Economics of Education Review 31 (2012) 961–969

Contents lists available at SciVerse ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Adolescent drug use and the deterrent effect of school-imposed penalties

G.R. Waddell[a,b,*,1]

[a] Department of Economics, University of Oregon, Eugene, OR 97403-1285, USA
[b] IZA, P.O. Box 7240, Bonn 53072, Germany

ARTICLE INFO

Article history:
Received 24 June 2011
Received in revised form 5 July 2012
Accepted 5 July 2012

JEL classification:
I1
I2
K42

Keywords:
Drug
Crime
Adolescent
Risky behavior
Expulsion

ABSTRACT

Estimates of the effect of school-imposed penalties for drug use on a student's consumption of marijuana are biased if both are determined by unobservable school or individual attributes. Reverse causality is also a potential challenge to retrieving estimates of the causal relationship, as the severity of school sanctions may simply reflect the need for more-severe sanctions. Using the National Longitudinal Study of Adolescent Health, I offer an instrumental-variables approach to retrieving an estimate of the causal response of marijuana use to sanctions and thereby demonstrate the efficacy of school-imposed penalties as a deterrent to adolescent drug use. This suggests that school sanctions may have important long-run benefits.

1. Introduction

There is a large literature that documents the long-run costs associated with drug use, and the implied gains to lowering adolescent drug use are not difficult to establish from existing research. For example, in an instrumental-variables design, DeSimone (2002) shows that employment probabilities are substantially reduced by marijuana and cocaine consumption. Considering the relationship between marijuana use in high school and future earnings, Ringel, Ellickson, and Collins (2006) suggest that a significant part of the negative relationship between substance use and earnings reflects an indirect mechanism by which early marijuana use affects human capital accumulation, which in turn affects earnings. The focus of this analysis is on the role of school policy in determining student's consumption of marijuana—the most common illicit substance used by adolescents.

Of course, the consequences of substance use are not restricted to labor-market outcomes. For example, Kaestner (1995) shows that drug users tend to delay marriage and, conditional on marriage, experience shorter marriage durations. Markowitz (2000) suggests that marijuana may also cause increased engagement in physical fights. Substance use has also been identified as a leading causal factor in suicidal thoughts and behaviors (Markowitz, Chatterji, Kaestner, & Dave, 2002). Clearly, there is the potential for a significant down side associated with adolescent drug use, which should motivate policy makers in their stewardship of adolescents.

* Correspondence address: Department of Economics, University of Oregon, Eugene, OR 97403-1285, USA. Tel.: +1 541 346 1259; fax: +1 541 346 1243.
E-mail address: waddell@uoregon.edu

1 I thank Jason Lindo, Joe Price, Larry Singell, and seminar participants at the 2010 Western Economic Association International meetings. Any errors remain my responsibility.

0272-7757/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.econedurev.2012.07.002

Somewhat surprisingly, however, the role of school policy in a student's choice to consume drugs has largely been ignored in the economics literature. Yet, among the established results in the literature, there are several empirical patterns that raise particular concern around this shortcoming. For example, Chatterji (2006) shows that marijuana use in high school is associated with lower levels of educational attainment, and concludes with an appropriate conjecture that "public policies that are effective in reducing substance use during high school should have some impact on educational attainment." Based on a relationship between marijuana use and lower high-school graduation rates, Yamada, Kendix, and Yamada (1998) also conclude with the suggestion that "high-school-based preventive programs which discourage alcohol consumption and marijuana use are highly recommended." The literature has also documented that the earlier one starts using a particular drug the less likely one is to stop using that drug (van Ours, 2006), which further supports considering the role of schools in influencing drug use. To the extent one believes that marijuana is a gateway to other (harder) substances, the benefits to curbing adolescent marijuana use also include mitigating this potential escalation and any costs associated with such escalation.[2]

In the end, the existing literature leaves us largely uninformed about the relationship between school policy and the substance use of youth.[3] Yet, there is reason to consider the influence of school policy in this regard, as educational institutions are well positioned to influence adolescent choices.[4]

Moreover, beginning with the Regan-Bush era drug-enforcement policies, although not without controversy, there are still growing numbers of schools moving toward "zero tolerance" policies with respect to drugs and alcohol, so much so that the application of zero tolerance is now quite common (Heaviside, Rowand, Williams, & Farris, 1998).[5] This analysis contributes to this area of policy by introducing estimates of the causal relationship between use and the severity of drug-related policy—the potential for school policy to influence a student's consumption of marijuana.

Specifically, I will model one's marijuana use as a function of the the penalty one's school would impose if one were to be caught consuming an illegal drug. In proceeding toward a preferred specification, I will be transparent about the empirical regularities in the data and report simple OLS specifications that highlight the potential endogeneity of penalties in such an environment.

For example, OLS estimates of the effect of school-imposed penalties for drug use on a student's consumption of marijuana would be biased if both are determined by unobservable school or individual attributes. That the severity of school sanctions may simply reflect the need for more-severe sanctions (i.e., drug use is high) is also a challenge to OLS estimates as this imparts positive bias. Alternatively, schools with well-behaved students and little marijuana use may have severe penalties because they so seldom need to follow through on them. This would introduce negative bias in OLS regressions.

Given the likely endogeneity of punishment levels, I will adopt an instrumental-variables approach to retrieving an estimate of the causal influence of sanctions on student behavior and, in the end, demonstrate the efficacy of school-imposed sanctions—stiffer sanctions for drug use cause students to be less likely to consume marijuana. In particular, the preferred estimates are identified off of variation in penalties imposed on second-time drug offenders across schools that issue the same penalties to first-time offenders. In this scenario, I instrument for the second-offence penalty with measures of how much the school *escalates* its penalties between first- and for second-time offences in non-drug areas of discipline.

In Section 2, I detail the data used in this analysis. In Section 3, I develop the empirical model and formally define the instrumental variables to be used to recover causal estimates of school-imposed penalties on marijuana use. I offer some discussion in Section 4 followed by concluding remarks in Section 5.

2. Data

2.1. Source

For our purpose, the National Longitudinal Study of Adolescent Health is a particularly fitting collection of information on adolescent behaviors as it is designed to investigate adolescent health and risk behaviors. The "Add Health" project is widely considered to be the largest and most comprehensive survey of adolescents ever undertaken, with a stratified sample of 80 high schools collectively representative of the U.S. school system with respect to region of country, urbanicity, school size, school

[2] Although, Bretteville-Jensen and Jacobi (2008) considers an alternative to a causal link between cannabis and subsequent hard-drug use, offering non-causal explanations for the observed "staircase" pattern.

[3] As an exception to the dearth of evidence on the role of institutions in drug use, although somewhat removed from the focus here, Mehay and Pacula (1999) exploits a drug-testing policy implemented by the military in 1981 and documents that rates of illicit-drug use among military personnel are significantly lower than civilian rates in years after the implementation of the program but not before, which they interpret as a sizable deterrence effect. Some 30 years have past since this policy change was initiated, though, and the nature of the policy change does not necessarily map into us learning about the implications of school policy toward drug use. Exploiting transaction-level data, Pacula, Kilmer, Grossman, and Chaloupka (2007) does find that changes in sanctions that lower the legal risks for users are associated with higher marijuana prices in the short-run. Anderson (2009) also offers some evidence that demand-side interventions to curbing drug use may be ineffective at changing consumption behavior, although this is focussed on methamphetamine use.

[4] In a related consideration, while the emphasis is more broadly on school crime (e.g., violent incidents), Cook, Gottfredson, and Na (2009) argues that crime in school is not a simple sum of students' criminal propensities—"that the organizational characteristics of the school have considerable influence."

[5] Zero tolerance polices are an outgrowth of the Reagan–Bush era drug enforcement policy, and such disciplinary measures are largely seen as attempts to send a message by punishing both major and minor incidents severely. For additional background on the history, philosophy, and effectiveness of zero tolerance school disciplinary strategies, see Skiba (2000).

Table 1
Summary statistics.

	Mean	Std Dev
Used marijuana in last 30 days	0.17	0.37
Grade level	10.49	1.09
Male	0.50	0.50
Black	0.22	0.41
Hispanic/Latino	0.20	0.40
Asian/Pacific	0.09	0.29
Other non-white	0.03	0.17
Parent Educ: high school	0.24	0.43
Parent Educ: some college	0.24	0.43
Parent Educ: college	0.12	0.32
Parent Educ: graduate	0.08	0.27
Weekly religious attendance	0.38	0.48
Unemployment rate, county	0.07	0.02
Proportion urban, county	0.69	0.38
School size: 401–1000	0.26	0.44
School size: 1001–4000	0.66	0.47
School area: urban	0.28	0.45
School area: suburban	0.54	0.50
School governance: private	0.08	0.26
Juvenile arrests per 100k, county	336.98	150.83
Arrests per crime, county	0.21	0.08
Median HH income, county	30,859.9	7851.1
Observations	12,642	

type, and ethnicity. For each of these schools, "feeder" schools were selected on the basis of student contributions to the chosen high school. An in-school questionnaire was administered to students in sampled schools between September 1994 and April 1995, and a random sample was selected from each of these schools for more detailed interviews, conducted in the respondents' homes between April and December 1995. It is this detailed "In-Home Survey" that is adopted in the current analysis. Summary statistics for the sample of students in 9th through 12th grades are shown in Table 1. This constitutes the largest sample of students employed in the analysis.[6]

2.2. Marijuana

While marijuana has been the most popular illicit substance among youths for some 25 years, there has been some variation in usage rates across time. With the first wave of Add Health collected in 1994/5, one should note that 1992 is generally thought of as a low in adolescent marijuana use. Marijuana use again rose after 1992.[7]

In the Add Health survey, the available information about marijuana use derives from responses to the question, "During the past 30 days, how many times did you use marijuana?" Roughly 14 percent of Add Health respondents report consuming marijuana in the 30 days prior to the interview. Given the mass at zero, I will report the results of a discretized version of this continuous response. For completeness, I will consider the intensive margin separately. Since I am relying on self-reported participation in potentially sensitive areas of disclosure, I note that for sensitive topics survey respondents listened to pre-recorded questions through earphones and entered their answers directly on laptops in order to maintain confidentiality and to minimize the potential for interviewer or parental influence. Rates of risky behaviors reported in Add Health are consistent with those measured in other sources (see Mocan & Tekin, 2005, 2006, 2010).

2.3. School-imposed penalties

Add Health records the penalties associated with both the first and second occurrences of student drug use, which will enable the identification strategy adopted below. Specifically, school administrators report the consequence a student faces when he is caught "using an illegal drug at school" for the first time and, separately, caught a second time.

All Add Health schools deal with first-time offenders with either an in-school suspension, an out-of-school suspension, or an expulsion. Possibly given the seriousness of drug use in adolescents, there is a clustering of sorts in how schools penalize drug-related offences. For example, among the largest sample used here, the minimum penalty is an in-school suspension, and only one school imposes such a penalty to first-time offenders, while 45 schools issue out-of-school suspensions and 24 schools issue expulsions. Of the 46 schools that do not expel first-time offenders, 29 schools will expel students upon a second occurrence.

3. Empirics

Point estimates from a simple model of drug use on school-level penalties for drug-related offences will be subject to some interpretive challenges. In particular, to the extent we anticipate that schools respond to higher drug use with more-severe penalties, OLS estimates of this relationship will be biased upward. In Section 3.2, I offer an instrumental-variables strategy through which I retrieve an estimate of the causal role of punishment severity on drug use. In motivating such a specification, I first present simple OLS models of the relationship in Section 3.1, and arrive at the sample of schools that will be used in identifying the causal estimate.

3.1. A baseline specification

Consider a general model of whether individual i has consumed marijuana as a function of the penalty associated with school-related drug offences,

$$Marijuana_{igs} = \gamma_g + \beta_1 Penalty_s^{drugs} + \gamma X_{is} + \epsilon_{igs}, \tag{1}$$

where i is in grade g at school s and $Penalty_s^{drugs}$ is the severity of penalty at school s (i.e., in-school suspension, out-of-school suspension, or expulsion). Unfortunately, there is no ability to separately identify marijuana consumed

[6] Performing a similar analysis on the AddHealth 7th and 8th graders reveals no significant relationship between penalty severity and marijuana use among this group. That said, the instrumental-variable technique does reverse the sign of the point estimate on penalty severity from positive to negative.

[7] Pacula et al. (2000) offer a nice analysis of this trend, linking time-series variation in consumption to changes in perceptions of the harms associated with regular marijuana use.

at school, where we would anticipate the largest treatment effects to arise from school-imposed penalties.

Since penalties do not vary within schools, $\hat{\beta}_1$ will be identified by the variation in *Penalty* that exists across schools. Grade-level fixed effects (γ_g) will be included throughout the analysis, so identification in all cases will be within grade-level, across schools. With no allowance for the inclusion of school-level fixed effects, I will control directly for the observable heterogeneity across schools with school size (i.e., small, medium, large), governance (i.e., public or private), urbanicity (i.e., urban, suburban, rural), and region (i.e., West, Midwest, South, Northeast).[8] I will also include county-level information on juvenile arrests per capita, arrests per crime, median household income, the proportion urban, the proportion rural, and the unemployment rate. At the individual level, included in X_{is} will be indicator variables for gender, race (i.e., black, Asian, Hispanic, other nonwhite), parent education (i.e., less-than high school, high school, some college, bachelor, graduate/professional), and religious participation (i.e., an indicator for weekly attendance). In all specifications I report standard errors that are corrected for clustering at the school level.

As a first pass, I report the estimated coefficients of a linear-probability model of the form (1), allowing level shifts in marijuana use with each first- and second-offence penalty observed in the data.[9] Since schools vary in their penalties for first-time offenders by imposing in-school suspensions, out-of-school suspensions, or expulsions, I include intercept shifters for out-of-school-suspensions and expulsions. Since schools vary in second-offence penalties only between out-of-school suspensions or expulsions, I allow for a difference in marijuana use by whether student *i*'s school expels students for second offences.

Table 2
OLS estimates of the relationship between marijuana use and school-imposed penalties.

Variables	(1)	(2)	(3)
Expulsion on 2nd occurrence	0.010		0.016
	(0.016)		(0.015)
Expulsion on 1st occurrence	−0.025	−0.013	
	(0.039)	(0.032)	
Suspension on 1st occurrence	−0.053	−0.044	
	(0.042)	(0.037)	
Male	0.037***	0.037***	0.037***
	(0.007)	(0.007)	(0.007)
Black	−0.005	−0.004	−0.007
	(0.014)	(0.014)	(0.014)
Hispanic/Latino	0.004	0.004	−0.000
	(0.013)	(0.013)	(0.014)
Asian/Pacific	−0.087***	−0.087***	−0.089***
	(0.015)	(0.015)	(0.016)
Other non-white	0.049	0.050	0.047
	(0.033)	(0.033)	(0.033)
Observations	12,642	12,642	12,642
Mean	0.166	0.166	0.166

The dependent variable is equal to one where the student reports to have consumed marijuana "in the last 30 days." Reported coefficients are least-squares estimates. All specifications also include controls for grade level (i.e., 9 through 12, less the omitted group), region (i.e., West, Midwest, South, Northeast), school size (i.e., small, medium, large), governance (i.e., public or private), and urbanicity (i.e., urban, suburban, rural), county-level juvenile arrests per capita, arrests per crime, median household income, the proportion urban, proportion rural, and unemployment rate, and individual-level indicators for parent education (i.e., less-than high school, high school, some college, bachelor, graduate/professional), and religious participation (i.e., an indicator variable for weekly attendance). Standard errors (in parentheses) are corrected for clustering at the school level.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

In Column (1) of Table 2, I cannot reject that any of these differences are zero—when uncorrected for endogeneity, there is no measurable difference in the reported marijuana use of students associated with their school's disciplinary response. This suggests that schools' drug-related penalties are ineffective in determining adolescent drug use. However, endogenous penalties imply that these estimates are likely biased upward. There are also no significant patterns in drug use revealed when first- and second-offence punishments are entered separately, as reported in columns (2) and (3).

3.2. An instrumental-variables approach

3.2.1. IV setup

In motivating the identification strategy below, one should have in mind an interpretation to the two penalties associated with drug use at a given school. One reasonable interpretation is that the penalty for a second occurrence captures what a school is ultimately prepared to do in response to this behavior and that the first, to the extent that it is less severe, is some measure of grace being afforded to "first-time offenders." If true, one would be particularly reluctant to consider the variation in first-offence punishment as exogenous to student behavior, since this "grace" might well be earned (e.g., in response to less drug use).

In order to retrieve an estimate of the causal effect of penalty severity, I will consider the variation in second-offence penalties in a sample of schools with common first-offence penalties. The obvious payoff from this restriction is in keeping any unobserved heterogeneity that is motivating differences in first-offence penalties from contributing to the estimated effect of school-imposed penalties on drug use—of $\hat{\beta}_1$. In this environment, I then instrument for each school's second-offence penalty with variation that is arguably exogenous, yielding possibly the cleanest environment available for answering the question of interest. This will ultimately serve as the preferred specification.

Given the breakdown of penalties, this amounts to restricting the sample of students to those who attend schools that issue out-of-school suspension to first-time

---

8 Although interesting, many other school-level attributes have insufficient variation to consider including as covariates. For example, greater than 98 percent of schools offer drug awareness and resistance education programs.

9 Results are robust to alternatives to estimating linear probabilities. However, discrete-type IV estimators, which will be required in subsequent specifications, assume that the endogenous regressors are continuous and are not appropriate for use with discrete endogenous regressors. Thus, reporting linear probabilities here allows for better comparison to the subsequent two-stage least squares estimates.

offenders. There are too few schools issuing in-school suspensions to reasonably interpret estimates from separate specifications and there is no comparable specification for students who are at schools that treat first offences with expulsion, as their second-offence consequences are irrelevant.[10] This highlights a tradeoff in the identification strategy—achieving cleaner identification by restricting the sample of schools by their first-offence penalties.

Conditional on being in such a school, then, I instrument for $Expulsion_s^{drugs}$ in a model of i's choice to consume marijuana,

$$Marijuana_{igs} = \gamma_g + \beta_1 Expulsion_s^{drugs} + \gamma X_{is} + \epsilon_{igs}, \tag{2}$$

where again, i is in grade g at school s and $Expulsion_s^{drugs}$ is the measure of penalty severity at school s for students caught using an illegal drug at school for a second time (i.e., an indicator variable for whether the school expels second-time offenders). In all cases, the counterfactual to expulsion remains an out-of-school suspension and, as before, errors are corrected for clustering at the school level.

3.2.2. The instruments

As with drug-related occurrences, the School Administrator Questionnaire in the Add Health survey includes first- and second-occurrence penalties for a variety of other offences. Being careful to avoid employing instruments that themselves may influence drug use and cannot be excluded from the second stage, I instrument for $Expulsion_s^{drugs}$ with the *difference in severity* between the first and second penalties associated with other infractions. For example, with the information on penalties for infractions of type j at school s, a potential instrument Z_s^j can be defined as,

$$Z_s^j = SecondPenalty_s^j - FirstPenalty_s^j. \tag{3}$$

Across j, this amounts to a set of school-specific "punishment trajectories" that are independent of level differences in penalty severity across schools. Below, I discuss the particular choice of infractions j, such that Z_s^j are unlikely to relate to substance use itself.

In order to quantify penalties (and the differences between first and second penalties) I impose a cardinal ranking on the penalties available to institutions. Penalties can range from "verbal warning" to "expulsion," which I simply map onto the range one through five.[11] As a result, the higher is a given Z_s^j the more school s tends to ramp up the severity across first and second offences of type j. For example, if school m imposes an expulsion for a second offence but only an out-of-school suspension for a first offence, then $Z_m^j = 5 - 4 = 1$, which would be equivalent to school n imposing an out-of-school suspension for a second offence and an in-school suspension for a first offence, with $Z_n^j = 4 - 3 = 1$. Quite clearly, Z_s^j is independent of *level* differences in penalty severity across schools.[12] As a robustness check, I consider a quadratic transformation of this cardinal ranking. This transformation yields a a similar point estimate to the IV estimates reported below.[13]

Among the trajectories that are arguably excludable, I adopt two as instruments throughout the analysis—the difference between first- and second-offence punishments for "Stealing school property" and for "Verbally abusing a teacher." The trajectories derived from several other categories of infraction are not considered as possible instruments, as the exclusion restrictions in these cases seem problematic. In particular, those related directly to substance use are unlikely to satisfy the exclusion restriction.[14] It is also questionable whether those associated with potential physical harm are excludable. Since estimates are not qualitatively different if they are not included as instruments, I discard them from the estimating equations.[15] The remaining contender, "Cheating," can likely be excluded from the drug equation but does not survive redundancy tests and is therefore not included in the reported specifications.[16]

Conditional on being in a school that treats first-time drug offenders equivalently, the identifying assumption is that one's marijuana use is not related to how much one's school increases it's penalty between first and second occurrences of stealing school property or verbally abusing a teacher. If one separately tabulates raw differences in observables across schools by the key variable of interest—whether or not the school punishes second occurrences with expulsion—anticipated differences emerge between the two types of punishment regimes, which further highlights that level difference in penalty severity is not clean variation and must be instrumented for in order

[10] As an alternative, one could include fixed effects for each first-offence penalty and instrument for the variation in second-offence penalties. But, given that second-offence consequences are irrelevant for schools that already expel first-time offenders, such an approach would merely add the students from the four schools that issue in-school suspension. The results are not sensitive to their inclusion and are not reported.

[11] The full set of possible penalties is, "(1) verbal warning," "(2) minor action," "(3) in-school suspension," "(4) out-of-school suspension," and "(5) expulsion." In reality, consequences need not span this entire range, however. For example, as suggested already, no school imposes less than an "in-school suspension" for drug offences.

[12] The benefit to identifying off of these trajectories is made all-the-more salient when considering the evidence offered by Babcock (2009), who suggests that high-school graduation and labor participation outcomes appear higher for students who attended schools with stricter discipline policies—notably, schools with higher average punishment *levels* over a range of disciplinary margins. Also identifying off of levels, Barton, Coley, and Welingsky (1998) find that stricter discipline policies in tenth grade to be associated with lower rates of delinquency in 12th grade.

[13] This is meaningful insofar as the cardinal ranking imposed may not reflect non-linearities in moving from one penalty to another, which one might worry are driving the result. For example, a linear ranking of penalties does not reflect that it may be more difficult for an administrator to increase from an out-of-school suspension on a first offence to an expulsion on the second, than to go from an in-school to an out-of-school suspension. In the above example, this is reflected as $Z_m^j = 5^2 - 4^2 = 9$, and $Z_n^j = 4^2 - 3^2 = 7$. The reported IV estimates are robust to this concern.

[14] These include smoking at school, drinking alcohol at school, possessing alcohol, and possessing an illegal drug.

[15] These included fighting with another student, injuring another student, possessing a weapon, and physically injuring a teacher.

[16] Including a cheating trajectory as part of the set of instruments yields slightly higher point estimates on second-offence expulsion. See Breusch, Qian, Schmidt, and Wyhowski (1999) for details on testing for instrument redundancy.

Table 3
Correlation of the instruments with mean school attributes.

Dependent variable	j = "Stealing school property"		j = "Verbally abusing a teacher"		Joint-test
	Z_s^j coefficient	P-value	Z_s^j coefficient	P-value	P-value
Age	0.203	0.243	0.222	0.136	0.187
Proportion male	−0.061	0.190	−0.059	0.139	0.160
Proportion black	0.003	0.963	−0.018	0.753	0.950
Proportion Hispanic/Latino	0.015	0.787	−0.009	0.850	0.944
Proportion Asian/Pacific	−0.033	0.215	0.013	0.584	0.379
Proportion other non-white	0.018	0.225	0.002	0.879	0.480
Parent: proportion high school	0.153	0.005	−0.005	0.918	0.019
Parent: proportion some college	−0.034	0.388	−0.034	0.309	0.437
Parent: proportion college	−0.023	0.252	0.046	0.007	0.010
Parent: proportion graduate	−0.048	0.169	0.015	0.625	0.327
Relig attend: Proportion weekly	−0.018	0.733	0.040	0.388	0.636
Unemployment rate, county	0.000	0.933	−0.005	0.220	0.467
Proportion urban, county	−0.198	0.018	−0.054	0.464	0.053
School size: 401–1000	0.071	0.548	−0.009	0.928	0.829
School size: 1001–4000	−0.072	0.541	0.045	0.662	0.739
School area: Urban	−0.118	0.293	−0.127	0.186	0.263
School area: Suburban	0.105	0.374	0.065	0.523	0.573
School governance: Private	−0.077	0.174	0.010	0.832	0.380
Juvenile arrests per 100k, county	−51.205	0.162	6.232	0.845	0.363
Arrests per crime, county	−0.009	0.541	−0.006	0.667	0.771
Median HH income, county	−2324.8	0.187	−496.6	0.745	0.410

Each coefficient represents a separate specification regressing the attribute on the instrument.

to retrieve estimates of the causal effect on drug use.[17] In Table 3, I report the estimated coefficients of a variety of school attributes regressed on each of the two instruments used below, and P-values associated with the null, $H_0{:}\hat{\beta} = 0$. I also report P-values for joint F tests on the two instruments together predicting each of the attributes. In almost all cases, the trajectory implied by the school's treatment of first and second non-drug offences does not vary significantly with these observable attributes. Moreover, the first-stage F-statistic (reported with the regression results in Table 4) is sufficient to reject the null that the excluded instruments are irrelevant in the first-stage regression (Stock & Yogo, 2005).

3.2.3. *The IV results*

In Column (1) of Table 4, I first report the OLS results from the restricted sample of schools that penalize first-time offenders with out-of-school suspensions. As in Table 2, marijuana use does not appear to respond to whether these schools expel second-time offenders or issue a second out-of-school suspension.[18] Again, however, this likely represents an upwardly biased estimate of the causal influence of expulsions to the extent that schools respond to drug use with stiffer penalties for second-time offenders, even when they share first-offence penalties.

IV estimates of the influence of second-offence expulsion on drug use are produced in Column (2). In short,

Table 4
IV estimates of the deterrent effect of expulsion on marijuana use.

Variables	(1) OLS	(2) IV
Expulsion on 2nd occurrence	0.002	−0.082**
	(0.018)	(0.040)
Male	0.039***	0.038***
	(0.009)	(0.009)
Black	−0.001	0.013
	(0.017)	(0.018)
Hispanic/Latino	0.009	0.015
	(0.018)	(0.017)
Asian/Pacific	−0.090***	−0.087***
	(0.017)	(0.013)
Other non-white	0.065	0.069
	(0.043)	(0.044)
Observations	9376	9376
1st-stage F	n/a	11.67
Mean	0.170	0.170
Impact (%)	1.432	−47.89
Effect size	0.006	−0.217

The dependent variable is equal to one where the student reports to have consumed marijuana "in the last 30 days." Reported coefficients are least-squares estimates, from a sample of schools with common first-offence penalties. All specifications also include controls for grade level (i.e., 9 through 12, less the omitted group), region (i.e., West, Midwest, South, Northeast), school size (i.e., small, medium, large), governance (i.e., public or private), and urbanicity (i.e., urban, suburban, rural), county-level juvenile arrests per capita, arrests per crime, median household income, the proportion urban, proportion rural, and unemployment rate, and individual-level indicators for parent education (i.e., less-than high school, high school, some college, bachelor, graduate/professional), and religious participation (i.e., an indicator variable for weekly attendance). Standard errors (in parentheses) are corrected for clustering at the school level.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

[17] The pattern is suggestive of heavier potential penalties tending to fall more on black students than on other non-white students, in higher-unemployment and lower-income areas, but also on students of more-educated parents with higher religious attendance.

[18] In unreported results, restricting the sample to the four schools that issue in-school suspension to first-time offenders suggests that the influence of "Expulsion on 2nd occurrence" is a larger positive but, unlike other models, estimates for this sample of schools are imprecise and sensitive to specification.

Table 5
IV estimates of the deterrent effect, by gender.

Variables	(1) Male	(2) Female
Expulsion on 2nd occurrence	−0.056	−0.107***
	(0.050)	(0.035)
Observations	4707	4669
1st-stage F	12.03	11.00
Impact (%)	−29.00	−71.91
Effect size	−0.141	−0.300
Mean	0.192	0.148

The dependent variable is equal to one where the student reports to have consumed marijuana "in the last 30 days." Reported coefficients are least-squares estimates. All specifications are equivalent to that in Table 4 (2). Standard errors (in parentheses) are corrected for clustering at the school level.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Table 6
IV estimates of the deterrent effect of expulsion on the intensive margin.

Variables	(1) Pooled	(2) Male	(3) Female
Expulsion on 2nd occurrence	−5.652	−8.940	−0.469
	(4.879)	(6.846)	(3.376)
Observations	1597	905	692
1st-stage F	10.73	13.35	7.866
Impact (%)	−43.04	−53.22	−5.622
Effect size	−0.140	−0.172	−0.033
Mean	13.13	16.80	8.342

The dependent variable is equal to the number of times the student reports to have consumed marijuana "in the last 30 days," conditional on consumption greater than zero. Reported coefficients are least-squares estimates. All specifications are equivalent to that in Table 4 (2). Standard errors (in parentheses) are corrected for clustering at the school level.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

correcting for the endogeneity of *Expulsion* reveals a very different relationship between school penalties and adolescent marijuana use. Within grade level and region, students attending schools that are equivalent with respect to their treatment of first-time offenders but that penalize second occurrences with expulsion (instead of suspension) are significantly less likely to report that they consume marijuana. The estimated difference is also reasonably large, suggesting a 0.082 decrease in the probability that one has consumed marijuana in the thirty days prior to the survey where schools expel on second occurrences. At the mean usage of 0.17 this implies a fairly high impact, of roughly 48 percent. On the other hand, the implied effect size from moving from out-of-school suspensions to expulsions is quite reasonable, reducing the proportion of students consuming marijuana by roughly 0.22 standard deviations.[19]

In Table 5, I separately identify the treatment effect by gender, which reveals that the influence of penalty severity on the propensity to consume marijuana is roughly twice as high among female students. The point estimate among females is both larger in magnitude and more-precisely estimated, suggesting that the takeaway is somewhat more nuanced. When separately identified, the preferred identification strategy continues to retrieve a negative estimate among male students (i.e., a sign change compared to the original OLS estimates that do not control for the endogeneity of penalty severity), but the precision of the estimate suggests that male responsiveness is no longer significant at conventional levels.

3.2.4. Intensive margin

In Table 6, I repeat the above analysis on the intensive margin, defining *Marijuana* as the number of times marijuana was used in the month prior to being interviewed and restricting the sample to include only those for which *Marijuana* > 0.[20] Similar patterns emerge in the IV estimate of *Expulsion*, with usage falling some 0.5 time (monthly) over a mean of 8.3 times. Again, the implied effect size is quite reasonable—roughly 0.03 standard deviations. Even though this sample is restricted to students who report using marijuana within the last 30 days, these estimates are close to those found around the extensive margin of use, reported in Table 4. That said, the IV estimates in the pooled sample, and in the gender-specific samples, are imprecise and it would be reasonable to conclude that the margin of importance is the extensive margin. With that caveat, the pointe estimates do suggest that the decline in drug consumption on the intensive margin is an order of magnitude higher among male students.

4. Discussion

4.1. Falsification

One may fear that expulsion regimes are merely identifying a "type" of student, as reflected in their marijuana use, but not an actual difference in drug behavior in response to penalty severity. In Table 7, I consider a potential falsification of the main result by running a similar specification but replacing drug use with a measure of the student's ability, which should not directly respond to drug-related expulsion.

In Column (1) I consider the implications of penalty severity on the student's score on a variant of the Peabody Picture Vocabulary Test (PPVT) that is administered to all survey respondents.[21] I find no systematic variation in PPVT with the severity of school-imposed penalties for drug use. In columns (2) and (3), I similarly find no relationship between severity and student ability when the models are run separately by gender. That variation in

[19] Similar estimates result from a quadratic transformation of the penalties used in the first-stage equation. Results are also qualitatively robust to a limited information maximum likelihood estimator (LIML), which can be more robust to the presence of weak instruments. That said, the reported specifications yield first-stage F statistics that far exceed the weak ID test critical values of Stock and Yogo (2005).

[20] While not reported, the anticipated bias correction is also apparent when the IV estimates in Fig. 6 are compared to the OLS estimates.

[21] The primary advantage the PPVT has over grade-based performance is that the test scores are comparable across schools and grade levels in a way that grade-base performance measures may not be.

Table 7
Falsification exercise: expulsion should not also predict student test scores.

Variables	(1) Pooled	(2) Male	(3) Female
Expulsion on 2nd occurrence	6.594	6.050	7.419
	(5.826)	(5.023)	(6.862)
Observations	8950	4458	4492
1st-stage F	10.73	13.35	7.866
Mean	48.86	50.72	47.00

The dependent variable is the student's score on the PPVT. Reported coefficients are least-squares estimates. All specifications are equivalent to that in Table 4 (2). Standard errors (in parentheses) are corrected for clustering at the school level.
$^{**}p < 0.1$.
$^{**}p < 0.05$.
$^{***}p < 0.01$.

student ability is independent around the treatment variable yields additional support to a causal interpretation.

4.2. Reported drug use

While tempting to consider the point estimate as the underlying true causal effect of penalty severity on drug use, one important caveat remains. Notably, one must bear in mind that the above analysis points to a causal response of *reported* drug use to a school-imposed deterrent. Thus, point estimates may still reflect both actual reduced usage and reductions in one's proclivity to report actual use. This is particularly important in this context as both usage and reporting may respond negatively to increases in penalty severity.[22]

Furthermore, recall that the identifying variation is coming from variation in penalty severity specifically around second-time offences. However, with no ability to separately identify students by whether they have been made subject to the prescribed penalty, that we are limited to self-reported use implies that one should interpret the point estimate retrieved by the model in such a way as to allow for the potential that there is lower overall use (i.e., inclusive of all students, whether they have been caught or not) in schools that penalize second-time offenders more severely. In particular, this is to say that we are not identifying the effect of the existing policy variation only on second-time offenders.

5. Conclusion

The focus of this analysis is on the potential for school policy to influence a student's consumption of marijuana. I model students' marijuana usage as a function of the penalties that would be imposed by the students' schools on those caught consuming an illegal drug. Given the potential endogeneity of these penalties, I adopt an instrumental-variables strategy to retrieve an estimate of the causal influence of expulsion on consumption.

In the end, estimates imply that increasing penalties for second-time offenders from out-of-school suspensions to expulsions reduces the proportion of students reporting 30-day marijuana consumption by roughly 0.22 standard deviations, or a 48 percent decrease from the mean propensity to consume of 0.15. I thereby demonstrate the efficacy of school-imposed penalties as a deterrent to adolescent drug use—the first evidence of such efficacy. Given what the literature has documented regarding the consequences of drug use—especially in school-aged youth—this research suggests that school sanctions may have important long-run benefits. However, effects sizes are larger and more-precisely estimated among female students, suggesting that gains to policy innovations may be gender specific.

[22] That said, care was taken to elicit truthful responses to potentially sensitive areas of disclosure. For example, the drug-use measure used in the current analysis were collected via audio-enhanced, computer-assisted self-interviewing protocols (Audio-CASI). Respondents answer the questions themselves, rather than telling the interviewer their answers. "When you get to the first question, the computer will read the question to you so that you can hear it through these headphones. It will also tell you what to do to enter your answer. We have made it very simple for you to use the computer. Let's take a look at how it works, by completing a couple of practice questions."

References

Anderson, D. M. (2009). *Does information matter? The effect of the Meth Project on Meth use among youths*. Working Paper University of Washington.

Babcock, P. (2009). The rational adolescent: Discipline policies, lawsuits, and skill acquisition. *Economics of Education Review*, *28*, 551–560.

Barton, P., Coley, R., & Welingsky, H. (1998). *Order in the classroom: Violence, discipline and student achievement*. Princeton, NJ: Policy Information Center, Educational Testing Service.

Bretteville-Jensen, A. L., & Jacobi, L. (2008). *Climbing the drug staircase: A Bayesian analysis of the initiation of hard drug use*. Discussion Paper 3879 IZA Bonn.

Breusch, T., Qian, H., Schmidt, P., & Wyhowski, D. (1999). Redundancy of moment conditions. *Journal of Econometrics*, *91*, 89–111.

Chatterji, P. (2006). Illicit drug use and educational attainment. *Health Economics*, *15*, 489–511.

Cook, P. J., Gottfredson, D. C., & Na, C. (2009). *School crime control and prevention*. Technical Report Duke University.

DeSimone, J. (2002). Illegal drug use and employment. *Journal of Labor Economics*, *20*, 952–977.

Heaviside, S., Rowand, C., Williams, C., & Farris, E. (1998). *Violence and discipline problems in U.S. Public Schools: 1996–97*. Technical Report Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Kaestner, R. (1995). *The effects of cocaine and marijuana use on marriage and marital stability*. NBER Working Papers 5038 National Bureau of Economic Research.

Markowitz, S. (2000). *The role of alcohol and drug consumption in determining physical fights and weapon carrying by teenagers*. NBER Working Papers 7500 National Bureau of Economic Research.

Markowitz, S., Chatterji, P., Kaestner, R., & Dave, D. (2002). *Substance use and suicidal behaviors among young adults*. NBER Working Papers 8810 National Bureau of Economic Research.

Mehay, S. L., & Pacula, R. L. (1999). *The effectiveness of workplace drug prevention policies: Does 'zero tolerance' work?* NBER Working Papers 7383 National Bureau of Economic Research.

Mocan, H. N., & Tekin, E. (2005). Drug use and juvenile crime: Evidence from a panel of siblings and twins. In B. Lindgren, & M. Grossman (Eds.), *Economics of substance use: Individual behavior, social interactions, markets, and politics* (pp. 91–120). Amsterdam: Elsevier.

Mocan, H. N., & Tekin, E. (2006). Guns and juvenile crime. *Journal of Law and Economics*, *49*, 507–531.

Mocan, H. N., & Tekin, E. (2010). Ugly criminals. *Review of Economics and Statistics*, *92*, 15–30.

van Ours, J. C. (2006). Dynamics in the use of drugs. *Health Economics*, *15*, 1283–1294.

Pacula, R. L., Grossman, M., Chaloupka, F. J., O'Malley, P. M., Johnston, L., & Farrelly, M. C. (2000). *Marijuana and youth*. NBER Working Papers 7703 National Bureau of Economic Research.

Pacula, R. L., Kilmer, B., Grossman, M., & Chaloupka, F. J. (2007). *Risks and prices: The role of user sanctions in marijuana markets*. NBER Working Papers 13415 National Bureau of Economic Research.

Ringel, J. S., Ellickson, P. L., & Collins, R. L. (2006). The relationship between high school marijuana use and annual earnings among young adult males. *Contemporary Economic Policy*, *24*, 52–63.

Skiba, R. J. (2000). *Zero tolerance, zero evidence: An analysis of school disciplinary practice*. Policy Research Report SRS2 Indiana Education Policy Center.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*. Cambridge University Press.

Yamada, T., Kendix, M., & Yamada, T. (1998). The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics*, *5*, 77–92.