laws, less common instruments found in the literature are briefly mentioned below.

4.3.2. Household characteristics—The presence of children or of a non-smoking partner in the household might encourage the respondent to reduce his/her smoking in order to avoid the negative health effects associated with second-hand smoke (Van Ours, 2004).

Although these instruments are assumed to have no direct effect on wages, the presence of children in the household might affect labor supply, a point noted by Van Ours.

4.3.3. Peer influences—Peer smoking is used to proxy for the individual's smoking environment when analyzing the effect of tobacco use on smoking risk perception (Lundborg, 2007), as studies have demonstrated a strong influence of peer smoking on the decision to use tobacco (Lundborg, 2006). Lundborg presents several intuitive arguments for why peer smoking is exogenous in Swedish schools: specifically, schools do not sort students across classes according to ability, parents cannot choose which class within the school their children will attend, and students cannot select schools or classes.

4.3.4. Personal beliefs/characteristics—Lagged cigarette consumption has been used when tobacco taxes or prices are not available. Clark and Etile (2002) examined the relationship between health while smoking and future cigarette use in Britain. The argument here is that current consumption is predicted in part by past consumption via the addictive properties of nicotine. To avoid first-order serial correlation, the authors used consumption that occurred two or more periods in the past, an approach suggested by Anderson and Hsiao (1982). Nevertheless, the instruments proved to be weak, and the IV strategy was ultimately abandoned.

Alternative instruments found in the tobacco literature include the respondent's satisfaction with life (Lundborg, 2007), a socioeconomic status measure (Van Ours, 2004), religiosity (Auld, 2005), the respondent's assessment of risk associated with smoking (Zarkin *et al.*, 1998), and an indicator of drinking or smoking before the age of 16 (Van Ours, 2004).

5. SUMMARY OF KEY FINDINGS

The key findings of this comprehensive literature review are itemized and discussed below.

1. A great deal of variety and irregularity exists in the use of instrumental variables for substance use, even among similar dependent variables and endogenous regressors. We have counted over 50 unique instrumental variables for alcohol use, over 30 for illicit drug use, and over 10 for tobacco use.

2. Some individual- and state-level instrumental variables are clearly more common and popular among researchers and, to some extent, more favored by reviewers. These include family history of alcohol use, religiosity, MLDA, and beer taxes for alcohol use; religiosity and marijuana decriminalization policies for illicit drug use; and cigarettes taxes and prices for smoking. On the other hand, many of these instruments are currently encountering disfavor in contemporary studies for reasons that will be discussed below.

3. Economists seem to have different beliefs about how to assess the appropriateness, strength, and validity of most instruments. For example, where earlier studies emphasized intuitive or theoretical support for instruments (Heien, 1996; Register and Williams, 1992), more recent studies prefer robust statistical evidence (Chatterji, 2006a; DeSimone, 2002; French and Maclean, 2006; Renna, 2007; Wolaver, 2002). This subtle change in practice over time can be partly explained by the fact that although concern about weak instruments was discussed in the econometrics literature by the late 1970s, it did not become a widespread concern until the work of Bound *et al.* (1995) and Staiger and Stock (1997).

4. To illustrate the previous point, among the 60 studies reviewed in this paper, 26 explicitly present results of the first stage regressions, 22 report the joint significance of their instruments (4 of these discuss it without providing the

results), and 25 report results of over-identification tests (5 of these discuss the tests without providing the statistics).

5. Ten of the studies analyze data from countries other than the US. Due to the lack of geographical variation in prices or policies, most of these employ individual-level instruments such as parental alcohol problems, presence of children in the household, a non-acute illness, or early onset of alcohol use. Lagged cigarettes prices are usually used as instruments in studies that use international data to study the potential consequences of tobacco use. Whether journal referees employ a different standard when evaluating the acceptability of instruments for studies using non-US data is impossible to prove but a topic for lively debate.

6. One concern raised in the literature centers on the fact that some state-level instruments (e.g., beer taxes, MLDA, marijuana decriminalization policies) may be correlated with unobserved determinants of cross-state variation in the outcome variables (e.g., criminal activity, educational attainment) (Chatterji, 2006a, 2006b; Chatterji and DeSimone, 2005; Dee, 1999a; Dee and Evans, 2003; Rashad and Kaestner, 2004). State-specific cultural attributes or social attitudes that lead a state to decriminalize the possession of marijuana or to set higher excise taxes on alcohol could also influence other state policies that directly affect the outcome of interest. To address this concern, Dee and Evans (2003) relied on within-state variation in the MLDA as an exogenous determinant of teen drinking. Alternatives would be including state dummies to capture time-invariant state-specific sentiments that affect policy enactment or controlling for state-level explanatory variables that are directly correlated with the selected outcome (e.g., state-level crime rates when examining the effect of substance use on criminal activity outcomes). Perhaps a better alternative when panel data are available would be to employ fixed-effects or first-difference models that remove the unobserved and time-invariant factors (Boden *et al.*, 2008; Tekin, 2004; Wooldridge, 2002). This approach is not a perfect solution, however, as any important time-varying factor is a potential remaining source of bias (Besley and Case, 2000). Changes in state-specific social attitudes could lead to changes in state policies or might be accompanied by changes in other factors that are correlated with the outcome. In such cases, panel data methods can be combined with IV estimation to derive consistent parameter estimates in the presence of time-varying endogeneity. Moreover, controlling for unobserved and time-invariant fixed effects might weaken the predictive power of state-specific policy instruments due to reduced within-state variation of these policies over time. Like most of the decisions involving IV methods, the analyst is faced with problematic tradeoffs.

7. A second source of concern is the fact that state-level instruments usually have less predictive power than individual-level instruments. As mentioned above, several procedural papers demonstrate that the use of weak instruments, even when they are not correlated with the error term in the structural equation, can lead to IV estimates that are more biased than single-equation estimates (Bollen *et al.*, 1995; Bound *et al.*, 1995; Staiger and Stock, 1997). Moreover, including more instruments will often make the problem worse (Angrist and Pischke, 2009). Since weak instruments will jeopardize the consistency and asymptotic efficiency of the parameter estimates, numerous studies erroneously supplement weak state-level instruments with more predictive individual-level variables that are not necessarily excludable.

8. Economists have used alternative empirical approaches to deal with concerns over the reliability of state-level policy instruments. Koch and Ribar (2001) used family fixed-effects models with the siblings' alcohol use initiation age as an identifying

instrument. Dee and Evans (2003) employed a two-sample IV strategy (Angrist and Krueger, 1992, 1995) that combines first-stage and reduced-form results using matched cohorts from two datasets.

9. Individual-level instruments such as family background of substance use or religiosity are usually highly predictive of substance use. Nevertheless, authors have recognized the likelihood that these variables are not legitimately excludable from the structural model (Auld, 2005; Mullahy and Sindelar, 1996; Renna, 2007; Zavala and French, 2003). As mentioned above, family background variables can be correlated with critical aspects of adolescents' living environment that also affect their educational, employment, or criminal activity outcomes as young adults and beyond. Strong religiosity might deter an individual from drinking, but could also directly impact criminal activity choices or time spent studying.

10. In their elusive search for better instruments, economists have sometimes reanalyzed data from previously published papers. For example, Terza (2002) re-examined the Mullahy and Sindelar (1996) study using a multinomial choice regression specification (i.e., multinomial logit with an endogenous treatment effect) that allows for non-linearity while accounting for the endogeneity of substance use (McFadden, 1973). Rashad and Kaestner (2004) critique the IV choices employed by Rees *et al.* (2001) and Sen (2002), and suggest alternative approaches. These and other examples suggest that some of the estimation methods and instrument choices used in earlier published studies would not necessarily meet the standards of contemporary peer review.

11. It is worth reiterating that in the presence of treatment effects heterogeneity, each valid instrumental variable estimates a unique parameter specific to the subpopulation of compliers for that instrument (Angrist and Pischke, 2009). We might therefore expect variability in the key results, even if different instruments for substance use are equally valid, as different causal parameters are estimated by each instrument. For the same reasons, the proper interpretation of policy implications demands familiarity with the subpopulation affected by an instrument (Kling, 2001; Meyer, 1995).

6. CONCLUSION

Ideally, instrumental variables should be grounded in theory, intuitively plausible, highly predictive of the endogenous substance use regressor, and uncorrelated with the error term in the structural equation. If candidate instruments were always conceptually sound and readily available, then empirical work would be less challenging and instrument discussions would be less spirited. Indeed, a colleague once facetiously argued that applied economists should choose a topic and find a data set only after first identifying and obtaining clever and defensible instrumental variables. Of course, such a strategy encourages instrument "fishing," but we suspect that a fair amount of fishing expeditions were conducted among the studies reviewed for this paper.

Absent the availability of ideal instruments for substance use, we offer a few recommendations that may improve the execution and findings of future IV studies. Methodological studies such as those by Murray (2006), Angrist and Krueger (2001), Angrist and Pischke (2009), and Rashad and Kaestner (2004) present guidelines for selecting reliable instruments and recognizing bad ones. Much of our advice is borrowed from these and other excellent sources.

Among substance use studies, it is common practice to select and utilize several instruments for one endogenous variable. While this is generally an effective approach to increasing

predictive power if all instruments are strong and valid, a few cautions must be kept in mind. First, we believe that standard statistical tests should always be used to assess the individual and joint explanatory power of the instruments. Moreover, results of these tests should be clearly reported and discussed in the manuscript. As previously mentioned, an F-statistic below 10 suggests that the instrumental variables might perform poorly in the IV models. In these cases, Murray (2006) recommends strategies to cope with weak instruments. A possible solution is constructing confidence intervals from a two-sided conditional likelihood ratio test (Andrews and Stock, 2005; Andrews *et al.*, 2006) that usually performs well regardless of whether instruments are weak. An alternative to proceeding with an IV analysis when instruments are weak is to use Fuller's (1977) estimators, originally proposed to modify limited information maximum likelihood estimation to obtain finite moments. Finally, because the F-statistic varies inversely with the number of weak instruments, the bias from IV estimation is an increasing function of the number of weak instruments. In this case, Angrist and Pischke (2009) suggest picking a single "best" instrument and estimating a just-identified model (i.e. using only one of the instruments to identify the endogenous variable). This suggestion follows from the fact that the just-identified model has the least bias because the number of instruments is lowest. Despite the availability of these second-best alternatives, the use of weak instruments should generally be avoided because the potential bias in the estimated effects could be equal to or greater than that from simple single-equation estimation (Bollen *et al.*, 1995; Bound *et al.*, 1995; Staiger and Stock, 1997).

Besides strongly predicting the potentially endogenous regressor, other key statistics with which to establish the validity of the instruments are the tests of over-identifying restrictions (Bollen *et al.*, 1995; Wooldridge, 2002) and Card's refutability test (Card, 1995). In conducting this literature review, we were surprised to discover that less than half of the studies actually conducted over-identification or related tests. Clearly reporting the results of these tests is strongly recommended as part of a complete assessment of instrument strength and validity. Supporting tests are not a perfect solution, however, as the reliability of the over-identification test is suspect when all instruments share a common characteristic (Murray, 2006; Wooldridge, 2002). An over-identification test is reliable only when a high level of confidence exists in the strength and validity of at least one of the instrumental variables. Whenever possible, we would recommend a small set of strong exogenous instruments that affect substance use through different mechanisms. Naturally, obtaining a blended set of 2–4 quality instrumental variables is easier said than done.

As noted above, some economists place greater weight on intuitive/theoretical/institutional evidence of good instruments while others prefer strong statistical support. In this area, we would recommend a reasonable balance of intuition, theory, institutional support, and statistics. After all, if an alcohol tax is a theoretically sound instrument for alcohol use but does a poor job predicting variation in drinking, then it is a weak and inappropriate instrument in this context. On the other hand, yearly rainfall totals in distinct geographical areas might pass all of the statistical tests for a strong and valid instrument, but it would be hard to defend this choice on conceptual grounds. Again, intuitive and theoretical arguments should be supplemented by empirical evidence from reduced-form regressions with the instruments and the exogenous variables as the regressors. To establish credibility and confidence, the estimated coefficient of an instrument in the reduced-form equation should be statistically significant and agreeable in sign and magnitude with intuition, theory, and institutional knowledge (Murray, 2006).

After an exhaustive search for the best available instruments, authors might be faced with an uncomfortable dilemma. How should one proceed when the best available instruments are theoretically sound but statistically weak (or vice-versa)? In these cases, we suggest running single-equation models to determine whether a substance use variable is significantly related

to a particular outcome measure while being careful to avoid any implications of causality. The analyst should also run a reduced-form regression of the outcome variable on the instruments and other exogenous regressors. If none of the instruments are significant, then a causal relationship between the substance use measure and the outcome of interest probably does not exist (Angrist and Krueger, 2001). When the substance use variable is dichotomous, another option is the use of propensity score matching to estimate "treatment effects" with observational data (Balsa *et al.*, 2008; Balsa and French, in press; Rosenbaum and Rubin, 1983). Of course, the reality of this predicament is that, depending on the topic, single-equation associations or propensity score matching in the face of a potentially endogenous regressor might not meet certain reviewers' threshold for a meaningful contribution to the literature.

In closing, IV methods have advanced considerably in the past two decades, particularly in the field of substance use research. Along with advanced technical capabilities comes an expectation among colleagues and reviewers that instrument choices will be convincingly defended and analyses will be carefully executed. Alas, this review article clearly illustrates that authors, reviewers, and readers are hardly in agreement as to which standards to apply and how vigorously to enforce them. As a profession, we are not yet at a stage where it is possible to confidently endorse a small set of reliable instrumental variables for economic studies of substance use consequences. Indeed, we firmly caution against using these findings alone to justify the choice of instrumental variables in future research. Given this predicament, perhaps it is appropriate to return to the quotation with which we began this paper – but with a slight adjustment: *There is certainly no absolute standard of [instrument strength, validity, and theoretical support]. That precisely is what makes its pursuit so interesting.*

Acknowledgments

REFERENCES

Alexandre PK, French MT. Further evidence on the labor market effects of addiction: chronic drug use and employment in metropolitan Miami. Contemporary Economic Policy. 2004; 22:382–393.

American Chamber of Commerce Researchers Association (ACCRA). ACCRA Cost of Living Index. The Council for Community and Economic Research; 2007. Available at: http://www.coli.org/

Anderson TW, Hsiao C. Formulation and estimation of dynamic models using panel data. Journal of Econometrics. 1982; 18:578–606.

Andrews DWK, Moreira MJ, Stock JH. Optimal two-sided invariant similar tests for instrumental variables regression. Econometrica. 2006; 74:715–752.

Andrews, DWK.; Stock, JH. Cowles Foundation Discussion Paper No. 1530. Yale University: 2005. Inference with weak instruments.

Angrist JD, Imbens G, Rubin D. Identification of causal effects using instrumental variables. Journal of the American Statistical Association. 1996; 91:444–455.

Angrist JD, Krueger AB. The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. Journal of the American Statistical Association. 1992; 87:328–336.

Angrist JD, Krueger AB. Split-sample instrumental variables estimates of the return to schooling. Journal of Business and Economic Statistics. 1995; 13:225–235.

Angrist JD, Krueger AB. Instrumental variables and the search for identification: from supply and demand to natural experiments. Journal of Economic Perspectives. 2001; 15:69–85.

Angrist, JD.; Pischke, J. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton, NJ: Princeton University Press; 2009.

Auld MC. Smoking, drinking, and income. The Journal of Human Resources. 2005; 15:504–518.

Averett SL, Rees DI, Duncan B, Argys L. Race, ethnicity, and gender differences in the relationship between substance use and adolescent sexual behavior. Topics in Economic Analysis & Policy. 2004; 4(1) Article 22.

Balsa AI, French MT. Alcohol use and the labor market in Uruguay. Health Economics. In press.

Balsa AI, Homer JF, Fleming M, French MT. Alcohol consumption and health among elders. The Gerontologist. 2008; 48:622–636. [PubMed: 18981279]

Barrett GF. The effect of alcohol consumption on earnings. The Economic Record. 2002; 78:79–96.

Beenstock M, Rahav G. Testing gateway theory: do cigarette prices affect illicit drug use? Journal of Health Economics. 2002; 21:679–698. [PubMed: 12146597]

Besley T, Case A. Unnatural experiments? Estimating the incidence of endogenous policies. The Economic Journal. 2000; 110:F672–F694.

Boden JM, Fergusson DM, Horwood LJ. Cigarette smoking and suicidal behavior: results from a 25-year longitudinal study. Psychological Medicine. 2008; 38:433–439. [PubMed: 17892622]

Bollen KA, Guilkey DK, Mroz TA. Binary outcomes and endogenous explanatory variables: tests and solutions with an application to the demand for contraception use in Tunisia. Demography. 1995; 32:111–131. [PubMed: 7774727]

Bound J, Jaeger DA, Baker RM. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. Journal of the American Statistical Association. 1995; 90:443–450.

Bray JW. Alcohol use, human capital, and wages. Journal of Labor Economics. 2005; 23:279–312.

Cameron L, Williams J. Cannabis, alcohol and cigarettes: substitutes or complements? Economic Record. 2001; 77:19–34.

Card, D. Using geographic variation in college proximity to estimate the return to schooling. In: Christofides, LN.; Grant, EK.; Swidinsky, R., editors. Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp. Toronto: University of Toronto Press; 1995.

Chaloupka FJ. Rational addictive behavior and cigarettes smoking. Journal of Political Economy. 1991; 99:722–742.

Chaloupka, F.; Grossman, M. NBER Working Paper No. W5740. Cambridge, MA: National Bureau of Economic Research; 1996. Price, tobacco control policies and youth smoking.

Chaloupka FJ, Wechsler H. Price, tobacco control policies and smoking among young adults. Journal of Health Economics. 1997; 16:359–373. [PubMed: 10169306]

Chatterji P. Does alcohol use during high-school affect educational attainment? Evidence from the National Education Longitudinal Study. Economics of Education Review. 2006a; 25:482–497.

Chatterji P. Illicit drug use and educational attainment. Health Economics. 2006b; 15:489–511. [PubMed: 16389630]

Chatterji, P.; DeSimone, J. NBER Working Paper No. 11337. Cambridge, MA: National Bureau of Economic Research; 2005. Adolescent drinking and high-school dropout.

Clark A, Etile F. Do health changes affect smoking? Evidence from British panel data. Journal of Health Economics. 2002; 21:533–562. [PubMed: 12146590]

Cook, PJ. Paying the Tab: The Economics of Alcohol Policy. Princeton, NJ: Princeton University Press; 2007.

Cook PJ, Moore MJ. Drinking and schooling. Journal of Health Economics. 1993; 12(4):411–429. [PubMed: 10131754]

Cook, PJ.; Moore, MJ. Alcohol. In: Culyer, AJ.; Newhouse, JP., editors. Handbook of Health Economics. Vol. vol. 1B. North-Holland: NY: 2000. p. 1629-1673.part 2

Cook, PJ.; Peters, B. NBER Working paper No. 11902. Cambridge, MA: National Bureau of Economic Research; 2005. The myth of the drinker's bonus.

DeCicca P, Kenkel DS, Mathios A. Putting out the fires: will higher taxes reduce the onset of youth smoking? Journal of Political Economy. 2002; 110:144–169.

DeCicca P, Kenkel DS, Mathios A. Cigarette taxes and the transition from youth to adult smoking: smoking initiation, cessation, and participation. Journal of Health Economics. 2008; 27:904–917. [PubMed: 18513811]

Decker, SL.; Schwartz, AE. NBER Working paper No. W7535. Cambridge, MA: National Bureau of Economic Research; 2000. Cigarettes and alcohol: substitutes or complements?.

Dee TS. The complementarity of teen smoking and drinking. Journal of Health Economics. 1999a; 18:769–793. [PubMed: 10847934]

Dee TS. State alcohol policies, teen drinking and traffic fatalities. Journal of Public Economics. 1999b; 72:289–315.

Dee TS, Evans WN. Teen drinking and educational attainment: evidence from two-sample instrumental variables estimates. Journal of Labor Economics. 2003; 21:178–209.

DeSimone J. Is marijuana a gateway drug? Eastern Economic Journal. 1998; 24:149–164.

DeSimone J. Illegal drug use and employment. Journal of Labor Economics. 2002; 20:952–977.

Dooley D, Fielding J, Levi L. Health and unemployment. Annual Review of Public Health. 1996; 17:449–465.

Evans WN, Ringel JS. Can higher cigarette taxes improve birth outcomes? Journal of Public Economics. 1999; 72:135–154.

Farrell S, Manning WG, Finch MD. Alcohol dependence and the price of alcoholic beverages. Journal of Health Economics. 2003; 22:117–147. [PubMed: 12564720]

Feng W, Zhou W, Butler JS, Booth BM, French MT. The impact of problem drinking on employment. Health Economics. 2001; 10:509–521. [PubMed: 11550292]

French MT, Maclean JC. Underage alcohol use, delinquency, and criminal activity. Health Economics. 2006; 15:1261–1281. [PubMed: 16786500]

French MT, Maclean JC, Ettner S. Drinkers and bettors: investigating the complementarity of alcohol consumption and problem gambling. Drug and Alcohol Dependence. 2008; 96(1–2):155–164. [PubMed: 18430523]

French MT, McGeary KA, Chitwood DD, McCoy CB. Chronic illicit drug use, health services utilization, and the cost of medical care. Social Science and Medicine. 2000; 50:1703–1713. [PubMed: 10798326]

French MT, Roebuck MC, Alexandre PK. Illicit drug use, employment, and labor force participation. Southern Economic Journal. 2001; 68:349–368.

Fuller WA. Some properties of a modification of the limited information maximum likelihood estimator. Econometrica. 1977; 45:939–954.

Gill AM, Michaels RJ. Does drug use lower wages? Industrial and Labor Relations Review. 1992; 45:419–434.

Greene, WH. Econometric Analysis. 6th Edition. Upper Saddle River, NJ: Prentice Hall; 2008.

Grossman M, Chaloupka FJ. The demand for cocaine by young adults: a rational addiction approach. Journal of Health Economics. 1998; 17:427–474. [PubMed: 10180926]

Hansen LP. Large sample properties of generalized method of moments estimators. Econometrica. 1982; 50:1029–1054.

Hausman JA. Specification tests in econometrics. Econometrica. 1978; 46:1251–1270.

Heien DM. Do drinkers earn less? Southern Economic Journal. 1996; 60:63–68.

Horowitz JL. Should the DEA's STRIDE data be used for economic analyses of markets for illegal drugs? Journal of the American Statistical Association. 2001; 96:1254–1271.

Johansson E, Alho H, Kiiskinen U, Poikolainen K. The association of alcohol dependency with employment probability: evidence from the population survey "Health 2000 in Finland.". Health Economics. 2007; 16:739–754. [PubMed: 17186570]

Johnson EO, Pickens RW. Familial transmission of alcoholism among non alcoholics and mild, severe, and dissocial subtypes of alcoholism. Alcoholism: Clinical and Experimental Research. 2001; 25:661–666.

Jones AM, Labeaga JM. Individual heterogeneity and censoring in panel data estimates of tobacco expenditure. Journal of Applied Econometrics. 2003; 18:157–177.

Jones AS, Miller DJ, Salkever DS. Parental use of alcohol and children's behavioral health: a household production analysis. Health Economics. 1999; 8:661–683. [PubMed: 10590469]

Kaestner R. The effect of illicit drug use on the wages of young adults. Journal of Labor Economics. 1991; 9:381–412.

Kaestner R. The effect of illicit drug use on the labor supply of young adults. The Journal of Human Resources. 1994a; 29:126–155.

Kaestner R. New estimates of the effect of marijuana and cocaine use on wages. Industrial and Labor Relations Review. 1994b; 47:454–470.

Kenkel DS, Ribar DC. Alcohol consumption and young adults' socioeconomic status. Brooking Papers on Economic Activity: Microeconomics. 1994; 1:119–161.

Kling JR. Interpreting instrumental variables estimates of the returns to schooling. Journal of Business and Economic Statistics. 2001; 19:358–364.

Koch SF, McGeary KA. The effect of youth alcohol initiation on high-school completion. Economic Inquiry. 2005; 43:750–765.

Koch SF, Ribar DC. A siblings analysis of the effects of alcohol consumption onset on educational attainment. Contemporary Economic Policy. 2001; 19:162–174.

Leigh JP, Schembri M. Instrumental variables technique: cigarette price provided better estimate of effects of smoking on SF-12. Journal of Clinical Epidemiology. 2004; 57:284–293. [PubMed: 15066689]

Levy DT, Chaloupka F, Gitchell J. The effects of tobacco control policies on smoking rates: a tobacco control scorecard. Journal of Public Health Management & Practice. 2004; 10:338–353. [PubMed: 15235381]

Lien DS, Evans WN. Estimating the impact of large cigarette tax hikes. The Journal of Human Resources. 2005; 15:373–392.

Luft HS. Impact of poor health on earnings. The Review of Economics and Statistics. 1975; 57:43–57.

Lundborg P. Having the wrong friends? Peer effects in adolescent substance use. Journal of Health Economics. 2006; 25:214–233. [PubMed: 15964090]

Lundborg P. Smoking, information sources, and risk perceptions – new results on Swedish data. Journal of Risk and Uncertainty. 2007; 34:217–240.

MacDonald Z, Shields MA. The impact of alcohol consumption on occupational attainment in England. Economica. 2001; 68:427–453.

MacDonald Z, Shields MA. Does problem drinking affect employment? Evidence from England. Health Economics. 2004; 13:139–155. [PubMed: 14737752]

Manning WG, Blumberg L, Moulton LH. The demand for alcohol: the differential response to price. Journal of Health Economics. 1995; 14:123–148. [PubMed: 10154654]

Mast BD, Benson BL, Rasmussen DW. Beer taxation and alcohol-related fatalities. Southern Economic Journal. 1999; 66:214–249.

McFadden, D. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P., editor. Frontiers of Econometrics. New York: Academic Press; 1973. p. 105-142.

McGeary KA, French MT. Illicit drug use and emergency room utilization. Health Services Research. 2000; 35:153–169. [PubMed: 10778828]

McGue M, Iacono WG, Legrand LN, Elkins I. Origins and consequences of age at first drink. II. Familial risk and heritability. Alcoholism: Clinical and Experimental Research. 2001; 25:1166–1173.

Meyer BD. Natural and quasi-experiments in economics. Journal of Business and Economic Statistics. 1995; 13:151–161.

Mullahy J, Portney PR. Air pollution, cigarette smoking, and the production of respiratory health. Journal of Health Economics. 1990; 9:193–205. [PubMed: 10107501]

Mullahy J, Sindelar J. Employment, unemployment, and problem drinking. Journal of Health Economics. 1996; 15:409–434. [PubMed: 10164037]

Murray MP. Avoiding invalid instruments and coping with weak instruments. Journal of Economic Perspectives. 2006; 20(4):111–132.

Pacula, RL.; Chriqui, JF.; King, J. NBER Working paper No. 9690. Cambridge, MA: National Bureau of Economic Research; 2003. Marijuana decriminalization: what does it mean in the United States?.

Rashad I. Structural estimation of caloric intake, exercise, smoking, and obesity. Quarterly Review of Economics and Finance. 2006; 46:268–283.

Rashad I, Kaestner R. Teenage sex, drugs and alcohol use: problems identifying the cause of risky behaviors. Journal of Health Economics. 2004; 23:493–503. [PubMed: 15120467]

Rees DI, Argys LM, Averett SL. New evidence on the relationship between substance use and adolescent sexual behavior. Journal of Health Economics. 2001; 20:835–845. [PubMed: 11558651]

Register CA, Williams DR. Labor market effects of marijuana and cocaine use among young men. Industrial and Labor Relations Review. 1992; 45:435–448.

Renna F. The economic cost of teen drinking: late graduation and lowered earnings. Health Economics. 2007; 16:407–419. [PubMed: 17031781]

Renna F. Teens' alcohol consumption and schooling. Economics of Education. 2008; 27:69–78.

Roebuck MC, French MT, Dennis ML. Adolescent marijuana use and school attendance. Economics of Education Review. 2004; 23:133–141.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Ross H, Chaloupka FJ. The effect of public policies and prices on youth smoking. Southern Economic Journal. 2004; 70:796–815.

Ross H, Powell LM, Tauras JA, Chaloupka FJ. New evidence on youth smoking behavior based on experimental price increases. Contemporary Economic Policy. 2005; 23:195–210.

Sargan JD. The estimation of economic relationships with instrumental variables. Econometrica. 1958; 26:393–415.

Sen B. Does alcohol-use increase the risk of sexual intercourse among adolescents? Evidence from the NLSY97. Journal of Health Economics. 2002; 21:1085–1093. [PubMed: 12475127]

Staiger D, Stock JH. Instrumental variables regression with weak instruments. Econometrica. 1997; 65:557–586.

Stock JH, Wright JH, Yogo M. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics. 2002; 20:518–529.

Tauras JA. Differential impact of state tobacco control policies among race and ethnic groups. Addiction. 2007; 102 Supp 2:95–103. [PubMed: 17850619]

Tekin E. Employment, wages, and alcohol consumption in Russia. Southern Economic Journal. 2004; 71:397–417.

Terza JV. Alcohol abuse and employment: a second look. Journal of Applied Econometrics. 2002; 17:393–404.

Van Ours JC. A pint a day raises a man's pay; but smoking blows that gain away. Journal of Health Economics. 2004; 23:863–886. [PubMed: 15353183]

Van Ours JC. The effects of cannabis use on wages of prime-age males. Oxford Bulletin of Economics and Statistics. 2007; 69:619–634.

Williams J. Habit formation and college students' demand for alcohol. Health Economics. 2005; 14:119–134. [PubMed: 15386667]

Williams J, Powell LM, Wechsler H. Does alcohol consumption reduce human capital accumulation? Evidence from the College Alcohol Study. Applied Economics. 2003; 35:1227–1239.

Wolaver AM. Effects of heavy drinking in college on study effort, grade point average, and major choice. Contemporary Economic Policy. 2002; 20:415–428.

Wooldridge, JM. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press; 2002.

Yamada T, Kendrix M, Yamada T. The impact of alcohol consumption and marijuana use on high school graduation. Health Economics. 1996; 5:77–92. [PubMed: 8653193]

Young DJ, Bielinska-Kwapisz A. Alcohol consumption, beverage prices, and measurement error. Journal of Studies on Alcohol. 2003; 64(2):235–238. [PubMed: 12713197]

Zarkin GA, Mroz TA, Bray JW, French MT. The relationship between drug use and labor supply for young men. Labour Economics. 1998; 5:385–409.

Zavala SK, French MT. Dangerous to your health. The role of chronic drug use in serious injuries and trauma. Medical Care. 2003; 41:309–322. [PubMed: 12555058]

Table I

Summary of Most Common Instrumental Variables for Alcohol Consumption

Instrumental variables	Analysis sample		
	Adolescents	**Young adults**	**Adults**
Family characteristics			
Number or presence of children			1, 2, 10
Parent with alcohol problem(s)		3, 4, 5, 6	7, 8, 25[*]
Other relative with alcohol problem(s)	9	3, 5	1, 25[*]
Resided with alcoholic relative (while under age 18)			1, 8
Parent smoking status			10, 11
Personal beliefs/characteristics			
Religiosity		3, 6, 13	12, 14
Smoked at age 18			15, 7
Chronic disease/health			7, 10, 11
State laws, taxes, policies, and prices			
BAC limits	16	13	
State Minimum Legal Drinking Age (MLDA)	17, 18, 3, 19, 20, 9, 21, 23	24	1
State beer taxes	17, 3, 19, 9, 23, 29	24, 4, 13	25, 8, 30
State ethanol/alcohol consumption/sales	19		25, 8, 31
State cigarette taxes	19	24	25, 8
County/state police expenditures per capita	26, 19, 27		
County arrest rates per crime[**]	26, 19[***], 27		
Percent of state's population residing in dry counties	29	5	28, 1
Alcohol prices	23	5	1, 12

Key to studies cited above: 1. Jones *et al.* (1999); 2. Van Ours (2004); 3. Renna (2007); 4. French and Maclean (2006); 5. Kenkel and Ribar (1994); 6. Wolaver (2002); 7. Johansson *et al.* (2007); 8. Terza (2002); 9. Koch and McGeary (2005); 10. MacDonald and Shields (2001); 11. MacDonald Shields (2004); 12. Auld (2005); 13. Williams *et al.* (2003); 14. Heien (1996); 15. Barrett (2002); 16. Williams (2005); 17. Cook and Moore (1993); 18. Dee and Evans (2003); 19. Sen (2002); 20. Dee (1999a); 21. Renna (2008); 23. Yamada *et al.* (1996); 24. Bray (2005); 25. Mullahy and Sindelar (1996); 26. Averett *et al.* (2004); 27. Rees *et al.* (2001); 28. Feng *et al.*, (2001); 29. Chatterji (2006a); 30. Balsa *et al*., (2008); 31. French *et al*., (2008).

Less common instrumental variables for alcohol consumption

Adolescents: penalty for alcohol possession for 1st offense in school (Averett *et al.*, 2004), state requires drug and alcohol education at middle school levels or earlier (Sen, 2002), parent absent while growing up (Williams, 2005), state of residence requires schools to offer alcohol and drug prevention education (Rees *et al.*, 2001), marijuana decriminalization (Yamada *et al.*, 1996).

Young adults: sibling's age of drinking onset (Koch and Ribar, 2001), state restrictions on happy hours, pitcher sales, state open container laws (Williams *et al*., 2003), percent of drinkers at college who obtained alcohol without an ID, resident assistant present, family disapproves of alcohol use, parent ever consumed alcohol (Wolaver, 2002), respondent's assessment of risk associated with using alcohol (Zarkin *et al.*, 1998).

Adults: proportion of sample members in the individual's local area who are abstainers and who are heavy drinkers (Barrett, 2002), mother or father has/had mental health problem(s), parents fighting, heavy alcohol use as indicated by high gamma-glutamyl transferase level (Johansson *et al*., 2007), state mandates/prohibits exclusive malt beverage distribution territories, state controls spirit and wine sales/spirit sales only, state allows

grocery store sales of alcohol, area per capita employees in retail liquor stores, area per capita retail liquor outlets, percent of population Mormon, percent of population Baptist, alcohol treatment availability (per capita alcohol treatment facilities, per capita CMHC/hospital/residential/outpatient alcohol treatment units, state per capita real expenditures for alcohol treatment, state per capita block grants for alcohol), medical care prices,

geographic area dummies, parents' height and weight without shoes, respondent's drinking prior to child conception, family income (Jones *et al.* ‡, 1999), indicator for stomach ulcer, indicator for living in an urban area, self-assessment of how much respondents drink (MacDonald and Shields, 2001), partner smoking status(MacDonald and Shields, 2004), alcohol or tobacco use before age 16, socioeconomic status indicator, presence of a

partner (Van Ours¶, 2004), tobacco price (Auld, 2005), respondent height, restricted off premises sale of alcohol on Sundays, exemptions for underage alcohol possession (French *et al.*, 2008), alcohol sales prohibited in gas stations, state bans on Sunday sales, merchandising prohibited in alcohol transactions, financial penalties for DUI, penalties for cocaine consumption (Balsa *et al*., 2008).

*
 Notes: The instrument is an indicator for whether respondent lived with alcoholic parent/other relative while under age 18.

**
  Number of arrests divided by total number of crimes in the county of residence.

***
   The instrument is the county-level juvenile DUI arrest rate per 100,000 population aged 10–17.

‡The list includes all the instruments used for the following endogenous variables: parental alcohol use, maternal time use (working or housekeeping), and access to routine care.

¶The analysis sample includes individuals ages 16 to over 65. list includes all the instruments used for the following endogenous variables: alcohol use and tobacco use.

Table II

Summary of Most Common Instrumental Variables for Illicit Drug Use

Instrumental variables	Analysis sample		
	Adolescents	Young adults	Adults
Policies and prices			
Cocaine price	1	2, 3	
Marijuana decriminalization	1, 4	2, 3	
Statutory jail terms for marijuana possession	1	5	
Fines for marijuana possession	1	5	
County/state police expenditures per capita	6, 7		
County arrest rates per crime[*]	6, 7		
State Minimum Legal Drinking Age (MLDA)	17	3	
Household/parental characteristics			
Parents with alcohol problem(s)		2, 5	
Household composition at age 14		2, 8	
Personal beliefs/characteristics			
Non-wage income		9, 10	
Number of delinquent acts		9, 10	
Presence or number of dependents		9, 10	11
Religiosity	12	9, 13, 8, 10	14, 15, 16

Key to Studies cited above: 1. Chatterji (2006b); 2. DeSimone (2002); 3. Grossman and Chaloupka (1998); 4. Yamada *et al.* (1996); 5. DeSimone (1998); 6. Rees *et al.* (2001); 7. Averett *et al.* (2004); 8. Kaestner (1994a); 9. Kaestner (1991); 10. Kaestner (1994b); 11. Van Ours (2007); 12. Roebuck *et al.* (2004); 13. Register and Williams (1992); 14. French *et al.* (2001); 15. French *et al.* (2000); 16. Zavala and French (2003); 17. Yamada *et al.* (1996).

Less common instrumental variables for illicit drug use:

Adolescents: school principal's perception of whether drugs are a moderate to serious problem at the school, indicator of whether school has a policy of expelling students if they are caught with illegal drugs on school property (Chatterji, 2006b), penalty for drug possession for 1st offense in school (Averett *et al.*, 2004), beer tax, liquor price (Yamada *et al.*, 1996), state of residence requires schools to offer alcohol and drug prevention education (Rees *et al.*, 2001).

Young adults: state excise tax on beer (DeSimone, 1998), respondent's assessment of risk associated with using various substances, respondent's assessment of difficulty in obtaining various substances (Zarkin *et al.*, 1998), respondent divorced in previous year, respondent resides in central city, parents' education (Register and Williams, 1992), perceived self-esteem (Kaestner, 1994a), drinking behavior, prior illegal activity (Gill and Michaels, 1992).

Adults: number of workers at individual and family services agencies per 100,000 residents in a particular zip code (Alexandre and French, 2004), parental cannabis use (Van Ours, 2007), frequency of seeing intoxicated individuals or individuals selling drugs in the neighborhood (McGeary and French, 2000).

[*] Notes: Number of arrests divided by total number of crimes in the county of residence.

Table III

Summary of Most Common Instrumental Variables for Tobacco Use

	Analysis sample		
Instrumental variables	**Adolescents**	**Adults**	**Households**
Policies, taxes, and prices			
Tobacco price		1, 2, 3[*], 13, 8[**]	12
Cigarettes tax	4	3[*], 5, 6, 7	
Clean indoor air act		5	
Household characteristics			
Presence of children		9	
Presence of a partner		9	
Peer influences			
Peer cigarettes use	10		
Personal beliefs/characteristics			
Lagged consumption		11	
Respondent's satisfaction with life	1 0		
Socioeconomic status		9	
Religiosity		13	
Respondent's assessment of risk associated with smoking cigarettes		14	
Drinking or smoking before age 16		9	

Key to Studies cited above: 1. Beenstock and Rahav (2002); 2. Leigh and Schembri (2004); 3. Chaloupka (1991); 4. Dee (1999a); 5. Rashad (2006); 6. Evans and Ringel † (1999); 7. Lien and Evans (2005); 8. Mullahy and Portney ‡ (1990); 9. Van Ours (2004); 10. Lundborg (2007); 11. Clark and Etile (2002); 12. Jones and Labeaga (2003); 13. Auld (2005); 14. Zarkin *et al.* (1998).

Notes: Table includes all instrumental variables used for tobacco use found in the literature.

[*] Lagged and lead smoking is instrumented by further lags and leads of cigarettes prices and taxes.

[**] A one year lagged value of tobacco prices was included.

† Analysis sample includes women ages 15 to 44.

‡ The study also includes the following variables in the instrument set: race, marital status, full-time student status, occupational status, education.