DRUGS AND JUVENILE CRIME: EVIDENCE FROM A PANEL OF SIBLINGS AND TWINS

H. Naci Mocan and Erdal Tekin

ABSTRACT

Using data from the U.S. National Longitudinal Study of Adolescent Health, this chapter investigates the impact of individual drug use on robbery, burglary, theft, and damaging property for juveniles. Using a variety of fixed-effects models that exploit variations over time and between siblings and twins, the results indicate that drug use has a significant impact on the propensity to commit crime. We find that the median impact of cocaine use on the propensity to commit various types of crimes is 11 percentage points. The impact of using inhalants or other drugs is an increase in the propensity to commit crime by 7 percentage points, respectively.

1. INTRODUCTION

The analysis of the determinants of juvenile risky behaviour in general, and juvenile crime in particular has become an important research question (Gruber, 2001; Levitt, 1998b; Mocan and Rees, 2005). Drug use is a potentially important determinant of criminal activity, although the extent

of the relationship between drugs and crime has not been identified clearly. Despite the strong evidence that drug use and criminal activity are positively correlated, the causal impact of drug use on crime has not been conclusively established (see the literature reviews of Chaiken & Chaiken, 1990; Harrison, 1992). Even though some recent studies using aggregate data provided evidence on the potential causal impact of drug use on crime (Corman & Mocan, 2000; Grogger & Willis, 2000), convincing cause-and-effect evidence from micro-data is missing.[1] The difficulty in identifying the causal impact of drug use on crime stems from the possibility that the observed positive correlation between drug use and crime may be due to the influence of an unobserved variable which has an impact on both drug use and criminal activity. For example, if the degree of risk aversion of the individual has an impact on both his drug use and criminal behaviour, then biased estimates of the impact of drug use on crime would be obtained in analyses that do not take into account the confounding due to risk aversion.

This chapter investigates the link between illicit drugs and juvenile crime using nationally representative individual-level data. To eliminate confounding due to unobservable variables, we exploit the longitudinal aspect of the data which include siblings and twins who live in the same household. The use of longitudinal data to eliminate time-invariant individual heterogeneity is a standard tool in micro-econometrics. As explained below in detail, the longitudinal nature of our data, and an unusually large number of personal and family background variables allow us to examine the impact of illicit drug use on an individual's criminal activity.

Data on twins have been employed by previous research to estimate returns to education, schooling and marriage decisions, and the impact of birth weight on infant health (Ashenfelter & Krueger, 1994; Miller, Mulvey, & Martin, 1995; Behrman, Rosenzweig, & Taubman, 1994, 1996; Almond, Chay, & Lee, 2002). All of these twin studies employed cross-sectional data on twins, and to the best of our knowledge, this is the first study that uses a *panel* of siblings and twins to control for both the impacts of time-invariant and time-varying unobservables.

We analyse four different crimes – robbery, burglary, damaging property, and theft. The four drug use indicators we employ are the use of cocaine, the use of inhalants, the injection of illegal drugs, and the use of other drugs. We address potential measurement error in drug use.

Drug use is found to increase the propensity to commit crime. Using cocaine, inhalants, and other drugs increases the propensity to commit crime from 6 to 11 percentage points; injecting drugs increases the probability of committing crime by 41 percentage points.

In this chapter, Section 2 presents the analytical framework, Section 3 discusses the measurement error in drug use, Section 4 describes the data, Section 5 displays the results, and Section 6 is the conclusion.

2. ANALYTICAL FRAMEWORK

The crime supply equation with the addition of drug use can be presented as follows:[2]

$$CR = f(X, A, F, DR) \qquad (1)$$

where CR stands for a measure of the extent of the criminal activity of the individual; X represents the characteristics of the person, such as age, race and ethnicity, and religious beliefs; A stands for location-specific deterrence and economic variables that impact criminal involvement, such as crime-specific arrest rates, police presence, and the unemployment rate; F is a vector of parent and family characteristics; and DR represents drug use of the individual.[3]

Drug consumption in Eq. (1) is a function of the price of drugs, buyers' income, and tastes for drug use, and specific penalties targeted at drug users. Using Goldstein's (1985) conceptual framework, drug use can affect criminal activity through three channels. First, the "pharmacological" effect is the direct impact of drug use on criminal activity because drug use may increase aggression. The second is the "economic" effect – that higher expenditures on drugs cause some users to finance these expenditures by committing crime. The third is the "systemic" effect – the violence due to the illegality of the drug market, because the participants cannot rely on contracts and courts to resolve disputes. If the "economic" effect is the dominant factor to influence criminal activity, the impact of drug use on crime could be ambiguous. For example, if the demand for drugs is price inelastic, then an increase in drug use, say due to a rightward shift of the supply of drugs, would be associated with an increase in drug consumption which is coupled with a decrease in drug spending. If the economic effect is more important than the pharmacological one, increased drug use would be associated with a reduction in crime.[4]

Empirical specification of the crime supply equation as a function of observable and unobservable personal characteristics (including biological attributes), deterrence measures, economic conditions, as well as the attributes of the family, and the extent of the drug use of the individual is

presented by the following equation:

$$CR_{jit} = \alpha + \delta X_{jit} + \gamma F_{jit} + \beta DR_{jit} + \tau A_{st} + \mu_{ji} + \lambda_{jit} + \Omega_{ji}^{F} + \Psi_{jit}^{F} + \varepsilon_{jit} \quad (2)$$

where CR_{jit} is the criminal activity measure of the ith individual of the jth sibling (or twin) pair at time t; X_{jit} represents observable individual characteristics such as age, race, gender and religiosity of the person, weekly allowance of the child, and measures of risk aversion such as whether the child wears seatbelt while driving; F_{jit} stands for observable family attributes, including parent characteristics and measures of the extent of supervision at home (the complete list of these variables is given in Section 4); DR_{jit} represents consumption of drugs; A_{st} stands for the deterrence measures faced by the individual, such as the arrest rates and the size of the police force, as well as local economic conditions in location s at time t where the child resides; μ_{ji} captures individual-specific time-invariant unobservables which include intellect; λ_{jit} represents person-specific time-varying unobservables; Ω_{ji}^{F} captures unobservable time-invariant family attributes; Ψ_{jit}^{F} is unobservable time-varying family attributes; and ε_{jit} is a standard error term.

Taking the first difference of Eq. (2) across time periods gives:

$$\Delta CR_{jit} = \delta \Delta X_{jit} + \gamma \Delta F_{jit} + \beta \Delta DR_{jit} + \tau \Delta A_{st} + \Delta \lambda_{jit} + \Delta \Psi_{jit}^{F} + \Delta \varepsilon_{jit} \quad (3)$$

where Δ stands for time differencing. Eq. (3) is a standard fixed-effects model, where time-invariant family and individual characteristics drop out, but time-varying heterogeneity remains.

Note that in Eq. (3), the change in individual's criminal activity between the 2 years depends, among other factors, on the change in local deterrence and economics variables (A_{st}). The values of these variables are not collected beyond the first year of our data; therefore, ΔA_{st} cannot be calculated. However, following Currie and Moretti (2003), and Cook and Ludwig (2002), we include state or county dummies to control for such factors. That is, we replace ΔA_{st} with state- or county-fixed effects for the first-differenced models.

There is variation in the consumption of illicit drugs between sibling pairs. This allows us to eliminate time-varying family effects by taking within-sibling differences of Eq. (3), which gives:[5]

$$\nabla \Delta CR_{jit} = \delta \nabla \Delta X_{jit} + \beta \nabla \Delta DR_{jit} + \nabla \Delta \lambda_{jit} + \nabla \Delta \varepsilon_{jit} \quad (4)$$

where ∇ stands for between-sibling differencing. This specification eliminates all heterogeneity with the exception of time-varying individual-specific unobservables (λ_{jit}). Note that the family environment and

location-specific economic, and deterrence variables drop out in Eq. (4) as they are the same for all siblings of the same household.

The analogue of Eq. (4) for twins is:

$$\nabla\Delta CR_{jit} = \delta\nabla\Delta X_{jit} + \beta\nabla\Delta DR_{jit} + \nabla\Delta\varepsilon_{jit} \tag{5}$$

In Eq. (5), time-varying individual-specific heterogeneity is eliminated under the assumption that it is biologically the same between twins. This may particularly be the case for monozygotic (identical, or MZ) twins. Therefore, Eq. (5) is estimated for all twins (monozygotic and fraternal), as well as for MZ twins.

3. MEASUREMENT ERROR

Data collection procedures were designed to minimize concerns about confidentiality, as described in detail in Section 4. For example, respondents were not provided with written questionnaires; rather they listened to sensitive questions on delinquent behaviour and drug use through earphones, and entered their answers directly on laptop computers. Nevertheless, it is still conceivable that drug use is reported with error. Furthermore, it is plausible that the reporting error is not symmetric in the classical sense, but it is one sided.

To demonstrate the impact of non-random measurement error in drug use in first-differenced data, consider the following equation:

$$\Delta CR_{it} = \beta\Delta DR_{it}^* + \Delta\varepsilon_{it} \tag{6}$$

where i stands for the ith individual and t is the time period. The subscript j and other covariates are dropped for ease of exposition. Let ΔDR_{it}^* be the actual drug use, DR_{it} stand for the reported drug use, and v_{it} represent the measurement error. The reported drug use is equal to the actual drug use plus the measurement error; that is, $DR_{it} = \Delta DR_{it}^* + v_{it}$. Note that $DR = 1$ if the individual reports using drugs, and $DR = 0$ if he/she reports no drug use. Similarly, $DR^* = 1$ if the actual drug use is positive and $DR^* = 0$ if actual drug use is zero. Let the probability distribution of v_{it} be:

$$\text{Prob}(DR_{it} = 1, DR_{it}^* = 1) = p_1, \quad \text{Prob}(DR_{it} = 1, DR_{it}^* = 0) = 0,$$

$$\text{Prob}(DR_{it} = 0, DR_{it}^* = 0) = p_2, \quad \text{Prob}(DR_{it} = 0, DR_{it}^* = 1) = q$$

That is:

DRit	DR_{it}^*	vit	Prob(vit)
1	1	0	p1
1	0	1	0
0	0	0	p2
0	1	−1	q

The first row in the table indicates that the probability of using drugs and reporting as such is p_1. The second row indicates that the probability of reporting positive drug use when the person in fact did not use drugs is zero. The probability of telling the truth when actual drug use is zero is p_2, and q stands for the probability of lying when the actual drug use is positive.

The estimated β in Eq. (6) is equal to:

$$\hat{\beta}_\Delta = \frac{\Sigma_i \Delta DR_{it} \Delta CR_{it}}{\Sigma_i \Delta DR_{it}^2} = \frac{\Sigma(\Delta DR_{it}^* + \Delta v_{it})(\Delta DR_{it}^* \beta + \Delta \varepsilon_{it})}{\Sigma \Delta DR_{it}^2}$$

Simplifying and taking the probability limit gives:

$$p \lim \hat{\beta} = \beta \, \frac{\text{var}(\Delta DR^*) + \text{cov}(\Delta v, \Delta DR^*)}{\text{var}(\Delta DR)}$$

It can be shown that (see Appendix A):

$$p \lim \hat{\beta} = \beta \left[\frac{p_1 p_2}{(p_1 - p_1^2)(1 - \rho)} \right] \tag{7}$$

where ρ is the autocorrelation coefficient of reported drug use between the time periods (i.e., $\rho = \text{cov}(DR_{it}, DR_{it-1})/\text{var}(DR_{it})$ assuming a covariance-stationary process for DR).

Following Ashenfelter and Zimmerman (1997), the probability limit of β in Eq. (7) can be substituted into Eq. (6) to obtain:

$$\Delta CR_{it} = \beta \left[\frac{p_1 p_2}{(p_1 - p_1^2)(1 - \rho)} \right] \Delta DR_{it} + \Delta \varepsilon_{it} \tag{8}$$

Note that p_1 is readily available in the data, which is the mean reported drug use. The medical literature contains detailed information regarding the reliability of self-reported substance use. For example, in an analysis the drinking patterns of college students, it has been found that the reliability of reporting in the quantity and frequency of drinking beer, wine, and spirits was high, with a reliability ratio of 0.84.[6] Reliability ratios range from 0.89

to 0.92 for items such as "driven a car while drinking," "missed a class because of hangover," and "damaged property because of drinking" (Weiss et al., 1998). An analysis of out-of-treatment drug users indicated a reliability ratio of 0.72 for self-reported cocaine use, 0.77 for heroin, and 0.82 for crack. The ratio was 0.88 for the number of times the person injected drugs. For both cocaine and opiate use, total agreement between self-reports and urinalysis was over 84 percent (Johnson et al., 2000). Utilizing this literature, we postulate that 80 percent of drug users reported their drug use correctly. This suggests that $p_2 = 1 - (p_1/0.8)$, $q = p_1/4$, and ρ is calculated from the data, separately for each drug use measure.[7] Variations in the reporting rate did not change the results in a meaningful way.

It is well known that classical measurement error in the explanatory variable attenuates its estimated coefficient, and the bias is exacerbated in first-differenced data (Levitt, 1998a; Griliches and Hausman, 1985). In our case, where we entertain the possibility of one-sided measurement error due to differential propensity of telling the truth about the use of illicit drug use, the bias depends on p_1, p_2, and ρ.

In models that employ time and sibling (or twin) differencing, we estimate models (suppressing other covariates) such as:

$$\nabla\Delta CR_{jit} = \beta\nabla\Delta DR^*_{jit} + \nabla\Delta\varepsilon_{jit}$$

The probability limit of the estimated coefficient is equal to (the details are in Appendix A):

$$p\lim\hat{\beta} = \beta\left[\frac{4p_1p_2}{(p_1 - p_1^2)(4 + \Phi)}\right]$$

where $\Phi = 2(-\rho_{DR_{j22}DR_{j21}} - \rho_{DR_{j22}DR_{j12}} + \rho_{DR_{j22}DR_{j11}} + \rho_{DR_{j21}DR_{j12}} - \rho_{DR_{j21}DR_{j11}} - \rho_{DR_{j12}DR_{j11}})$. Φ is calculated from reported drug use as it depends on observed correlations in reported drug use over time and between siblings or twins. Thus, in models with time and sibling (or twin) differencing we have,[8]

$$\nabla\Delta CR_{jit} = \beta\left[\frac{4p_1p_2}{(p_1 - p_1^2)(4 + \Phi)}\right]\nabla\Delta DR_{jit} + \nabla\Delta\varepsilon_{jit} \qquad (9)$$

4. DATA

The data used in the analyses are drawn from the two waves of the National Longitudinal Study of Adolescent Health (Add Health).[9] Add Health is a

nationally representative study of adolescents in grades 7–12. An in-school questionnaire was administered to every student who attended one of the sampled 132 U.S. schools on a particular day during the period between September 1994 and April 1995. A random sample of approximately 200 adolescents from each high school/feeder school pair was selected for in-home interviews which were conducted from April to December 1995.[10] The in-home interviews constituted the core sample and contained about 12,000 adolescents. In addition to the core sample, several special samples (e.g., ethnic and genetic) were also drawn on the basis of in-school interviews. The core and the special samples provide a total number of 20,745 adolescents for Wave I. The adolescents are interviewed for the second time from April to August 1996 for Wave II. In Wave II, 14,738 adolescents were interviewed.[11] Data are gathered from adolescents, from their parents, siblings, friends, romantic partners and fellow students, and from school administrators. The survey was designed to provide detailed information on teen behaviour, including their criminal and substance use/abuse.

One feature of Add Health that we utilize in this chapter is the genetic oversample. The genetic sample consists of pairs of siblings (full, half, and stepsiblings), identical twins, and fraternal twins. Eligibility for the genetic sample was determined based on the responses provided by adolescents in the in-school questionnaire. All mixed sex twin pairs were classified as fraternal, or dizygotic (DZ). In addition to asking each twin if they were MZ or DZ, each twin was also given a set of questions on confusability of appearance (if they looked like two peas in a pod as young children, and three questions on whether they are confused by strangers, teachers, or family members). A zygosity scale is created, which is an average of the confusability item scores over the reports of both twins. When self-reported data on appearance was missing, mother's report of confusability of appearance was used. If there was conflict between the twins' self-reports of zygosity and the classification based on confusability of appearance, the twins are classified as "uncertain zygosity." Using the responses from Wave I questionnaire, those classified as uncertain zygosity were asked in Wave II for cheek samples for DNA analysis. There are 43 twin pairs that refused to provide a sample for testing, and they are deleted from our sample. After deleting twins with undetermined zygosity, the raw sample of siblings (including twins) consists of 4,030 individuals. Of these, 1,986 are full siblings, 700 are half siblings, 821 are DZ twins, and 523 are MZ twins. The sample of twins contains the DZ and MZ twins; and the sample of identical twins consists of the 523 MZ twins. Twins constitute 7 percent of the sample.[12] There is one set of triplets and no quadruplets. The triplets are coded as three sets of twins.

The survey includes a number of detailed questions about delinquent behaviour of adolescents. Specifically, respondents were asked whether they had committed any of the following acts in the 12 months prior to the interview date – robbery, burglary, damaging property, and theft. Adolescents were also asked about whether they had used different types of illicit drugs such as cocaine, other drugs (heroin, LSD, etc.), inhalants, or ever injected any illegal drugs with a needle. In wave I, the juveniles were asked if they ever used these drugs. In wave II, they were asked if they used these drugs since the last interview. Survey administrators took several steps to maintain data security and to minimize the potential for interviewer or parental influence. First, respondents were not provided with any printed questionnaires. Rather, all data were recorded on laptop computers. Second, for sensitive topics, such as delinquent behaviour and substance use/abuse, the adolescents listened to pre-recorded questions through earphones and entered their answers directly on the laptops.[13]

Definitions of the variables used in empirical analyses based on the siblings and twins samples and their descriptive statistics are reported in Table 1.[14] The first two columns of Table 1 report the weighted means and standard deviations of the sibling sample of Wave I. The next column displays the standard deviations of the first-differenced variables, and the last column presents the standard deviations for the first-and-sibling-differenced variables. Some personal and household characteristics, such as race, ethnicity, gender, and whether parents were born in the U.S.A. do not change between the waves. Therefore, these variables are not reported in Table 1. The deterrence variables in Wave I, such as arrest rates, pertain to 1992, and they were not collected in Wave II. This is not a drawback because sibling or twin differencing eliminates all variables that are the same across twins or siblings. Put differently, siblings of the same household are exposed to the same time-series variation in contextual variables, such as local economic and social conditions, and deterrence measures.

5. RESULTS

In Table 2, we report the estimated coefficients of drug use indicators using the sample of siblings (including twins). The top panel presents results pertaining to Eq. (3), which is the time-differenced model. As noted above these models include state dummies as controls for the change in the local deterrence variables across periods since these variables do not exhibit variation in the data. Models with county-fixed effects did not change the

Table 1. Descriptive Statistics.

	Definition	Wave I Cross Section Siblings and Twins		First Difference Siblings and Twins	First and Sibling Difference
		Mean	Standard deviation	Standard deviation	Standard deviation
Damage	Dummy variable (= 1) if deliberately damaged someone else's property that did not belong to you in the past 12 months, 0 otherwise	0.208	0.406	0.422	0.582
Burglary	Dummy variable (= 1) if went into a house or building to steal something in the past 12 months, 0 otherwise	0.054	0.226	0.254	0.359
Theft	Dummy variable (= 1) if took something from a store without paying for it, or took something worth more than 50 dollars in the last 12 months	0.251	0.434	0.462	0.656
Robbery	Dummy variable (= 1) if used or threatened to use a weapon to get something from someone in the past 12 months, 0 otherwise	0.047	0.211	0.240	0.322
Cocaine	Dummy variable (= 1) if ever used any kind of cocaine (including powder, freebase, or crack cocaine) in life, 0 otherwise	0.027	0.161	0.177	0.222
Inhale	Dummy variable (= 1) if ever used inhalants, such as glue or solvents in your lifetime, 0 otherwise	0.068	0.251	0.230	0.303
Other drugs	Dummy variable (= 1) if ever used any other type of illegal drug, such as lysergic acid diethylamide (LSD), phencyclidine (PCP), ecstasy, mushrooms, speed, ice, heroin, or pills, without a doctor's prescription in your lifetime, 0 otherwise	0.073	0.260	0.248	0.337
Inject	Dummy variable (= 1) if ever injected (shot up with a needle) any illegal drug, such as heroine or cocaine, in your lifetime, 0 otherwise	0.002	0.047	0.075	0.107

Variable	Description				
Allowance	Allowance per week	4.730	7.872	10.934	13.882
Welfare	Dummy variable (= 1) if any parent is on welfare, 0 otherwise	0.131	0.337	0.312	
Seatbelt	Dummy variable (= 1) if wears seatbelt every time in a car, 0 otherwise	0.863	0.344	0.357	0.487
Tattoo	Dummy variable (= 1) if had a permanent tattoo, 0 otherwise	0.039	0.193	0.203	0.402
Piercing	Dummy variable (= 1) if has both ears pierced, 0 otherwise	0.541	0.498	0.254	0.668
No chance to live until 35[a]	Dummy variable (= 1) if the perceived chance of living until age 35 years is less than 50%, 0 otherwise	0.036	0.187	0.246	0.341
Good chance to live until 35	Dummy variable (= 1) if the perceived chance of living until age 35years is more than 50%, 0 otherwise	0.871	0.336	0.404	0.565
Gut feeling — Yes[b]	Dummy variable (= 1) if agrees with the statement "I usually go with 'gut feeling' when making decisions without thinking too much about the consequences," 0 otherwise	0.403	0.491	0.600	0.823
Gut feeling — Neutral	Dummy variable (= 1) if neither agrees nor disagrees with the statement "I usually go with 'gut feeling' when making decisions without thinking too much about the consequences," 0 otherwise	0.212	0.409	0.539	0.755
Perceived IQ — Below average[c]	Dummy variable (= 1) if in comparison to other people of the same age, the perceived intelligence is below average, 0 otherwise	0.066	0.248	0.296	0.405
Perceived IQ — Average	Dummy variable (= 1) if in comparison to other people of the same age, the perceived intelligence is about average, 0 otherwise	0.393	0.488	0.523	0.717
GPA	Average GPA from math, science, history, and English classes	2.772	0.865	0.808	1.092
Chooses own friends	Dummy variable (= 1) if parents allow the respondent to decide with whom to hang around, 0 otherwise	0.853	0.354	0.447	0.758

Table 1. (*Continued*)

Definition		Wave I Cross Section Siblings and Twins		First Difference Siblings and Twins	First and Sibling Difference
		Mean	Standard deviation	Standard deviation	Standard deviation
Decides television time	Dummy variable (= 1) if parents allow respondent to decide how much television to watch, 0 otherwise	0.817	0.387	0.475	0.840
Decides own curfew on weekends	Dummy variable (= 1) if parents allow the respondent to decide about the time to be at home on weekend nights, 0 otherwise	0.317	0.465	0.546	0.985
Decides own curfew on weeknights	Dummy variable (= 1) if parents allow the respondent to decide about the time to be at home on weeknights, 0 otherwise	0.629	0.483	0.554	0.980
Height	Height in centimetres	168.378	10.600	4.715	6.177
Weight	Weight in kilogrammes	64.083	16.379	5.488	7.180
Alcohol available	Dummy variable (= 1) if alcohol is available at home, 0 otherwise	0.264	0.441	0.590	
Drugs available	Dummy variable (= 1) if illegal drugs are easily available at home, 0 otherwise	0.024	0.154	0.209	
		$n = 3,394$		$n = 3,039$	$n = 1,294$

[a]The omitted category is: Dummy variable (= 1) if the perceived chance of living until age 35 years is 50%, 0 otherwise.
[b]The omitted category is: Dummy variable (= 1) if disagrees with the statement "I usually go with 'gut feeling' when making decisions without thinking too much about the consequences," 0 otherwise.
[c]The omitted category is: Dummy variable (= 1) if in comparison to other people of the same age, the perceived intelligence is above average, 0 otherwise.

Table 2. The Impact of Drug Use on Crime: All Siblings (Including Twins).

	Without Measurement Error Correction				With Measurement Error Correction			
	Robbery	Burglary	Theft	Damage	Robbery	Burglary	Theft	Damage
First-Differenced Data								
Cocaine	0.146***	0.134***	0.140***	0.165***	0.115***	0.106***	0.111***	0.131***
	(0.047)	(0.040)	(0.052)	(0.053)	(0.037)	(0.032)	(0.042)	(0.042)
Inhale	0.168***	0.153***	0.191***	0.181***	0.133***	0.121***	0.152***	0.143***
	(0.036)	(0.033)	(0.043)	(0.042)	(0.029)	(0.026)	(0.034)	(0.033)
Other drugs	0.038	0.051*	0.079**	0.126***	0.024	0.032*	0.050**	0.080**
	(0.032)	(0.030)	(0.041)	(0.041)	(0.020)	(0.019)	(0.026)	(0.026)
Inject	0.199	0.331***	0.497***	0.312*	0.176	0.293***	0.440***	0.276*
	(0.158)	(0.114)	(0.111)	(0.161)	(0.139)	(0.101)	(0.097)	(0.142)
First-and-Sibling-Differenced Data								
Cocaine	0.142**	0.155***	0.154*	0.118	0.098**	0.107***	0.107*	0.082
	(0.062)	(0.058)	(0.086)	(0.086)	(0.043)	(0.04)	(0.06)	(0.06)
Inhale	0.128***	0.168***	0.210***	0.155**	0.094***	0.124***	0.155***	0.114**
	(0.041)	(0.043)	(0.065)	(0.063)	(0.03)	(0.032)	(0.048)	(0.046)
Other drugs	0.004	0.066	0.065	0.093	0.002	0.040	0.039	0.056
	(0.039)	(0.042)	(0.062)	(0.059)	(0.024)	(0.026)	(0.038)	(0.036)
Inject	0.403**	0.511***	0.754***	0.404**	0.357**	0.453***	0.669***	0.358**
	(0.159)	(0.151)	(0.135)	(0.163)	(0.141)	(0.134)	(0.120)	(0.145)

*, **, and *** indicate statistical significance at <10%, <5%, and <1% levels, respectively. Robust standard errors are in parentheses.

In the upper panel of the table, sample sizes range from 3,018 to 3,024 for robbery, burglary, from 3,015 to 3,021 for theft, and from 3,019 to 3,025 for damage. In the lower panel, sample sizes range from 1,304 to 1,307 for robbery, from 1,300 to 1,303 for burglary, from 1,304 to 1,307 for theft, and from 1,303 to 1,306 for damage. All regressions include 21 control variables as described in the text.

results significantly. In addition to drug use measures, the regressions include the following explanatory variables – seatbelt use, height, weight, grade point average (GPA), perceived IQ *below average*, perceived IQ *average*, welfare, alcohol at home, drugs at home, guns at home, allowance, tattoo, piercing, no chance to live until 35, good chance to live until 35, decides own curfew on weekends, decides own curfew on weeknights, decides television time, chooses own friends, gut feeling *yes*, and gut feeling *neutral*.

The bottom panel of Table 2 displays the results obtained from Eq. (4), which involves time differencing as well as sibling differencing. In this specification, variables pertaining to family attributes as well as the state dummies drop out as they do not vary between siblings. Both panels of

Table 2 display two sets of results. The left-hand side presents the results unadjusted for measurement error in drug use, where the right-hand side displays the results with measurement error adjustment. For each drug variable we used specific values of Φ obtained from the data.

As Table 2 demonstrates, drug use coefficients are positive and significantly different from zero in almost all cases in the top panel, and a similar picture emerges in the bottom panel, with the exception of the impact of *other drugs*.[15] We estimated all models with the inclusion of an additional variable which controls for the age difference between the siblings. The results remained the same.

Table 3 presents the results for twins. As in Table 2, the upper panel displays the results of the fixed-effects model (Eq. (3)), while the lower panel contains the results obtained from time and within-twin differencing.[16] We do not analyse injection because of the very small number of twins who injected drugs. To the extent that individual unobserved time-varying heterogeneity is the same between twins, this specification is represented by Eq. (5). Although the sample size goes down to about 450 in case of twins, drug use coefficients remain significant in many cases, even in models with time and twin differencing. For example, in the lower panel, cocaine consumption impacts theft and damage. The use of other types of drugs influences burglary, theft, and damage.

Table 4 displays the results for identical twins. Although there are only 400 observations in the fixed-effects model (top panel), with the exception of theft, we observe statistically significant associations between crime and drug measures. For example, robberies are influenced by using cocaine and inhalants, burglaries are influenced by inhalants, and damage is influenced by using inhalants and other drugs. In the lower panel where the results of fixed-effects and within-twin differences are reported, the sample size goes down to 176, and therefore, the coefficients are not estimated with precision.

The results in Tables 2–4 demonstrate the positive impact of drug use on crime. Although the precision of the estimated coefficients goes down as the sample gets smaller, the point estimates of individual drug variables are stable across specifications. We calculated the median point estimate for each drug category across crime types. In models with first differences, the median impact on crime of using cocaine is 11 percentage points. The impact of using inhalants is 13 percentage points. The median impacts on crime of other drugs and injecting drugs are 5 and 28 percentage points, respectively. In double-differenced models the median impacts are 11 percentage points for cocaine, 7 percentage points for inhalants, 7 percentage points for other drugs, and 41 percentage points for injection.

Table 3. The Impact of Drug Use on Crime: All Twins.

	Without Measurement Error Correction				With Measurement Error Correction			
	Robbery	Burglary	Theft	Damage	Robbery	Burglary	Theft	Damage
First-Differenced Data								
Cocaine	0.173*	0.111*	0.116	0.176**	0.131*	0.084*	0.088	0.133**
	(0.094)	(0.063)	(0.095)	(0.083)	(0.071)	(0.048)	(0.072)	(0.063)
Inhale	0.224***	0.146***	0.175**	0.237***	0.174***	0.113***	0.136**	0.183***
	(0.072)	(0.055)	(0.087)	(0.075)	(0.056)	(0.042)	(0.067)	(0.058)
Other drugs	0.110**	0.054	0.088	0.168**	0.065**	0.032	0.052	0.099**
	(0.057)	(0.044)	(0.066)	(0.070)	(0.034)	(0.026)	(0.039)	(0.042)
First-and-Twin–Differenced Data								
Cocaine	0.183	0.088	0.268**	0.278*	0.119	0.057	0.175**	0.181*
	(0.111)	(0.109)	(0.128)	(0.149)	(0.072)	(0.071)	(0.084)	(0.097)
Inhale	0.078	0.118	0.168	0.081	0.058	0.087	0.124	0.060
	(0.079)	(0.086)	(0.124)	(0.097)	(0.058)	(0.064)	(0.092)	(0.072)
Other drugs	0.082	0.121*	0.174*	0.258***	0.049	0.073*	0.105*	0.155***
	(0.068)	(0.063)	(0.098)	(0.09)	(0.041)	(0.038)	(0.059)	(0.054)

*, **, and *** indicate statistical significance at $<10\%$, $<5\%$, and $<1\%$ levels, respectively. Robust standard errors are in parentheses.
In the upper panel of the table, sample sizes are 1,023 or 1,024 for robbery and burglary, 1,022 or 1,023 for theft, and 1,024 or 1,025 for damage. In the lower panel, sample sizes are 452 or 453 for robbery and burglary, 453 or 454 for theft and damage. All regressions include 21 control variables as described in the text.

The results for injection should be taken with caution because of the small number of users in this case.

5.1. Undifferenced Estimates

To investigate the impact of unobserved heterogeneity, we estimated models using cross-sectional data from Wave I. We added a number of additional variables that could not be included in the first- and first-and-sibling-differenced models. These are time-invariant characteristics of the child and the parents, such as child's race, religious affiliation, gender, whether the child is born in the U.S.A., and parent education. We estimated these cross-sectional models with all siblings and twins. We also estimated them using all available observations (all children). The estimates for drug use variables were always positive and 4 to 7 percentage points larger than the ones obtained from first- and first-and-twin-differenced models reported earlier.

Table 4. The Impact of Drug Use on Crime: Identical Twins.

	Without Measurement Error Correction				With Measurement Error Correction			
	Robbery	Burglary	Theft	Damage	Robbery	Burglary	Theft	Damage
First-Differenced Data								
Cocaine	0.418***	0.108	0.055	0.121	0.331***	0.086	0.044	0.096
	(0.123)	(0.107)	(0.181)	(0.127)	(0.098)	(0.085)	(0.144)	(0.101)
Inhale	0.337***	0.167*	0.111	0.280**	0.229***	0.114*	0.076	0.190**
	(0.113)	(0.095)	(0.131)	(0.138)	(0.077)	(0.064)	(0.089)	(0.094)
Other drugs	0.095	0.072	0.016	0.208**	0.059	0.044	0.009	0.128**
	(0.103)	(0.072)	(0.120)	(0.105)	(0.064)	(0.043)	(0.074)	(0.065)
First-and-Twin-Differenced Data								
Cocaine	0.219	0.001	0.219	−0.195	0.16	0.000	0.16	−0.143
	(0.149)	(0.230)	(0.144)	(0.149)	(0.109)	(0.168)	(0.105)	(0.109)
Inhale	0.043	0.080	−0.105	−0.251	0.027	0.050	−0.065	−0.156
	(0.113)	(0.183)	(0.172)	(0.177)	(0.07)	(0.113)	(0.107)	(0.110)
Other drugs	0.043	0.138	0.173	0.249***	0.025	0.080	0.100	0.144***
	(0.088)	(0.1000)	(0.155)	(0.088)	(0.051)	(0.058)	(0.09)	(0.051)

*, **, and *** indicate statistical significance at $<10\%$, $<5\%$, and $<1\%$ levels, respectively.
Robust standard errors are in parentheses.
Sample size is 397 in the upper panel models and 176 in the lower panel models. All regressions include 21 control variables as described in the text.

This suggests that unobserved propensity to commit crime, which cannot be controlled for in cross-sectional regressions, tends to be positively correlated with drug use.

5.2. Reverse Causality

Although taking first and sibling (or twin) differences eliminates unobserved heterogeneity that would otherwise have been included in the error terms, it can still be argued that drug use may be influenced by reverse causality from property crimes if committing these crimes is associated with increased income. To account for this potential reverse causality, we considered a reduced-form drug use equation, where the instruments that impact the drug use of the juvenile include the following variables – whether at least one of the three best friends smokes at least one cigarette a day, whether at least one of the three best friends drinks alcohol at least once a month, and whether at least one of the three best friends uses marijuana at least once a month. While it can plausibly be argued that friends' consumption of

cigarette, alcohol, and marijuana may be correlated with own drug use, it is less obvious that these instruments are uncorrelated with own criminal activity. Unfortunately no better instruments are available. State- or county-level alcohol and drug prices are not viable candidates to identify the effect of drug use as they do not vary between siblings and twins. School-based policy variables are not useful either, because all twins and most siblings attend the same school.[17] Estimation of the double-differenced crime and double-differenced reduced-form drug use equations with full information maximum likelihood revealed that although the magnitudes of the estimated drug use coefficients are similar to those reported in Tables 2–4, most coefficients are not estimated with precision. The imprecision of the estimated coefficients is most likely due to weak instruments, but the data set does not include better instrument candidates. (The explicit specification is reported in Appendix B.)

In Table B1 in Appendix B, we report the coefficients for inhale and inject for siblings, twins, and identical twins. These are the drugs that created the most precise estimates. For siblings, the use of inhalants has a positive impact on burglary, theft, and damage. Injection has an impact on robbery and burglary. In case of all twins, injection has an impact on theft, and in the sample of identical twins injection influences burglary and theft.

6. CONCLUSION

The causal effect of drug use on crime has not yet been credibly established due to statistical difficulties. The propensity to use drugs may be correlated with unobserved attributes and characteristics of the individual. If these attributes, such as risk aversion or intelligence, have an influence on criminal activity, then estimates of drug use on crime are biased because of this confounding.

In this chapter we employ the Add Health data, which is a nationally representative panel data set of high school students in the U.S.A. that contains an oversample of siblings and twins. In addition to an unusually large number of interesting variables that aim to gauge personal characteristics, family background and family supervision, the data set contains detailed information about drug use and criminal activity of the juveniles. In particular, consumption of cocaine, injecting drugs, using inhalants, and using other drugs are measured. The crimes we analyse are robberies, burglaries, and thefts committed by juveniles and whether they damaged property.

The variation of drug use between siblings and twins allows us to exploit within-sibling differences of the first-differenced data. This enables us to filter out time-varying unobservables that are common to each household (therefore to each sibling). In addition, taking the twin differences of the first-differenced data enables us to eliminate the genetic component of criminal activity common to both twins. We adjust for measurement error in drug use with an algorithm that allows for non-symmetric measurement error.

The results indicate that drug use increases the propensity to commit crime. The median impact of injecting drugs on the probability of committing robbery, burglary and theft, and creating property damage is 41 percentage points, although this result should be taken with caution because it is based on small number of individuals who inject drugs. The median impact of cocaine is an increased criminal propensity of 11 percentage points. The use of inhalants generates a (median) 7 percentage point increase in the propensity to commit crime; and other drugs increase the propensity to commit crime by 6 percentage points.

NOTES

1. A few papers analysed micro-data in related context. Markowitz (2000) investigated the impact of alcohol and drug prices on violent crime using the National Crime Victimization Survey, Jofre-Bonet and Sindelar (2002) analysed the impact of drug treatment on criminal behaviour.

2. Theoretical justification of the inclusion of drug use in the crime equation can be found, among others, in Ehrlich (1973).

3. Empirical evidence from aggregate data on the impact on crime of deterrence, economic conditions and drug use can be found, among others, in Corman and Mocan (2002), Corman and Mocan (2000), Levitt (1998b, 1999), and Raphael and Winter-Ebmer (2001).

4. For a more detailed discussion, see Mocan and Corman (1998).

5. Note that this procedure also eliminates time-varying economic and deterrence measures as they do not exhibit variation between the individuals in the same household. Subjective probabilities of apprehension and conviction may differ between siblings. However, to the extent that they are approximated by such measures as arrest and conviction rates in the locality, they do not vary.

6. The reliability ratio is the proportion of individuals who are confirmed to have provided correct information about their drug use. Confirmation is typically based on drug tests.

7. Variations in the reporting rate did not change the results in a meaningful way.

8. It is assumed that $\nabla \Delta DR$ and ΔDR are uncorrelated with other covariates, which is supported by the data. Other covariates are assumed to contain no measurement error.

9. The Add Health project is a programme project designed by J. Richard Udry (PI) and Peter Bearman, and funded by grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation by the National Cancer Institute; the National Institute of Alcohol Abuse and Alcoholism; the National Institute on Deafness and Other Communication Disorders; the National Institute on Drug Abuse; the National Institute of General Medical Sciences; the National Institute of Mental Health; the National Institute of Nursing Research; the Office of AIDS Research, NIH; the Office of Behaviour and Social Science Research, NIH; the Office of the Director, NIH; the Office of Research on Women's Health, NIH; the Office of Population Affairs, DHHS; the National Center for Health Statistics, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Office of Public Health and Science, DHHS; the Office of the Assistant Secretary for Planning and Evaluation, DHHS; and the National Science Foundation. Persons interested in obtaining data files from The National Longitudinal Study of Adolescent Health should contact Add Health Project, Carolina Population Center, 123 West Franklin Street, Chapel Hill, NC 27516-2524, U.S.A. (e-mail: addhealth@unc.edu).

10. Participating high schools were asked to identify junior high or middle schools that were expected to provide at least five students to the entering class of the high school. These schools are called feeder schools. Their probability of selection was proportional to the percentage of the high school's entering class that came from that feeder.

11. The sample for the Wave II in-home interview was composed of the respondents of the Wave I in-home interview with the following exceptions: A respondent who was in the 12th grade in Wave I and who was not part of the genetic sample was not interviewed in Wave II. Respondents who were only in Wave I's disabled sample were not re-interviewed.

12. The proportion of twins in total births has been rising steadily over the last two decades. When most of the adolescents of the sample were born around 1980, twin births were about 2 percent of total births (*National Vital Statistics Report*, 1999).

13. For less sensitive questions, the interviewer read the questions aloud, and entered the respondent's answers.

14. Questions in Wave II are worded as "Since the last interview...". Therefore, the change in behaviour between the two waves is easily identifiable.

15. Note that the relative sample size of the first- and first-and-sibling-differenced data depends on the number of siblings in households. For example, if a household consists of two siblings, the first-differenced (time-differenced) data will contain two observations, and the first-and-sibling-differenced data will contain one observation. On the other hand, if the household consists of three siblings A, B, and C, the first-differenced data will contain three observations, and the first-and-sibling-differenced data will also contain three observations (it will consist of ΔA–ΔB, ΔA–ΔC, and ΔB–ΔC, where ΔA is a first-differenced variable of sibling A). In case of a household with four siblings, the first-differenced data have four observations and first-and-sibling-differenced data have six observations.

16. The sample size of the first-and-twin-differenced data is not half of the first-differenced twin sample because of missing values in some variables.

17. Grossman, Kaestner, and Markowitz (2002) highlight the same point in their analysis of drug use on teenage sexual activity.

ACKNOWLEDGEMENTS

REFERENCES

Almond, D., Chay, K. Y., & Lee, D. S. (2002). *Does low birth weight matter? Evidence from the U.S. population of twin births.* Center for Labor Economics Working Paper no. 53. University of California, Berkeley.

Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic returns to schooling from a new sample of twins. *American Economic Review, 84*(5), 1157–1173.

Ashenfelter, O., & Zimmerman, D. J. (1997). Estimates of the returns to schooling from sibling data: Fathers, sons, and brothers. *The Review of Economics and Statistics, 79*(1), 1–9.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1994). Endowments and the allocation of schooling in the family and in the marriage market: The twins experiment. *The Journal of Political Economy, 102*(6), 1131–1174.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1996). College choice and wages: Estimates using data on female twins. *The Review of Economics and Statistics, 78*(1), 672–685.

Chaiken, J. M., & Chaiken, M. R. (1990). Drugs and predatory crime. In: M. Tonry & J. Q. Wilson (Eds), *Drug and crime.* Chicago: University of Chicago Press.

Cook, P., & Ludwig, J. (2002). *The effects of gun prevalence on burglary: Deterrence vs. inducement.* NBER Working Paper no. 8926.

Corman, H., & Mocan, H. N. (2000). A time-series analysis of crime, deterrence, and drug abuse in New York city. *American Economic Review, 90*(3), 584–604.

Corman, H., & Mocan, H. N. (2002). *Carrots, sticks and broken windows.* NBER Working Paper no. W9061.

Currie, J., & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics, 188*(4), 1495–1532.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy, 81*(3), 521–565.

Goldstein, P. (1985). The drugs/violence nexus: A tripartite conceptual framework. *Journal of Drug Issues, 15*(4), 493–506.

Griliches, Z., & Hausman, J. (1985). Errors in variables in panel data. *Journal of Econometrics, 31*, 93–118.

Grogger, J., & Willis, M. (2000). The emergence of crack cocaine and the rise in urban crime rates. *The Review of Economics and Statistics, 82*(4), 519–529.

Grossman, M., Kaestner, R., & Markowitz, S. (2002). *Get high and get stupid: The effect of alcohol and marijuana use on teen sexual behavior.* NBER Working Paper no. 9216.

Gruber, J. (Ed.) (2001). *Risky behavior among youths.* Chicago and London: University of Chicago Press.

Harrison, L. D. (1992). The drug–crime nexus in the USA. *Contemporary Drug Problems, 19*(2), 203–246.

Jofre-Bonet, M., & Sindelar, J. L. (2002). *Drug treatment as a crime fighting tool.* NBER Working Paper no. W9038.

Johnson, M. E., Fisher, D. G., et al. (2000). Reliability and validity of not-in-treatment drug users' follow-up self-reports. *Aids and Behavior, 4*(4), 373–380.

Levitt, S. D. (1998a). Why do increased crime arrest rates appear to reduce crime: Deterrence, incapacitation, or measurement error? *Economic Inquiry, 36*(3), 353–372.

Levitt, S. D. (1998b). Juvenile crime and punishment. *Journal of Political Economy, 106*(6), 1156–1185.

Levitt, S. D. (1999). The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *The Quarterly Journal of Economics, 111*(2), 319–351.

Markowitz, S. (2000). *An economic analysis of alcohol, drugs, and violent crime in the National Crime Victimization Survey.* NBER Working Paper no. W7982.

Miller, P., Mulvey, C., & Martin, N. (1995). What do twins studies reveal about the economic returns to education? A comparison of australian U.S. findings. *American Economic Review, 85*(3), 586–599.

Mocan, N. H., & Corman, H. (1998). An economic analysis of drug use and crime. *Journal of Drug Issues, 28*(3), 613–629.

Mocan, N. H., & Rees, D. (2005). Economic conditions, deterrence, and juvenile crime: Evidence from micro data. *American Journal of Law and Economics, 7*(2).

National Vital Statistics Report. (1999). Vol. 47, No. 24. National Center for Health Statistics, Hyattsville, Maryland.

Raphael, S., & Winter-Ebmer, W. (2001). Identifying the effect of unemployment on crime. *The Journal of Law and Economics, 44*, 259–283.

Weiss, R. D., Najavits, L. M., et al. (1998). Validity of substance use self-reports in dually diagnosed outpatients. *American Journal of Psychiatry, 155*(1), 127–128.

Grossman, M., Kaestner, R., & Markowitz, S. (2002). *Get high and get stupid: The effect of alcohol and marijuana use on teen sexual behavior.* NBER Working Paper no. 9216.

Hu, W. Y. (1999). Child support, welfare delinquency, and woman's labor supply. *Journal of Human Resources, 34*(1), 71–103.

Mocan, N. H., & Tekin, E. (2003). Nonprofit sector and part-time work: An analysis of employer–employee matched data of child care workers. *Review of Economics and Statistics, 85*(1), 38–50.

Mocan, N. H., Tekin, E., & Zax, J. S. (2004). The demand for medical care in urban China. *World Development, 32*(2), 289–304.

APPENDIX A

General Framework

$DR_{it} = DR_{it}^* + v_{it}$, where DR_{it}^* is the actual drug use $(0 = \text{No}, 1 = \text{Yes})$ and DR_{it} is the reported drug use $(0 = \text{No}, 1 = \text{Yes})$.

Let probability distribution of v_{it} be:

$$\text{Prob}(DR_{it} = 1, DR_{it}^* = 1) = p_1$$

$$\text{Prob}(DR_{it} = 1, DR_{it}^* = 0) = 0$$

$$\text{Prob}(DR_{it} = 0, DR_{it}^* = 0) = p_2$$

$$\text{Prob}(DR_{it} = 0, DR_{it}^* = 1) = q$$

In other words:

DR_{it}	DR_{it}^*	v_{it}	$\text{Prob}(v_{it})$
1	1	0	p_1
1	0	1	0
0	0	0	p_2
0	1	-1	Q

$$p_1 + p_2 + q = 1$$

$$E(v) = -q$$

$$E(DR^*) = 1 - p_2$$

$$\text{var}(v) = q - q^2$$

$$\text{var}(DR^*) = p_2 - p_2^2$$

$$\text{cov}(v, DR^*) = -p_2 q$$

$$\text{var}(DR) = p_1 - p_1^2$$

Probability Limit of the Coefficient of Drug Use in First-Differenced Data
The probability limit is:

$$p \lim \hat{\beta} = \beta \frac{\text{var}(\Delta DR_t^*) + \text{cov}(\Delta v_t, \Delta DR_t^*)}{\text{var}(\Delta DR_t)}$$

Note that by definition:

$$\text{var}[\Delta DR_t^*] = \text{var}[\Delta DR_t] - \text{var}[\Delta v_t] - 2\text{cov}[\Delta DR_t^*, \Delta v_t]$$

Thus:

$$\text{cov}(\Delta DR_t^*, \Delta v_t) = \frac{1}{2}\text{var}(\Delta DR_t) - \frac{1}{2}\text{var}(\Delta v_t) - \frac{1}{2}\text{var}(\Delta DR_t^*)$$

Substitution for $\text{cov}(\Delta DR^*, \Delta v)$ provides:

$$p\lim \hat{\beta} = \beta\, \frac{0.5\big[\text{var}(\Delta DR_t) + \text{var}(\Delta DR_t^*) - \text{var}(\Delta v_t)\big]}{\text{var}(\Delta DR_t)}$$

Note that:

$$\text{var}(\Delta DR_t) = \text{var}(DR_t - DR_{t-1}) = \text{var}(DR_t) + \text{var}(DR_{t-1}) - 2\text{cov}(DR_t, DR_{t-1})$$

Assuming that DR is covariance stationary, that is:

$$\text{var}(DR_t) = \text{var}(DR_{t-1})$$

$$\text{var}(\Delta DR_t) = 2\sigma_{DR}^2(1 - \rho_{DR_t, DR_{t-1}})$$

where $\sigma_{DR}^2 = \text{var}(DR_t)$.

Similarly:

$$\text{var}(\Delta DR_t^*) = 2\sigma_{DR^*}^2(1 - \rho_{DR_t^*, DR_{t-1}^*}) \ \text{ and } \ \text{var}(\Delta v_t) = 2\sigma_v^2(1 - \rho_{v_t v_{t-1}})$$

where $\sigma_v^2 = \text{var}(v_t)$ and $\sigma_{DR^*}^2 = \text{var}(DR_t^*)$.

Substituting the variances of DR, DR^* and v into the probability limit formula, one obtains:

$$p\lim \hat{\beta} = \beta\, \frac{0.5[2\sigma_{DR}^2(1 - \rho_{DR_t, DR_{t-1}}) + 2\sigma_{DR^*}^2(1 - \rho_{DR_t^*, DR_{t-1}^*})2\sigma_v^2(1 - \rho_{v_t, v_{t-1}})]}{2\sigma_{DR}^2(1 - \rho_{DR_t, DR_{t-1}})}$$

$$\text{(A1)}$$

Note that:

$$\text{cov}(DR_t, DR_{t-1}) = \text{cov}(DR_t^* + v_t, DR_{t-1}^* + v_{t-1})$$
$$= \text{cov}(DR_t, DR_{t-1}) = \text{cov}(DR_t^*, DR_{t-1}^*) + \text{cov}(DR_{t-1}^*, v_t)$$
$$+ \text{cov}(DR_t^*, v_{t-1}) + \text{cov}(v_t, v_{t-1}) \qquad \text{(A2)}$$

Also:

$$\text{cov}(DR_t, DR_{t-1}) = \rho_{DR_t, DR_{t-1}}\, \sigma_{DR}^2$$

$$\text{cov}(DR_t^*, DR_{t-1}^*) = \rho_{DR_t^*, DR_{t-1}^*}\, \sigma_{DR^*}^2$$

$$\text{cov}(v_t, v_{t-1}) = \rho_{v_t, v_{t-1}}\, \sigma_v^2$$

Therefore, one can rewrite Eq. (A2) as:

$$\rho_{DR_t,DR_{t-1}}\sigma^2_{DR} = \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*} + \rho_{v_t,v_{t-1}}\sigma^2_v + \text{cov}(DR_{t-1}^*, v_t) + \text{cov}(DR_t^*, v_{t-1})$$

As:

$$\text{cov}(DR_{t-1}^*, v_t) = \text{cov}(DR_{t-1}^*, DR_t - DR_t^*) = \text{cov}(DR_{t-1}^*, DR_t) - \text{cov}(DR_{t-1}^*, DR_t^*)$$

Assuming $\text{cov}(DR_{t-1}^*, DR_t) = 0$, one obtains:

$$\text{cov}(DR_{t-1}^*, v_t) = -\rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*}$$

Similarly:

$$\text{cov}(DR_t^*, v_{t-1}\,\text{cov}(DR_t^*, DR_{t-1} - DR_{t-1}^*) = \text{cov}(DR_t^*, DR_{t-1}) - \text{cov}(DR_t^*, DR_{t-1}^*)$$

Assuming $\text{cov}(DR_t^*, DR_{t-1}) = 0$, one gets:

$$\text{cov}(DR_t^*, v_{t-1}) = -\rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*}$$

Therefore, Eq. (A2) is equivalent to:

$$\rho_{DR_t,DR_{t-1}}\sigma^2_{DR} = \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*} + \rho_{v_t,v_{t-1}}\sigma^2_v - \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*} - \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*}$$

$$\rho_{DR_t,DR_{t-1}}\sigma^2_{DR} = \rho_{v_t,v_{t-1}}\sigma^2_v - \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*}$$

Solving for $\rho_{v_t,v_{t-1}}$ yields:

$$\rho_{v_t,v_{t-1}} = \frac{\rho_{DR_t,DR_{t-1}}\sigma^2_{DR} + \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*}}{\sigma^2_v} \tag{A3}$$

Substitution of Eq. (A3) into Eq. (A1) gives:

$$p\lim\hat{\beta} =$$

$$\beta\frac{0.5\left[2\sigma^2_{DR}(1 - \rho_{DR_t,DR_{t-1}}) + 2\sigma^2_{DR^*}(1 - \rho_{DR_t^*,DR_{t-1}^*}) - 2\sigma^2_v\left(1 - \frac{\rho_{DR_t,DR_{t-1}}\sigma^2_{DR} + \rho_{DR_t^*,DR_{t-1}^*}\sigma^2_{DR^*}}{\sigma^2_r}\right)\right]}{2\sigma^2_{DR}(1 - \rho_{DR_t,DR_{t-1}})}$$

Simplification yields:

$$p\lim\hat{\beta} = \beta\frac{\sigma^2_{DR} + \sigma^2_{DR^*} - \sigma^2_v}{2\sigma^2_{DR}(1 - \rho_{DR_t,DR_{t-1}})} = \frac{p_1 p_2}{(p_1 - p_1^2)(1 - \rho_{DR_t,DR_{t-1}})}$$

Therefore:

$$p \lim \hat{\beta} = \beta \, \frac{p_1 p_2}{(p_1 - p_1^2)(1 - \rho_{DR_t, DR_{t-1}})}$$

Probability Limit of the Coefficient of Drug Use in Double-Differenced Data

The probability limit is:

$$p \lim \hat{\beta} = \beta \frac{\text{var}(\nabla \Delta DR_{jit}^*) + \text{cov}(\nabla \Delta v_{jit}, \nabla \Delta DR_{jit}^*)}{\text{var}(\nabla \Delta DR_{jit})}$$

By definition:

$$\text{var}(\nabla \Delta DR_{jit}^*) = \text{var}(\nabla \Delta DR_{jit}) - \text{var}(\nabla \Delta v_{jit}) - 2\text{cov}(\nabla \Delta DR_{jit}^*, \nabla \Delta v_{jit})$$

Therefore:

$$\text{cov}(\nabla \Delta DR_{jit}^*, \nabla \Delta v_{jit}) = \frac{1}{2} \text{var}(\nabla \Delta DR_{jit}) - \frac{1}{2} \text{var}(\nabla \Delta v_{jit}) - \frac{1}{2} \text{var}(\nabla \Delta DR_{jit}^*)$$

and

$$p \lim \hat{\beta} = \beta \, \frac{0.5[\text{var}(\nabla \Delta DR_{jit}) + \text{var}(\nabla \Delta DR_{jit}^*) - \text{var}(\nabla \Delta v_{jit})]}{\text{var}(\nabla \Delta DR_{jit})} \tag{A4}$$

Note that:

$$\begin{aligned}
\text{var}(\nabla \Delta DR_{jit}) &= \text{var}(DR_{j2t} - DR_{j2t-1} - DR_{j1t} + DR_{j1t-1}) \\
&= \text{var}(DR_{j2t}) + \text{var}(DR_{j2t-1}) + \text{var}(DR_{j1t}) + \text{var}(DR_{j1t-1}) \\
&\quad - 2\text{cov}(DR_{j2t}, DR_{j2t-1}) - 2\text{cov}(DR_{j2t}, DR_{j1t}) \\
&\quad + 2\text{cov}(DR_{j2t}, DR_{j1t-1}) + 2\text{cov}(DR_{j2t-1}, DR_{j1t}) \\
&\quad - 2\text{cov}(DR_{j2t-1}, DR_{j1t-1}) - 2\text{cov}(DR_{j1t}, DR_{j1t-1})
\end{aligned}$$

where the subscripts 1 and 2 represent the first and the second individuals in sibling (twin) pair *j*, and $t-1$ and t represent the first and the second time periods. Suppressing *j* for ease of notation, we can rewrite the above equation as follows:

$$\begin{aligned}
\text{var}(\nabla \Delta DR_{it}) &= 4\sigma_{DR}^2 - 2\rho_{DR_{2t}, DR_{2t-1}} \sigma_{DR}^2 - 2\rho_{DR_{2t}, DR_{1t}} \sigma_{DR}^2 \\
&\quad + 2\rho_{DR_{2t}, DR_{1t-1}} \sigma_{DR}^2 + 2\rho_{DR_{2t-1}, DR_{1t}} \sigma_{DR}^2 \\
&\quad - 2\rho_{DR_{2t-1}, DR_{1t-1}} \sigma_{DR}^2 - 2\rho_{DR_{1t}, DR_{1t-1}} \sigma_{DR}^2
\end{aligned}$$

Assuming constant variances of drug use between siblings and over time, one obtains:

$$\text{var}(\nabla \Delta DR_{it}) = \sigma_{DR}^2 (\Phi + 4)$$

where

$$\Phi = -2\rho_{DR_{2t},DR_{2t-1}} - 2\rho_{DR_{2t},DR_{1t}} + 2\rho_{DR_{2t},DR_{1t-1}} + 2\rho_{DR_{2t-1},DR_{1t}} - 2\rho_{DR_{2t-1},DR_{1t-1}}$$
$$- 2\rho_{DR_{1t},DR_{1t-1}} \tag{A5}$$

Similarly:

$$\text{var}(\nabla\Delta DR_{it}^*) = \sigma_{DR^*}^2(\Psi + 4)$$

where

$$\Psi = -2\rho_{DR_{2t}^*,DR_{2t-1}^*} - 2\rho_{DR_{2t}^*,DR_{1t}^*} + 2\rho_{DR_{2t}^*,DR_{1t-1}^*} + 2\rho_{DR_{2t-1}^*,DR_{1t}^*} - 2\rho_{DR_{2t-1}^*,DR_{1t-1}^*}$$
$$- 2\rho_{DR_{1t}^*,DR_{1t-1}^*}$$

and,

$$\text{var}(\nabla\Delta v_{it}) = \sigma_v^2(\Omega + 4)$$

where

$$\Omega = -2\rho_{v_{2t},v_{2t-1}} - 2\rho_{v_{2t},v_{1t}} + 2\rho_{v_{2t},v_{1t-1}} + 2\rho_{v_{2t-1},v_{1t}} - 2\rho_{v_{2t-1},v_{1t-1}} - 2\rho_{v_{1t},v_{1t-1}}$$

Substituting the expressions for the variances in Eq. (A4) gives:

$$p\lim\hat{\beta} = \beta\,\frac{\text{var}(DR_{it}^*)(4 + \Psi) + \frac{1}{2}(\text{var}(DR_{it})(4 + \Phi) - \text{var}(v_{it})(4 + \Omega) - \text{var}(DR_{it}^*)(4 + \Psi))}{(4 + \Phi)\text{var}(DR_{it})}$$

$$p\lim\hat{\beta} = \beta\,\frac{1}{2}\,\frac{(\text{var}(DR_{it})(4 + \Phi) - \text{var}(v_{it})(4 + \Omega) + \text{var}(DR_{it}^*)(4 + \Psi))}{\text{var}(DR_{it})(4 + \Phi)}$$

Note that:

$$\text{cov}(DR_{1t}, DR_{2t-1}) = \text{cov}(DR_{1t}^* + v_{1t}, DR_{2t-1}^* + v_{2t-1})$$

$$\text{cov}(DR_{1t}, DR_{2t-1}) = \text{cov}(DR_{1t}^*, DR_{1t}^*) + \text{cov}(DR_{1t}^*, v_{1t}) + \text{cov}(v_{2t-1}, DR_{2t-1}^*)$$
$$+ \text{cov}(v_{2t-1}, v_{2t-1})$$

As

$$\text{cov}(DR_{1t}, DR_{2t-1}) = \rho_{DR_{1t},DR_{2t-1}}\sigma_{DR}^2$$

$$\text{cov}(DR_{1t}^*, DR_{2t-1}^*) = \rho_{DR_{1t}^*,DR_{2t-1}^*}\sigma_{DR^*}^2$$

$$\text{cov}(v_{1t}, v_{2t-1}) = \rho_{1t,v_{2t-1s}}\sigma_v^2$$

we obtain

$$\text{cov}(DR_{1t}, DR_{2t-1}) = \rho_{DR_{1t}, DR_{2t-1}} \sigma^2_{DR} = \rho_{DR^*_1, DR^*_{2t-1}} \sigma^2_{DR^*} + \rho_{v_{1t}, v_{2t-1}} \sigma^2_v$$
$$+ \text{cov}(DR^*_{1t}, v_{2t-1}) + \text{cov}(v_{1t}, DR^*_{2t-1}) \qquad (A6)$$

The third term in Eq. (A6) is:

$$\text{cov}(DR^*_{1t}, v_{2t-1}) = \text{cov}(DR^*_{1t}, DR_{2t-1} - DR^*_{2t-1})$$
$$= \text{cov}(DR^*_{1t}, DR_{2t-1}) - \text{cov}(DR^*_{1t}, DR^*_{2t-1})$$

Assuming $\text{cov}(DR^*_{1t}, DR_{2t-1}) = 0$, we obtain:

$$\text{cov}(DR^*_{1t}, v_{2t-1}) = -\text{cov}(DR^*_{1t}, DR^*_{2t-1}) = -\rho_{DR^*_{1t}, DR^*_{2t-1}} \sigma^2_{DR^*} \qquad (A7)$$

The fourth term in Eq. (A6) is:

$$\text{cov}(v_{1t}, DR^*_{2t-1}) = \text{cov}(DR_{1t} - DR^*_{1t}, DR^*_{2t-1})$$
$$= \text{cov}(DR_{1t}, DR^*_{2t-1}) - \text{cov}(DR^*_{1t}, DR^*_{2t-1})$$

Again, assuming $\text{cov}(DR_{1t}, DR^*_{2t-1}) = 0$, we obtain:

$$\text{cov}(v_{1t}, DR^*_{2t-1}) = -\text{cov}(DR^*_{1t}, DR^*_{2t-1}) = -\rho_{DR^*_{1t}, DR^*_{2t-1}} \sigma^2_{DR^*} \qquad (A8)$$

Substituting Eqs. (A7) and (A8) into Eq. (A6), one obtains:

$$\rho_{DR_{1t}, DR_{2t-1}} \sigma^2_{DR} = -\rho_{DR^*_{1t}, DR^*_{2t-1}} \sigma^2_{DR^*} + \rho_{v1, v_{2t-1}} \sigma^2_v$$

or

$$\rho_{DR_{1t}, DR_{2t-1}} = \frac{-\rho_{DR^*_{1t}, DR_{2t-1}} \sigma^2_{DR^*} + \rho_{v_{1t}, v_{2t-1}} \sigma^2_v}{\sigma^2_{DR}}$$

One can obtain similar expressions for:

$$\rho_{DR_{2t}, DR_{1t}}, \quad \rho_{DR_{2t}, DR_{1t-1}} \rho_{DR_{2t-1}, DR_{1t}}, \quad \rho_{DR_{2t-1}, DR_{1t-1}}, \quad \rho_{DR_{1t}, DR_{1t-1}}$$

Substituting each of these correlation coefficients into Eq. (A5) one obtains

$$\Phi = \frac{\sigma^2_v \Omega - \sigma^2_{DR^*} \Psi}{\sigma^2_{DR}}$$

$$\Omega = \frac{\sigma^2_{DR} \Phi + \sigma^2_{DR^*} \Psi}{\sigma^2_v}$$

Substitution of Ω into Eq. (A4) gives:

$$p \lim \hat{\beta} = \beta \, \frac{1}{2} \, \frac{\left[\sigma_{DR}^2 (4 + \Phi) - \sigma_v^2 \left(4 + \frac{\sigma_{DR}^2 \Phi + \sigma_{DR*}^2 \Psi}{\sigma_v^2} \right) + \sigma_{DR*}^2 (4 + \Psi) \right]}{\sigma_{DR}^2 (4 + \Phi)}$$

which simplifies to:

$$p \lim \hat{\beta} = \beta \, \frac{2[\sigma_{DR}^2 - \sigma_v^2 + \sigma_{DR*}^2]}{\sigma_{DR}^2 (4 + \Phi)}$$

Substituting the expressions for variances one obtains:

$$p \lim \hat{\beta} = \beta \, \frac{4 p_1 p_2}{(p_1 - p_1^2)(4 + \Phi)}$$

APPENDIX B

To address potential reverse causality, we specify Eqs. (B1) and (B2):

$$\nabla \Delta CR_{jit} = \delta \nabla \Delta X_{jit} + \beta \nabla \Delta DR_{jit} + \mu_{jit} \qquad \text{(B1)}$$

$$\nabla \Delta DR_{jit} = B \nabla \Delta X_{jit} + \zeta \nabla \Delta Z_{jit} + \eta_{jit} \qquad \text{(B2)}$$

In Eq. (B2) Z represents the instruments that impact the drug use of the juvenile which include the following variables – whether at least one of the three best friends smokes at least one cigarette a day, whether at least one of the three best friends drinks alcohol at least once a month, and whether at least one of the three best friends uses marijuana at least once a month. While it can plausibly be argued that friends' consumption of cigarette, alcohol, and marijuana may be correlated with own drug use, it is less obvious that these instruments are uncorrelated with own criminal activity. Unfortunately, no better instruments are available. State- or county-level alcohol and drug prices are not viable candidates to identify the effect of drug use as they do not vary between siblings and twins. School-based policy variables are not useful either, because all twins and most siblings attend the same school.[B1]

Eqs. (B1) and (B2) are estimated jointly using full information maximum likelihood. We allow for a correlation between the error terms in Eqs. (B1) and (B2) using the discrete factor method (DFM). The DFM assumes that the correlation between these two equations is governed by a common factor, the distribution of which can be approximated by a step function. The common

Table B1.   The Impact of Drug Use on Crime with Reverse Causality. Full Information Maximum Likelihood Estimates of Drug Use.

		Without Measurement Error Correction				With Measurement Error Correction			
		Robbery	Burglary	Theft	Damage	Robbery	Burglary	Theft	Damage
All Siblings	Inhale	0.136	0.179*	0.196*	0.176*	0.100	0.132*	0.144*	0.130*
		(0.095)	(0.095)	(0.095)	(0.095)	(0.07)	(0.07)	(0.07)	(0.07)
	Inject	0.401*	0.515*	0.768	0.424*	0.356*	0.457*	0.681	0.376*
		(0.264)	(0.264)	(0.264)	(0.264)	(0.234)	(0.234)	(0.234)	(0.234)
All Twins	Inhale	0.064	0.137	0.128	0.096	0.047	0.101	0.095	0.071
		(0.179)	(0.179)	(0.18)	(0.18)	(0.132)	(0.132)	(0.133)	(0.133)
Identical Twins	Inhale	−0.003	0.071	−0.133	−0.206	−0.002	0.044	−0.082	−0.128
		(0.338)	(0.336)	(0.335)	(0.335)	(0.21)	(0.208)	(0.208)	(0.208)

*, **, and *** indicate statistical significance at $<10\%$, $<5\%$, and $<1\%$ levels, respectively.
Robust standard errors are in parentheses.

discrete factor is then integrated out of the model as in the standard random effects approach. This method is less restrictive than the specifying functional form, such as joint normality. See Hu (1999), Mocan, Tekin, and Zax (2004), and Mocan and Tekin (2003) for applications of the DFM.

NOTE

B1. Grossman, Kaestner, and Markowitz (2002) highlight the same point in their analysis of drug use on teenage sexual activity.