What America's Users Spend on Illegal Drugs: 2000-2010

Technical Report

February 2014

What America's Users Spend on Illegal Drugs: 2000-2010
Technical Report

Prepared for:
Office of National Drug Control Policy
Office of Research and Data Analysis
Under HHS contract number: HHSP23320095649WC
Contract Officer Representative: Michael Cala, PhD

Prepared by:
RAND Corporation
1776 Main Street
Santa Monica, CA 90401-3297
B. Kilmer, S. Everingham, J. Caulkins, G. Midgette,
R. Pacula, P. Reuter, R. Burns, B. Han, R. Lundberg

Table of Contents

List of Tables and Figures

Section 1. Introduction

In January 2012, the White House Office of National Drug Control Policy asked RAND to generate estimates of the total number of users, total expenditures, and total consumption for four illicit drugs from 2000 to 2010: cocaine (including crack), heroin, marijuana, and methamphetamine (or meth). The Main Report provides an overview of our methodology and presents our results.

This Technical Report presents additional details about our methods and calculations. Similar to the Main Report, it initially focuses on the analyses for cocaine, heroin, and meth. Section 2 provides a detailed account of the regression models which helped us understand which state- and substate-level variables predict the number of adult male arrest events involving a positive drug test for cocaine, heroin/opiates, and meth. Section 3 explains how we used the information from these models to generate estimates of the total number of chronic cocaine, heroin, and meth users throughout the country. Sections 4 and 5 then explain how we converted these figures regarding chronic drug users (CDUs) into expenditure and consumption estimates. All of these sections explain how the methods differ from what was used in the previous version of the report.

Sections 6 and 7 exclusively focus on marijuana. The methods employed to estimate users, expenditures, and consumption are less complex than what was required for the harder drugs. But since these estimates are rooted in the National Survey on Drug Use and Health (NSDUH), which undercounts users, Section 6 considers multiple approaches to inflating these estimates. The Main Report is the first in the *What America's Users Spend on Illicit Drugs* series addressing changes in marijuana potency. Section 7 assesses changes in marijuana potency over the decade and how it influences the expenditure estimates.

Section 8 describes the contents of the spread sheets that generate the results, figures, and tables described in the Main Report and Technical Report.

Section 2. Predicting Positive Drug Tests in ADAM

2.1. Predicting Positive Drug Tests Among Adult Male Arrest Events in ADAM Jurisdictions

Drug use among adult male arrestees is a function of both individual and market-level demand factors, and the relationships can differ by substance. For example, Dave (2008) finds that race, ethnicity, age, the drug possession arrest rate, and STRIDE-derived street-level cocaine prices are statistically associated with the probability that an arrestee in the ADAM program tests positive for cocaine; for heroin, age did not matter, but the street-level price of heroin, possession arrest rate, race, and employment status did. Our challenge was to find state- and substate-level variables that both predicted the share of arrestees testing positive for each drug in ADAM counties and were available for all, or substantially all, counties in the country. These variables are discussed later, but we begin with our construction of the dependent variable.

Dependent Variable

As noted by previous studies (Abt, 2001; ONCDP, 2012c), a large proportion of chronic cocaine, heroin, and methamphetamine users are subject to a nontrivial arrest risk; hence, they are within the purview of ADAM data. There were roughly 22,000 adult male respondents in the 43 ADAM I counties (2000–2003)[1] and 4,000 in the 10 ADAM II counties (2007–2010) who submitted to a toxicology screen for a battery of substances (Table 2.1).

[1] Not all 43 jurisdictions participated every year. The average number of participating counties was 36.

Table 2.1: ADAM Locations by Regions

1. Northeast	2. East North Central	3. West North Central	4. South Atlantic	5. South Central	6. Mountain	7. Pacific
Manhattan Philadelphia Albany Boston	*Indianapolis* Detroit *Chicago* Cleveland	Omaha Kansas City *Minneapolis* Des Moines Woodbury	*Washington, D.C.* Ft. Lauderdale Miami *Atlanta* *Charlotte* Tampa	Houston New Orleans Dallas Birmingham San Antonio Laredo Oklahoma City Tulsa	Phoenix *Denver* Albuquerque Tucson Las Vegas Salt Lake City Rio Arriba	*Portland* San Diego Los Angeles San Jose *Sacramento* Anchorage Seattle Spokane Honolulu

NOTE: Italicized cities included in both ADAM I and ADAM II

For each county and year for which data were collected, we collapsed urinalysis results to a county-year mean positive rate for cocaine, methamphetamine, and opiates. We used a linear regression specification, so it was possible to estimate infeasible prevalence rates that were either negative or greater than 1. To eliminate that risk, we employed a logistic transformation on county-year prevalence rates before estimation so as to guarantee predicted prevalence rates fall within the range of 0 and 1:

$$Logistic[R]_{at} = Log(R_t/(1-R_t)), \qquad (2.1)$$

where R is the observed prevalence rate for each ADAM county a in year t.

Independent Variables

The goal of the independent variables is to predict what value would have been measured for the share of adult male arrest events involving a positive drug test had ADAM been conducted in every county. For example, we would expect the proportion of arrestees testing positive to be higher in counties with greater overall prevalence of use, greater demand for treatment, more job applicants testing positive, and more overdose events. So, in non-ADAM counties with high rates of those predictors, our model would predict that a high proportion of arrestees would test positive.

To capture the overall prevalence of use in the area, we included the NSDUH household survey–reported past-year use rates for each drug as predictors. All subnational NSDUH data are pooled in three-year groups to achieve a baseline sample size, and only the cocaine series is available as substate use estimates. For cocaine, we assign the appropriate substate regional estimate to each county in that region and divide by population to generate a rate. The data for methamphetamine and heroin were reported for metro and non-metro users at the state level (made available by a special request to the Substance Abuse and Mental Health Services Administration [SAMHSA]). However, as non-metro area estimates for these drugs are very small and imprecise, we assume constant NSDUH use rates within states and generate a rate series by dividing by total state population.

We used treatment admissions recorded in Treatment Episode Daily Systems (TEDS) as our measure of treatment demand.[2] TEDS data include information on the substance or substances leading to admissions to treatment facilities receiving public funding.[3] For every location, we calculated the ratio of admissions with a primary diagnosis, including the drug of interest to total admissions. While not universal in coverage, TEDS captures treatment facilities in all states and many counties in the United States, so the data offer a good view of treatment episodes tied to drug dependence and abuse. The TEDS data were acquired at the primary metropolitan statistical area (PMSA) level for metropolitan areas and state levels excluding PMSAs for the rest of the nation. All ADAM sites except Rio Arriba fall into a PMSA. Counties fall exclusively within a single PMSA or non-urban area, so their TEDS admission rates are simply the rate corresponding to their PMSA or state non-urban area. Admission rates may vary according to local publicly funded treatment facility capacity or cost, or where private drug treatment facilities are more common. To account for this disparity in the model, we tested interactions with the region indicators and retained these interaction terms when significant. Additionally, we included National Survey of Substance Abuse Treatment Services (N-SSATS) methadone treatment reports per capita in the heroin model.

Our third predictor was toxicology screen data from Quest Diagnostics. Quest data offer compelling advantages, including being available down to the 3-digit ZIP code level and being based on an objective assay. The data were manipulated using ArcGIS, a mapping and spatial analysis software package, to apportion the test information to each county based on land area. When a county overlaps multiple ZIP

[2] Individuals enter the data for each time they are admitted to a covered treatment facility in a year. Data on individuals, rather than episodes, are not available at the national level. This is only problematic if the rate of same-year repeat-admissions differs systematically across geographies. For example, if two counties have an identical number of people in need of treatment, but due to admissions rules or availability constraints individuals in one county average twice as many treatment admissions per person than those in another, the former's reported treatment rate in TEDS will be twice as large for the latter. However, we expect that such a situation is unlikely.

[3] From the Office of Applied Studies: "TEDS does not include all admissions to substance abuse treatment. It includes admissions at facilities that are licensed or certified by the State substance abuse agency to provide substance abuse treatment (or are administratively tracked for other reasons). In general, facilities reporting TEDS data are those that receive State alcohol and/or drug agency funds (including Federal Block Grant funds) for the provision of alcohol and/or drug treatment services" (SAMHSA, undated).

codes, its rate is calculated as the area-weighted average of the ZIP code areas it is a member of. These data cover all tests between 2000 and 2010 for cocaine, methamphetamine, and opiates.[4] The opiates data were excluded from our analysis as we hope to capture only heroin in our model, rather than other substances including prescription opioids.

After careful consideration, we decided not to include Quest meth testing data in our analyses for multiple reasons. First, the series is not available for two key ADAM years (2000 and 2001). The overall rate of meth use reported in the Quest data accelerates rapidly from 0.19 percent in 2002 to 0.32 percent in 2003, and it would be difficult to extrapolate a trend back to 2000 with much confidence. Second, the number of tests administered grew rapidly, from roughly 40,000 per month in 2002 to nearly 80,000 at the peak in 2007, then rapidly fell to 40,000 once again by late 2008. This suggests that the groups captured in testing may not be comparable over time. While problematic, these wrinkles alone are not sufficient to justify dropping the Quest meth series. However, given the possibility of this changing base and very low overall prevalence in more recent years (0.10 percent between 2008 and 2010), even small changes in the number of positive tests observed in any area result in relatively large changes in prevalence rates entering our models. At this point, we cannot reliably account for the variability in these unobserved factors affecting the data.

For each county, we include the area-weighted average positive rate for each drug in our model of cocaine prevalence in ADAM. However, these data represent only the sample of people submitting to drug tests evaluated by Quest for occupational and medical testing; the majority of tests in this sample are pre-employment tests ordered by people's prospective place of employment. As a result, there may be variation in Quest positive rates due to the composition of people testing in a given location if the types of jobs requiring testing and the people applying for those jobs are different from region to region. Again, we examine and attempt to control for this variability explicitly using region interactions.

Perhaps the most familiar data series recording adverse drug-related health events is from the Drug Abuse Warning Network (DAWN), but this is only available in selected cities (and then weighted to generate national estimates). Mortality data, by contrast, are available throughout the United States, and are based on totals, not samples. We created time series of drug-related deaths per 10,000 county residents for cocaine and heroin from accidental poisoning occurrences in the Multiple Cause-of-Death Mortality Data from the National Vital Statistics System of the Centers for Disease Control (CDC)

[4] Quest data for methamphetamine begins in 2002.

National Center for Health Statistics. We were not able to create an equivalent series for methamphetamine, as the dataset does not differentiate between different types of psychostimulants. However, psychostimulants overdose information may inform heroin and cocaine prevalence predictions, so we included the series as a potential covariate in those models.

In addition, a base set of standard demographic variables were included in all models. The county-level share of total population between 18 and 24 from the annual intercensal population estimates, along with each county's poverty rate, the share of the population who are high school graduates, and log-transformed population from the Census Bureau captured basic demographic variation and trends.

Finally, for series missing values in a given county-year, we impute using linear trends when possible. If data from 2000 or 2010 are missing, we impute using the closest nonmissing annual value for that county.

2.2. Empirical Specification

Our main empirical specification is:

$$Logistic[R_{dat}] = \beta C_{at} + \gamma D_{dat} + \alpha I_A + \theta(D_{at} I_A) + \{\lambda_t\} + \varepsilon_{at}, \qquad (2.2)$$

where $Logistic[R_{dat}]$ is the logistic transformation of the observed positive drug test rate for drug d in county a in year t. C_{at} represents a matrix of demographic characteristics including log(population), population share aged 18–24, population share with high school or more education and poverty rate for county a in time t. D_{dat} represents a matrix of drug market indicators including treatment admissions, Quest Diagnostics-observed positive drug test rates (cocaine), CDC's accidental poisoning rate per 10,000 people (cocaine, heroin), and NSDUH-reported past-year use rates for drug d in county a in time t. I represents six region indicator variables, where each county a belongs to exactly one region A and A contains at least one ADAM county. λ_t represents a vector of time effects tested in some specifications of the models either as year fixed-effects, an ADAM II indicator, or linear, squared, and cubed terms for years. β, γ, α, θ represent vectors of coefficients that are estimated in the model (the latter for region-market demand indictor interaction effects). ε_{at} are the residual error terms for each observation.

We estimate these models using STATA 12, assuming errors are clustered at the county level to account for heteroskedasticity. We weighted each county-year based on the reciprocal of the variance in observed prevalence rate for each drug, so larger, less varied city-years were weighted relatively more than rates based on smaller, more varied observations. For the few counties with no variance in their rates—all arrestees tested negative for a drug—weights were defined as the reciprocal of their sample size. These observations were all based on less than 25 arrestees; they should not contribute as much to their region's predicted prevalence as another county with significantly more observations.

It is important to note that we did not use the sampling weights provided with the ADAM dataset for several reasons. The weights initially developed for ADAM I observations became obsolete when the methodology changed for the 2007 survey. Corrected weights were developed after ADAM II was re-initiated, but the new weights are only available for a subset of the original ADAM I locations. Specifically, the new weights are only available for a subset of observations in nine of the 43 counties that appear in ADAM I, reducing our total sample size from 97,641 to 9,226 and significantly diminishing regional representation.[5] Additionally, the sample-wide aggregate figures using the new weights produced what we believe to be reasonable estimates, but suggested erratic trends in population estimates within counties over time (see Section 3.2 for further discussion of ADAM weights). We chose to base our analyses on the unweighted observations to improve the sample size and regional representation rather than in-county representativeness.

2.3. Model Selection

We selected models using an iterative process, attempting to simultaneously minimize in-sample prediction error and out-of-sample standard error in extrapolating from ADAM counties to the nation. Region-fixed effects and the demographic information series were included in every run. The drug-related variables from NSDUH, TEDS, Quest, and CDC were interacted with the region-fixed effects. These interaction terms were included along with the other covariates for each of the models. Starting with a saturated model of all demographic, drug use, and region-fixed effects variables, and drug use by region interaction terms, we chose a final specification for each drug model empirically. We selected models based on whether any coefficients in the group were significant (at the 10-percent level) in the saturated model and refined each specification to maximize model performance given the risk of overfitting due to the inclusion of too many covariates. The performance of the models was measured using the Akaike Information Criterion (AIC; which measures in-sample performance well in these

[5] Corrected weights were not computed for Washington, D.C., which provides data for only 2003 in ADAM I and 2007–2010 in ADAM II.

models) and the Bayesian Information Criterion (BIC; which measures out-of-sample performance, including extrapolation to the non-ADAM counties).

The saturated models included the substate- and state-level variables and their interactions with region for all three drugs in each model. Some of these indicators dropped out for cocaine prevalence, and all but one dropped out for the other two drugs. If all interaction terms for a market demand indicator variable were not significant at the 10 percent level, we dropped that group of interaction terms from the model. If the base covariate is also insignificant in the resulting parsimonious model, it is also dropped. For cocaine, several interactions came up significant, though not all did. For heroin, the TEDS heroin series were significant.

We went to great effort to make sure time was specified correctly in these models. First we considered year-fixed effects, which offer the greatest amount of flexibility in the model by not imposing any shape restrictions in the regression. However, these fixed effects were rarely independently statistically significant and required ad hoc assumptions about the influence of time on prevalence rates during the three-year period from 2004 to 2006 not covered by ADAM when performing the extrapolation.

We also considered including a linear trend for time, along with time-squared and cubed terms in the models. While linear time was significant in several versions of the models, indicating a systematic shift in prevalence rates, the higher-order terms were insignificant in the cocaine and heroin models, only contributing to error due to "overfitting." However, in the methamphetamine model, they were significant. We assume that this is due to relatively weak drug market indicators for methamphetamine use, leaving more prediction variance to be explained by time-varying factors not captured by the market demand indicators than for cocaine and heroin.

We found very similar results for a model not controlling for time explicitly versus inclusion of an indicator for ADAM II (i.e., an indicator for years 2007–2010), in the cocaine and heroin models. Using the latter ADAM II indicator variable approach, we assumed the effect size was one-half of what was estimated in the model for the non-ADAM years. As this variable was typically not statistically significant and its effect was small, we reduced standard errors somewhat by dropping this ADAM period time variable while not significantly altering predictions with our arbitrary assumption for 2004–2006. In summary, we chose not to explicitly control for time in the cocaine and heroin models, and we controlled for time using a linear time trend, squared, and cubed terms for year in the methamphetamine model.

Ultimately, the best performing model specifications for each drug were as follows:

Cocaine:

Logistic[Cocaine Prevalence Rate$_{at}$] = β_1(log(population$_{at}$)) + β_2(poverty rate$_{at}$) + β_3(HS or more$_{at}$) + β_4 (% 18-24$_{at}$) + γ_1(NSDUH past year cocaine use rate$_{at}$) + γ_2(Quest cocaine rate$_{at}$) + γ_3(TEDS cocaine rate$_{at}$) + γ_4(TEDS meth rate$_{at}$) + γ_5(TEDS heroin rate$_{at}$) + γ_6(CDC cocaine OD rate$_{at}$) + γ_7(CDC psychostimulants OD rate$_{at}$) + θ_{1A}(Quest cocaine rate$_{at}$ × I_A) + θ_{2A}(TEDS cocaine rate$_{at}$ × I_A) + θ_{3A}(TEDS meth rate$_{at}$ × I_A) + θ_{4A}(CDC cocaine OD rate$_{at}$ × I_A) + αI_A+ ε_{at}

Heroin:

Logistic[Opiates Prevalence Rate$_{at}$] = β_1(log(population$_{at}$)) + β_2(poverty rate$_{at}$) + β_3(HS or more$_{at}$) + β_4 (% 18-24$_{at}$) + γ_1(Quest cocaine rate$_{at}$) + γ_2(TEDS cocaine rate$_{at}$) + γ_3(TEDS meth rate$_{at}$) + γ_4(TEDS heroin rate$_{at}$) + γ_5(CDC cocaine OD rate$_{at}$) + γ_6(CDC psychostimulants OD rate$_{at}$) + γ_7(CDC heroin OD rate$_{at}$) + γ_8(N-SSATS methadone rate$_{at}$) + θ_{1A}(TEDS heroin rate$_{at}$ × I_A) + αI_A+ ε_{at}

Methamphetamine:

Logistic[Methamphetamine Prevalence Rate$_{at}$] = β_1(log(population$_{at}$)) + β_2(poverty rate$_{at}$) + β_3(HS or more$_{at}$) + β_4 (% 18-24$_{at}$) + γ_1(NSDUH past year meth use rate$_{at}$) + γ_2(Quest cocaine rate$_{at}$) + γ_3(TEDS cocaine rate$_{at}$) + γ_4(TEDS meth rate$_{at}$) + γ_5(TEDS heroin rate$_{at}$) + γ_6(CDC cocaine OD rate$_{at}$) + αI_A + λ_1(year$_{at}$) + λ_2(year$_{at}^2$) + λ_3(year$_{at}^3$) + ε_{at}

Table 2.2: Regression Estimates for Adult Male Drug User Arrest Events Using ADAM

Covariate	Cocaine	Heroin/ Opiates	Meth
log(Population)	0.238[c]	-0.186[b]	0.243[a]
	-4.24	-2.73	-1.98
Poverty rate	6.088[c]	0.595	-3.642
	-3.82	-0.28	-1.23
High school or more rate	3.716[c]	-1.052	-0.409
	-4.1	-0.77	-0.16
Percent Population 18-24	-1.436	5.707	27.38[c]
	-0.54	-1.32	-4.31
NSDUH past-year cocaine rate	-3.666		
	-1.25		
NSDUH past-year meth rate			359.0[c]
			-7.06
Quest cocaine rate	105.9[c]	-15.51	43.62
	-3.53	-0.87	-1.79
TEDS cocaine rate	-1.249	-1.622[c]	-3.495[c]
	-0.45	-4.48	-6.36
TEDS meth rate	-79.1	-1.272[c]	3.942[c]
	-1.69	-3.64	-4.25

Covariate	Cocaine	Heroin/ Opiates	Meth
TEDS heroin rate	-0.478	2.901	0.0908
	-0.32	-1.9	-0.12
CDC cocaine overdose share	0.736	-0.321	-0.685 [a]
	-0.8	-1.3	-2.14
CDC meth overdose share	-0.859	0.307	
	-1.76	-0.39	
CDC heroin overdose share		-0.0259	
		-0.06	
N-SSATS methadone rate		2.256[c]	
		-6.64	
Quest cocaine × region 2	-149.9[b]		
	-3.1		
Quest cocaine × region 3	87.14		
	-1.83		
Quest cocaine × region 4	-24.85		
	-0.61		
Quest cocaine × region 5	-46.29		
	-1.3		
Quest cocaine × region 6	-23.93		
	-0.63		
Quest cocaine × region 7	15.64		
	-0.28		
TEDS cocaine × region 2	4.537		
	-1.45		
TEDS cocaine × region 3	7.024[a]		
	-2.1		
TEDS cocaine × region 4	2.284		
	-0.81		
TEDS cocaine × region 5	1.38		
	-0.49		
TEDS cocaine × region 6	2.613		
	-0.91		
TEDS cocaine × region 7	0.8		
	-0.27		
TEDS meth × region 2	73.86		
	-1.57		
TEDS meth × region 3	82.43		
	-1.77		
TEDS meth × region 4	82.95		
	-1.78		
TEDS meth × region 5	78.5		
	-1.68		
TEDS meth × region 6	78.57		
	-1.68		
TEDS meth × region 7	77.87		
	-1.67		
TEDS heroin × region 2	-3.268	-1.017	
	-1.93	-0.63	
TEDS heroin × region 3	13.14[a]	4.571	
	-2.03	-1.43	

Covariate	Cocaine	Heroin/ Opiates	Meth
TEDS heroin × region 4	-1.863	5.863[b]	
	-0.84	-2.7	
TEDS heroin × region 5	-0.401	-1.171	
	-0.25	-0.67	
TEDS heroin × region 6	-1.987	2.702	
	-1.09	-1.12	
TEDS heroin × region 7	0.294	-2.275	
	-0.19	-1.48	
CDC OD cocaine × region 2	1.329		
	-1.09		
CDC OD cocaine × region 3	-2.276		
	-1.2		
CDC OD cocaine × region 4	-1.483		
	-1.41		
CDC OD cocaine × region 5	-1.951		
	-1.61		
CDC OD cocaine × region 6	-0.452		
	-0.48		
CDC OD cocaine × region 7	-0.0873		
	-0.08		
Region 2	-0.898	0.691	-0.798[c]
	-0.47	-1.27	-3.77
Region 3	-4.449	-0.101	-0.878[a]
	-1.97	-0.2	-2.46
Region 4	-0.731	-0.583	-0.349
	-0.4	-1.11	-1.28
Region 5	-0.337	0.535	-0.619[a]
	-0.18	-1.01	-2.24
Region 6	-0.682	0.286	-1.480[b]
	-0.37	-0.48	-2.73
Region 7	-0.709	0.887	-1.377[c]
	-0.39	-1.78	-3.59
Year			0.551[c]
			-3.63
$Year^2$			-0.0397
			-1.23
$Year^3$			-0.0000822
			-0.04
Constant	-7.522[c]	-0.368	-10.89[b]
	-3.41	-0.19	-3.01
N	183	183	183
R-sq	0.899	0.733	0.822
AIC	45.08	219.1	315.2
BIC	199.1	299.2	379.4
Covariates	47	24	19

t statistics below coefficient estimates

[a] $p<0.05$

[b] $p<0.01$

[c] $p<0.001$

Since several of the market demand indicators were highly correlated, their interaction terms often inflated BIC while minimizing AIC.[6] As BIC increased, so too did our confidence intervals, while the final estimates of arrestees produced by the models were not greatly affected.[7] We eliminated market-demand-indicator-by-market-demand-indicator interactions (e.g., TEDS admission rates by Quest positive rates). Differences in summary statistics of model covariates for the ADAM counties versus non-ADAM counties showed that our initial naïve assumption that minimizing out-of-sample error by minimizing a tenfold error statistic or AIC value among nested models looking at only ADAM counties was not sufficient.[8] That is to say, the distribution of observed values for many of the market demand indicators is sometimes very different for the nation as a whole than it is for the almost exclusively urban ADAM counties. These covariates and interactions with means in ADAM counties that are in the tails of their respective national distributions are problematic for at least two reasons. First, the fact that data in ADAM counties are different than in the rest of the nation suggests the two groups may not be equivalent, and that the estimated coefficient based on ADAM data does not accurately represent national data. Second, when included, these covariates and interactions inflated our already large confidence intervals, typically without changing our final estimated number of CDUs very much.

[6] For both AIC and BIC, a lower test statistic is better.

[7] Discussed in further detail in Section 2.4.

[8] A *k*-fold cross-validation procedure randomly excludes 1/k share of observations from the estimation, predicting their value based on the other 1-1/k share of observations *k* times, and using the deviation from the true estimate to construct a root mean square error statistic based on out-of-sample predictions. In this case, we created ten partitions and ran the estimation over ten trials.

Section 3. Estimating the Number of Hard-Drug Users

3.1. Previous Approach

ONDCP (2012c) established estimates for 2000–2003 using ADAM, TEDS, and an assumption of a constant proportionality between chronic drug users in a county and treatment admissions nationally. The basic model is:

$$C = C_a * (V / V_a), \qquad (3.1)$$

where C_a represents the number of chronic users in ADAM counties, V_a represents the number of treatment admissions in ADAM counties, V represents the number of treatment admissions in the country, and C represents the number of chronic users in the country. In essence, the number of CDUs in ADAM counties is scaled up by the ratio of treatment admissions in the country relative to those counties.

The number of CDUs in ADAM counties is extrapolated from the number of adult male CDUs to include females and juveniles. The number of adult male CDUs in ADAM counties, in turn, is estimated by using responses to ADAM to identify adult male arrestees who are CDUs (i.e., past-month users of cocaine, heroin, or methamphetamine who used four or more times), and then generating probability-based estimates of arrests of adult male chronic users from respondents' 12-month calendar of major events (including arrests). The latter assumes that the probability distribution of arrest given chronic drug use follows a particular distribution, known as the Poisson distribution. If behavior follows that distribution, then one can infer how many CDUs there are who did *not* get arrested by looking only at data on CDUs who *were* arrested—and, hence, were within ADAM's sampling frame. The danger, of course, is that behavior might instead follow a zero-inflated Poisson distribution, which is a fancy way of saying that there could be CDUs who are not arrested in any given year, not because they were lucky (i.e., participated in criminal activity but managed not get arrested), but because they did not participate in criminal activity that carried any appreciable risk of arrest (e.g., committed no crimes besides drug use and only used those drugs in private).[9]

To overcome the loss of ADAM program data between 2004 and 2006, ONDCP (2012c) estimated the number of CDUs in those years using treatment rates calculated with self-reported treatment histories

[9] As explained in Section 3.2, we took a different approach to estimating numbers of CDUs who were not arrested.

from ADAM I 2000–2003 data and treatment numbers from TEDS. They adjust the treatment rate to try to account for differential treatment rates in ADAM I and non-ADAM I localities. They ultimately rely on outpatient treatment only, because they suspect the ADAM-derived inpatient treatment rates are less reliable (ONDCP, 2012c, p. D-4). They also chose to count only treatment admissions where the client indicated the substance was either a primary or secondary reason for seeking treatment. The TEDS-based estimate leads to numbers that are somewhat higher than those generated by their previous ADAM-based approach.[10] So as not to disregard the information used to generate the ADAM-based estimates, ONDCP (2012c) employs a hybrid that retains the earlier estimates for 2000–2003, and rescales the estimate downward to match the earlier level.

3.2. Current Approach

We consider a different approach that first estimates the probability (per arrest) that an adult male arrestee will test positive for use of a specific drug in a given county and year, using not only ADAM data but also a variety of other county- and state-level variables that are correlated with chronic drug use (previously described in Section 2). Next, we project year-by-year national estimates of the number of adult male arrestees who test positive to Uniform Crime Report (UCR) data for counties with coverage indicator scores greater than or equal to 65, and then separately extend the estimation to the rest of the country. We then adjust these year-by-year national estimates by a number of factors that account for the fraction of positive tests that are from CDUs, how many times an arrestee is arrested per year, CDUs who are not arrested, females, and juveniles (See Table 2.2).

Our approach differs from the previous approach in important ways. First, the ADAM I and ADAM II data—while very useful (and an integral part of our analysis)—are erratic, both in the sense of varying in inexplicable ways over time in certain counties and by not always mirroring trends in other indicators. Some, but only a modest proportion, of these irregularities can be pinned on problems with sampling weights, because puzzles remain even for the subset of locations and years for which the weights producing less volatile population estimates are available. The weights are meant to allow analysts to scale each observation up to represent a proportion of arrestees who have similar characteristics, increasing the representation of an arrestee who is deemed by a statistical matching algorithm to represent a group of similar criminal offenders in the population. However, these weights sometimes

[10] And unexpectedly so, since TEDS captures only about 70% of treatment admissions in the U.S. – though TEDS could at the same time generate an overestimate of the number of users since some people may be admitted into treatment more than once in a year.

generate questionable estimates at the county level as they result in improbably large swings in the estimated count of arrest events in cities from year to year. For example, the percent change (in absolute value) of the total number of arrests said to be represented by the weighted ADAM sample is 1 percent from 2000 to 2001, 8 percent from 2001 to 2002, and 9 percent from 2002 to 2003. These changes are not implausible; however, the percent change of the location-specific estimates is frequently 25 percent or more, which we consider implausible based on our understanding that the underlying phenomenon are likely to be relatively stable from year to year.[11]

We believe this makes ADAM stronger for estimating a proportion—specifically the proportion of arrestees testing positive—than a count.

3.2.1. Projecting Chronic Drug Use among Adult Male Arrest Events

Once we determined the best performing model for each drug's estimated prevalence rate among arrest occurrences in ADAM counties, we use the covariates' estimated coefficients to predict logistic-transformed county-year positive rates for the 43 ADAM I and roughly 3,100 non-ADAM counties based on their observed values for the covariates in each drug's model. After transforming the logistic rates back to estimated prevalence rates between 0 and 1, we have vectors of prevalence rates for all counties in the United States:

$$\hat{R} = exp(logistic[\hat{R}])/(1+exp(logistic[\hat{R}])), \qquad (3.2)$$

where $\hat{R}$ is the predicted prevalence rate for each drug. We also calculate standard errors for these values individually and transform upper- and lower-bound 95-percent confidence intervals for each using the same basic formula:

$$\hat{R}_{95\%} = exp(logistic[\hat{R}] \pm se(logistic[\hat{R}]) \times 1.96))/(1+exp(logistic[\hat{R}] \pm se(logistic[\hat{R}]) \times 1.96)))$$

(3.3)

At this point, we have a predicted value for each county-year, as well as 95-percent upper and lower bounds on these estimated prevalence rates. We then estimate the number of male arrest events using county-level Federal Bureau of Investigation (FBI) UCR data for those counties with coverage indicator scores of greater than 65 percent for all years between 2000 and 2010, 78 percent of the nation's

[11] For example, the estimated number of arrests in ADAM jumps nearly 60 percent in Washington, D.C., from 2007 to 2008, while falling nearly 25 percent in Portland over the same period.

roughly 3,150 counties.[12] Summing these totals, we have a point estimate for the number of male arrests for each year for counties in the United States with consistently high arrest reporting rates. This step mitigates the risk of building erratic county-level arrest reporting into our estimates. UCR national totals for adult male arrests already adjust for reporting uncertainty, so we can scale our estimates from counties with acceptable reporting rates to the Nation using their ratio:

$$\textit{National adult male arrests for drug } X = \Sigma(R_{xz} \times \textit{Male Arrests}_z) \times \textit{Total National Male Arrests}/\Sigma(\textit{Male Arrests}_z), \quad (3.4)$$

where R is the positive drug test prevalence rate for drug X in county Z; there are 2,415 counties with UCR coverage indicator scores greater than or equal to 65 for each year from 2000 to 2010. We use the same formula to calculate ranges for these national estimates, replacing our predicted R with its upper and lower values from the formula defined previously. We now have annual estimates for male arrest events involving an arrestee testing positive for cocaine, opiates, or methamphetamine. Table 3.1 shows these estimates.

Table 3.1: Estimated Adult Male Arrest Events per Year by Drug (in thousands)

	Cocaine			Heroin			Meth		
	Total	95% Confidence Interval		Total	95% Confidence Interval		Total	95% Confidence Interval	
2000	2,037	1,347	351	137	751	3,064	809	419	1,370
2001	1,851	1,252	529	233	989	2,766	800	426	1,326
2002	1,721	1,130	786	380	1,337	2,651	749	392	1,250
2003	1,725	1,128	999	512	1,611	2,653	753	402	1,232
2004	1,883	1,217	1,214	638	1,896	2,908	761	414	1,240
2005	1,992	1,283	1,472	817	2,210	3,047	716	388	1,160
2006	2,022	1,305	1,433	789	2,155	3,105	707	382	1,169
2007	1,880	1,217	1,297	687	2,000	2,879	703	387	1,152
2008	1,691	1,119	1,078	544	1,720	2,556	769	421	1,266
2009	1,549	1,004	952	464	1,560	2,376	867	471	1,452
2010	1,415	898	832	375	1,421	2,213	885	477	1,488

[12] Closely following Pacula et al. (2010), we do not consider arrest data from counties with coverage indicators < 65. The coverage indicator is an indicator of the quality of the data made available to the FBI and ranges from 0 (no information) to 100 (complete information). Imputations are made for counties with coverage indicators less than 100.

3.2.2. Projecting Chronic Drug Use for the Rest of the Population

We develop national chronic user estimates for cocaine, heroin, and methamphetamine based on the number of male arrest events estimates previously detailed.

Adjustment factor 1: Not everyone who tests positive is a CDU, because people who have used only one to three days in the past month can also test positive. So, we need to convert from arrests with a positive test result to arrests with a positive test result stemming from chronic use. We use the arrestees' self-reported drug use frequency from ADAM to calculate the share of arrest events with a positive drug test that is attributable to chronic drug users. The observed rates fluctuated modestly over time due to sampling variation, the composition of the arrestee pool, and fluctuations in arrest rates, but with no apparent trend. Since there was no trend, we assume that the rate is stable over time and multiply our total male arrest events estimates by the average proportion over the 2000–2010 time period, as indicated in Table 3.2.

Table 3.2: Adjustment Factor 1—Of Those Who Test Positive and Report Any Use in Last Month, What Share Reported Using on Four or More Days per Month?

		Cocaine	Heroin/Opiates	Methamphetamine
Total CDU share of arrests with positive tests	**Average**	**72.6%**	**89.0%**	**74.8%**
	Min	68.8%	86.4%	68.1%
	Max	75.2%	93.4%	79.4%

Adjustment factor 2: We next convert from total male arrest events to total male arrestees using past-year arrest information observed in ADAM.[13] These rates vary by city, but are stable over time. However, the sample sizes in many counties are too small to allow county-specific rates. Additionally, we combine near-daily (21+ days per month) users with more than weekly users (11 to 20 days per month) to maintain adequate sample size (see Table 3.3). Dividing the estimates of male CDU arrest events by this adjustment factor yields estimates of the number of adult males arrested who both test positive and were CDUs.

[13] We chose to disregard arrests resulting from outstanding warrants to eliminate the risk of combining arrests attributable to criminal activity that occurred further in the past and may not be tied to their current drug use. However, this exclusion had very little effect on our estimates.

Table 3.3: Adjustment Factor 2—Arrests per Arrestee

	One to Ten Days Past-Month Use	11+ Days Past-Month Use
Cocaine	1.51	1.68
Opiates	1.50	1.50
Meth	1.44	1.51

Adjustment factor 3: We supplement the result of the previous step with an estimate of adult male CDUs who were criminally active but who happened not to get arrested. In particular, dividing the number of CDUs arrested by the probability a CDU is arrested in a given year inflates the count to include those who happened not to have been arrested. This adjustment has similar logic to earlier ONDCP estimates. In particular, if arrests among criminally active CDUs follow a Poisson distribution, the probability of arrest in the past year for criminally active adult male CDUs is:

$$P\{ N > 0 \} = 1\text{-}exp(\text{-}\mu), \qquad (3.5)$$

where μ can be computed numerically from adjustment factor 2 via the formula:

$$Adjustment\ factor\ 2 = \mu/(1\text{-}exp(\text{-}\mu)) \qquad (3.6)$$

These μ and corresponding adjustment factors are shown in Table 3.4. The primary difference between this and what ONDCP (2012c) did is that we use the Poisson assumption only to extrapolate to *criminally active* CDUs who did not get arrested. Adjustment factor 4 supplements this with an NSDUH-based estimate of CDUs who are effectively at no meaningful risk of arrest.

Table 3.4: Adjustment Factor 3—Probability of Arrest in the Past Year for Criminally Active Chronic Drug Users

	One to Ten Days Past-Month Use		11+ days Past-Month Use	
	μ	**P{ N > 0 }**	**μ**	**P{ N > 0 }**
Cocaine	0.88	0.586	1.15	0.682
Opiates	0.88	0.584	0.87	0.583
Meth	0.78	0.541	0.89	0.591

Adjustment factor 4: ADAM collects data on people who are arrested (and booked). No matter the statistical adjustments, there really is no way to use those data on arrestees to estimate the size of a population of CDUs who are at no risk of arrest. Technically, of course, every CDU is at some risk of arrest because, by definition, they are using illegal drugs. Practically speaking, however, people who commit no offense other than drug use and who both purchase and consume inside a private residence, not in public, may be at very little risk of arrest.

We try to estimate this hidden—from ADAM—population using the household survey. Our (imperfect) proxy for a CDU being at negligible risk of arrest is never having been arrested at all, not just having avoided arrest in the last 12 months. Table 3.5 presents the average annual number of CDUs who have never been arrested in NSDUH, based on data from 2000 through 2010.

Table 3.5: Average Annual Amount of CDUs Who Have Never Been Arrested in NSDUH, 2000–2010

	Cocaine	Heroin	Meth
21+ days in past month	17,767	4,760	8,549
11–20 days	34,325	4,419	13,820
4–10 days	108,368	6,939	20,288
Less than 4 days in past month	291,025	8,326	14,292

Even though these individuals are not otherwise criminally involved, they are CDUs. Similar to what was done in the previous study with respect to inflating occasional users in NSDUH, we multiply the NSDUH-based estimates by a factor of four to compensate for underreporting in the household survey. By adding this product to the previously estimated number of criminally active CDUs, we generate the national estimate of adult male CDUs.

Adjustment factor 5: We next need to inflate the estimate to include adult females. There are a variety of data systems that reflect chronic drug use and record gender, including NSDUH, TEDS, CDC overdose data, and DAWN. Roughly speaking, males account for approximately two-thirds of cocaine mentions in these data systems. The equivalent rates for heroin and methamphetamine are 0.65 and 0.6, respectively. We scale our adult male CDU estimates to cover all adults by multiplying them by the reciprocal of the male share (1.5 for cocaine, 1.53 for heroin, and 1.67 for methamphetamine).

Adjustment factor 6: We follow essentially the same logic to inflate from adult CDUs (male and female) to all CDUs, including juveniles, with one minor addition. Whereas the male proportion of all mentions seemed stable over time, the juvenile proportion showed a trend over time for cocaine and meth (but not heroin). So for those two drugs, the adjustment to account for juveniles varies by year. In particular, the multiplier to include juveniles falls from adding roughly 5.1 percent to the adult total in 2000 to 1.7

percent in 2010 for cocaine, and from 8.3 percent to 4.6 percent for methamphetamine over the same period. We estimate that the adjustment to include juvenile heroin CDU remains constant at 3 percent over the period. Obviously, the bulk of CDUs are adults, so this adjustment is of minor importance.

3.3. Misreporting in ADAM

It is likely that some drug users in the ADAM sample misreport their frequency of use, and there are potentially several reasons for misreporting. First, responses to the question about use days in the past month tend to cluster around multiples of five, suggesting that many users are probably rounding up or down. Since we classify users into four groups (>20, 11–20, 4–10, and <4 past-month use days), misreporting due to responses clustering along these intervals could influence our consumption estimates, though the direction of the potential bias is difficult to guess.

While there is a decent amount of literature about whether drug users lie in self-report surveys, much less is known about the accuracy of use frequency conditional upon admitting some consumption. Morral et al. (2000) examined the validity of frequency self-reports using research questionnaires and clinical urinalysis test results from 700 self-reports made by a sample of methadone maintenance patients. Since the respondents were told that their ASI responses were for research purposes and would not be shared with their counselors, Morral et al. note that "although there were clinical incentives to reduce drug use, there were not programmatic incentives to underreport use on the ASI." Using a simulation approach, they estimated that 51 percent of those reporting one to ten days of heroin use had actually used on more than ten days; the comparable figure for cocaine was 22 percent (ten days is the threshold ONDCP used to denote hardcore use). While it is unclear how these results apply to nontreatment populations, with much depending about how much they trusted the researchers administering the ASI, this suggests that frequency underreporting could be an important issue.

One typically thinks the bias would be toward underreporting frequency—e.g., reporting use on only a handful of days when in reality use was daily. But there is no logical or theoretical reason why the opposite could not occur. For example, Leigh et al. (1998) found that relative to daily diary records, adolescents' retrospective statements about the past-month underreported their frequency of drinking but overreported their frequency of sexual activity. Admittedly, this is a very different population from ADAM respondents, but one could imagine some recent arrestees overstating frequency of use to make it seem like it was an extended binge that led to the criminal activity. They may also just misremember.

We believe our results are sensitive to misreporting, but ultimately the magnitude and direction of potential bias in our estimates, after the adjustments we make in our model, is ambiguous.

Morral et al. (2000) focused on heroin and cocaine, with heroin underreporting being a much larger problem. Since nearly 70 percent of the heroin CDUs in the ADAM sample already report using 21 or more days in the past month, there is not much room for underreporting for this drug. Even if we arbitrarily "promoted" 25 percent of the four-to-ten-day group to the 11–20-day group and another 25 percent to the 21-or-more group, that would only increase estimated total consumption (in pure metric tons) by about 5 percent. Considering cocaine, arbitrarily promoting 12.5 percent of the four-to-ten-day group to the 11–20-day group and 12.5 percent to the 21-or-more group, would increase total consumption by roughly 5 percent, depending on the year (6 percent for 2000; 4 percent for 2010). This type of underreporting could be a bigger issue for the stimulants, since the 21-or-more-days group for those drugs does not dominate the distribution of frequencies of use. However, a similar exercise for meth increases total estimated consumption by from 5 percent to 8 percent, depending on the year.

This thought experiment illustrates that this topic deserves further research and we make this point in the conclusion (Section 8). But for now, given questions about the applicability and size of the bias, as well as the fact that the possible changes are likely well within the 95-percent confidence intervals, we do not adjust for frequency underreporting since any such adjustment would be arbitrary.

Another concern is that we do not adjust for CDUs who test clean. This can happen for three reasons:

- they didn't use much during the month (but still more than three days)
- they used a lot in the month, but little or none in the few days preceding the drug test
- the test was reported a false negative.

While this can be an issue, there is also a possibility of a false positive or that they misreported use days. The bias due to the specificity and sensitivity of the urinalysis test can work in both directions, and we expect the effects to be countervailing and relatively small among the sources of uncertainty in the analysis.

Where this could have more of an effect is how we distribute CDUs across the three use-day groups. If we are more likely to miss those in the four-to-ten group, that means we are inflating the share of CDUs in the more frequent categories, and thus, total consumption. It is hard to imagine this having a large effect on the results and it is somewhat encouraging that it would likely bias the results in the opposite direction of the previous concern.

Section 4. Estimating Expenditures on Hard-Drug Users

4.1. Previous Estimates

The approach for generating monthly drug expenditures in the previous report is fairly straightforward, but impossible for us to replicate with the available information.[14] Using data from those who paid cash as well as those who bartered for at least some of the product, the report regressed monthly expenditure (value of last purchase multiplied by the number of times purchased on that day, multiplied by the number of days purchased in the past 30) on past-month use days for those who reported using everything they obtained (i.e., they did not sell or give away any). The report uses the estimated parameter on use days to impute values for arrestees who did gift or sell, as well as those in the first group who had "unbelievably large" monthly expenditures (possibly because of the aforementioned random incidence problem). The average monthly expenditures are calculated for each site and then weighted up by the number of chronic drug users in the site. These results are reported in Table 4.1 and then multiplied by 12 to generate an annual estimate.

Table 4.1: ONDCP (2012) Estimated Monthly Expenditures by Chronic Drug Users (nominal $)

	2000	2001	2002	2003
Cocaine	$979 ($847-$1,110)	$960 ($829-$1,091)	$1,001 ($866-$1,137)	$1,005 ($874-$1,136)
Heroin	$1,019 ($808-$1,229)	$1,077 ($814-$1,340)	$1,072 ($915-$1,228)	$1,054 ($907-$1,201)
Methamphetamine	$1,101 ($915-$1,288)	$1,062 ($876-$1,249)	$1,101 ($924-$1,278)	$1,022 ($841-$1,204)

Source: ONDCP (2012c) Table 3.13. Uncertainty intervals equal plus or minus two standard errors

After comparing their expenditure estimates for heroin, cocaine, and methamphetamine with others from the literature, they conclude: "Thus the estimates from others are broadly consistent with the estimates that appear in Table 3.13." To generate values for 2004–2006, ONDCP (2012) extrapolates the average monthly expenditure from 2000–2003; thus, the value for all three years is assumed to be constant. They acknowledge that "this is not a satisfying estimate and it may be wrong" (p. 40).

[14] As administrators of several years of the ADAM survey, Abt Associates, the previous authors of this report, leveraged detailed ADAM site booking census information, which is not publicly available, to scale survey respondents to all male arrestees.

4.2. Current Approach

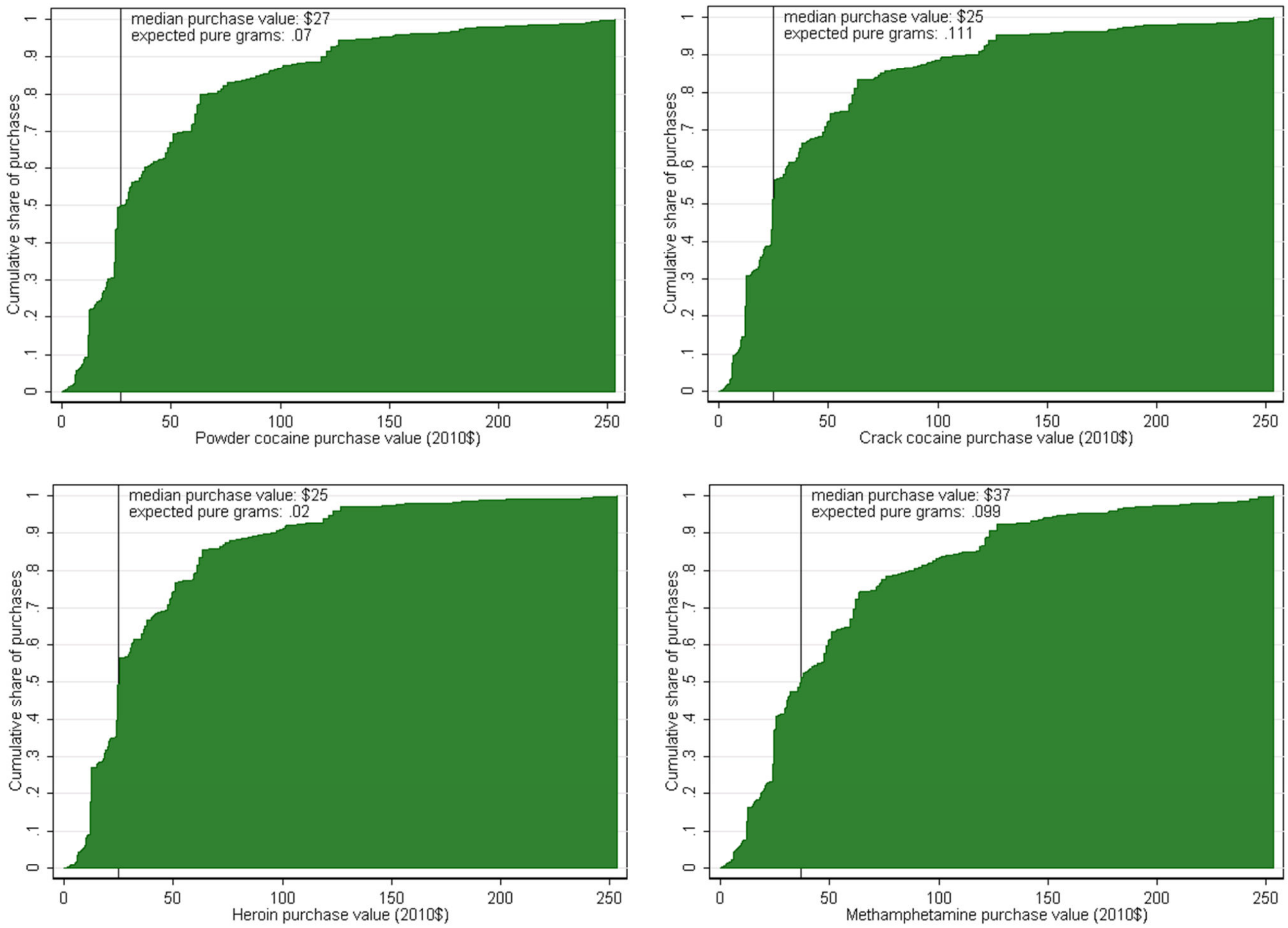
4.2.1. Arrest-Level Monthly Expenditure Estimates

For each city and user category, we calculated average expenditure by multiplying each observation's reported purchase amount by its fraction claimed for personal use. We incorporate this information so as not to overstate the amount of drugs purchased by end users by including expenditures on drugs meant for resale or sharing with others.[15] We further cleaned the expenditure series by excluding individual purchases with values greater than $200 assuming that expenditures greater than that critical value are meant for resale. The extreme values ignored make up between 4 and 10 percent of purchase observations for each drug in ADAM. We show the distribution of what we believe to represent street-level purchases for each drug in Figure 4.1. In these graphs, the median purchase value (in 2010 dollars) is marked by a vertical line. We imputed the pure weight in grams of this median purchase value based on expected purity values.[16] For all three drugs, more than 50 percent of purchases are for less than $40, and more than 80 percent of all purchases are for less than $70. This shows that our estimates are sensitive to the number of purchases per person we allow in each month. As shown in Section 4, the amount of pure weight typically purchased also has important consequences for our price-per-pure-gram estimates, and ultimately the value of total consumption we estimate for each drug. While there is likely considerable variation in purity and prices even within cities, for each purchase there is typically a very small amount of pure drug weight sold.

[15] More than half of reported drug purchases were designated to be purely for use by the buyer, but sharing or resale is more common among some drugs—76 percent of heroin purchases were for the buyer alone, as opposed to 52 percent of methamphetamine. We do not have complete confidence that drug users accurately report the share of drugs they plan to sell or give away; responses clustered at 0, 25, 50, 75, and 100 percent.

[16] We describe our Expected Purity Hypothesis methodology in Section 4 of this document.

Figure 4.1: Cumulative Density Functions, Median Purchase Values and Expected Pure Weight for Powder Cocaine, Crack, Heroin, and Methamphetamine (2010 $)



To scale arrestees' single-purchase information provided in ADAM to monthly estimates, we multiply by the number of past-day and then by past-month purchase days reported. As some buyers report questionably large same-day purchases, we limit the total number of past-month purchases to 120. By using this approach rather than independently capping the number of purchases on the same day, or the number of days in which a respondent purchased, or both, we allow for binge use to be incorporated. This does not prevent potentially improbable data points from influencing the estimates—there are several, particularly in the crack cocaine estimates where same-day purchases are very common among chronic users—but it allows for flexibility as we refine the model for each drug.

4.2.2. County-Level Monthly Expenditure Estimates

We create three pools of arrest-level expenditure estimates at the county level. The first is all ADAM I data between 2000 and 2003 for counties. The second and third included the subset of ten ADAM II counties for 2000 to 2003 and 2007 to 2010, respectively. In order to reduce volatility, only counties with at least 40 reported purchases over the pool's four-year period were included. For heroin and meth, drugs known to have geographically concentrated use, this step excluded ten of the 42 counties for which drug purchase data were available. The cleaned arrest-level expenditure estimates were averaged within each group. For most counties, this produced reasonable estimates. To minimize the influence of the bottom and top 5 percent of the distribution of expenditures for each drug, we replaced bottom and top 5 percent with the 5th and 95th percentile responses, respectively.[17] At this point, we have county-level expenditure estimates for medium, heavy, and daily/near-daily users within the chronic using arrestee population.

Pooled ADAM Site-Level Monthly Expenditure Estimates

Next, we use UCR adult male arrests to extrapolate our ADAM county estimates to represent county-level drug expenditures among arrestees nationally. We assume that the share of arrests reporting light, medium, heavy, and daily/near-daily use in ADAM is representative of the true mix of use among arrestees, and generate a single estimate of expenditures among arrestees for the ADAM counties with reliable UCR arrest data for adult males.

UCR provides a coverage indicator for each year of data. However, several counties noted to have excellent coverage exhibited improbable swings in arrest counts from year to year, including New York and Portland (which are traditionally large markets for several illicit drugs). This step results in a reduction of counties underlying the estimate to 29 for ADAM I and six for ADAM II.

To estimate expenditure figures for the 12 drug-by-frequency category combinations, we multiply each county's unique CDU category expenditure estimate by the county's estimated number of arrests we predict to be associated with that category from our earlier analyses. This step results in estimated expenditures for low, medium, heavy, and near-daily drug users in ADAM I, in ADAM II, and in the six ADAM II counties we retain during the ADAM I years 2000 to 2003. These estimates are built up from

[17] This cap was mostly an issue for crack cocaine, where it may be that frequent small daily purchases were reported independently along with the aggregate amount spent on that day, resulting in extremely large initial expenditure estimates.

county-level estimates weighted by the size of each county's arrestee population and each drug's prevalence therein.

The basic formula for generating the average monthly drug expenditure for each pool of ADAM counties and years is:

$$\text{Avg Monthly Expenditure}_g = \Sigma_c((\text{County Expenditure Avg})_{c,g} \times (\text{Male Arrests in County})_{c,g} / (\text{Male Arrests in County})_g), \qquad (4.1)$$

where g indexes each of the four use categories: light, medium, heavy, and near-daily; and c indexes the counties in each ADAM pool.

The change between the periods 2000–2003 and 2007–2010 in ADAM II average expenditures was used to extrapolate the growth between these periods in the larger set of ADAM I counties, for which only 2000–2003 data are available. We assume that the ADAM I counties more closely represent the nation as a whole, and thus use these estimates to build our national drug market expenditure estimates.

4.3. Adjusting for Occasional Users

Those using a drug less than four times per month contribute a small portion to overall consumption. As this population is not necessarily well-described by ADAM, or by a treatment or mortality series, we rely on use rates reported in NSDUH to estimate this group. We calculate total spending by occasional users to be 12 percent, 3 percent, and 6 percent of total use for cocaine, heroin, and methamphetamine, respectively. These estimates are based on the ratios of total use to chronic use observed in treatment admissions from TEDS, and days of use among arrestees in ADAM II counties and NSDUH respondents over 2000–2010. Assuming underreporting of occasional use, we use ONDCP's (2012c) assumption that NSDUH captures 25 percent of actual occasional hard-drug users. Our estimates for each drug are shown in Table 4.2.

Table 4.2: Millions of Occasional Users by Substance

	Cocaine	Heroin	Meth
2000	2.36	.17	3.26
2001	3.33	.13	2.74
2002	4.54	.21	2.87
2003	6.03	.13	2.67
2004	3.98	.12	2.86
2005	5.47	.18	2.75
2006	5.26	.38	3.06
2007	4.58	.15	1.80
2008	4.23	.24	1.31
2009	4.16	.34	2.00
2010	3.85	.33	1.47

Notes: This estimate is based on annual NSDUH estimate for number of light users (one to three days in the past month), multiplied by four to account for underreporting. Estimates for 2000 and 2001 are not comparable to other years due to changes in survey design. Meth estimates prior to 2007 are not comparable to later years due to changes in NSDUH.

Section 5. Estimating Hard-Drug Consumption

5.1. Previous Approach to Estimating the Average Price Paid Per Pure Gram Purchased

The previous report followed a six-step process to develop price series for user-level cocaine, heroin, and methamphetamine purchases based on STRIDE undercover purchase data between 1998 and March 2007 (ONDCP, 2012c). They first limited the dataset to include only purchases between $20 and $200 in nominal terms to mitigate the risk of nonretail-level purchases entering the data. Next, they excluded purchase of greater than 10 pure grams for the same reason. Finally, they defined a price-purity metric as pure grams per dollar spent and exclude purchases with values greater than 0.1 as outliers—to rephrase, they excluded purchases with values below $10 per pure gram for these drugs for being unbelievably low.

Second, they imputed purity values for 95 percent of observations with reported purity levels of zero because many of these were not true zeros, but instead were not tested due to sample amounts being too small. They arrived at the 95-percent threshold after examining the distribution of zero-purity observations on their raw mass using a simple logistic regression. There is no true way to detect actual zeros, so they discarded 5 percent of zero-purity observations randomly. For the remaining 95 percent, they imputed purity values of 30 percent for crack cocaine, 7.5 percent for powder cocaine, 3.5 percent for heroin, and 1 percent for methamphetamine. These values were approximately the second percentile of detected non-zero purities observed in the data.

Third, they generated an expected pure amount purchased per dollar measure, the reciprocal of price per pure gram, by regressing the observed values of that metric on dummies for calendar quarters and Metropolitan Statistical Area (MSA) and interaction terms of price paid by MSA. As defined, their model does not include a covariate for price paid alone. They also state that they used both ordinary least squares and weights, though their weighting method is not defined in this step.

Fourth, they averaged predicted values of their pure grams per dollar metric for each MSA in each quarter for which there were data. They state that these were weighted using ADAM, but the mechanism is not defined in this step.

Fifth, they weighted the MSA-level pure grams per dollar values by the ratio of purchase observations in that MSA to total purchase observations, thereby assuming that the spatial distribution of drug market activity and prices is accurately depicted by undercover purchases entering the data. To minimize the influence of purchases concentrated in a single geography, they discard observations from Washington, D.C. DEA heroin Domestic Monitoring Program data were also excluded from the calculations because they were overrepresented in the data and may not have accurately represented the street price of heroin.

Sixth, they developed a national average price series for each drug by dividing a nationally representative weighted sum of expenditures reported in ADAM by a pure gram estimate derived by combining estimated purities from STRIDE with quantities reported in ADAM, again scaled up to a nationally representative sum. As a result, their price series is not evaluated at a referent quantity, and instead represents the price for the weighted average purchase size in the nation for a given year. The underlying average purchase size is not recorded in their report. ONDCP (2012c) prices are relatively lower than the Expected Purity Hypothesis price estimates developed by the Institute for Defense Analysis (IDA; Fries et al., 2008) for the smallest pure gram quantity interval that is closest to the price of each drug faced by users (see Table 5.1). Given that the same STRIDE data were used, this means that the implicit quantity they evaluate prices at is higher than the referent quantity used by IDA (and RAND previously).

Table 5.1: ONDCP (2012c) Price per Gram Series (rows listed as "WAUSID") Compared to Fries et al. Values at Four Purchase Sizes (nominal dollars)

		2000	2001	2002	2003	2004	2005	2006
Cocaine powder	0.1g <= AMT <= 2g @ 0.75g*	$186.37	$194.18	$137.13	$147.54	$134.02	$132.28	$130.37
	2g < AMT <= 10g @ 5g*	$113.14	$93.52	$84.36	$85.80	$73.83	$66.74	$59.98
	WAUSID	$106.69	$119.04	$109.24	$95.84	$88.99	$81.39	$75.97
Crack cocaine	0.1g <= AMT <= 1g @ 0.3g*	$252.41	$226.85	$206.94	$187.91	$178.67	$161.23	$152.71
	1g < AMT <= 15g @ 5g*	$111.59	$101.46	$93.18	$85.98	$83.16	$72.23	$69.64
	WAUSID	$141.63	$160.19	$148.98	$123.83	$111.12	$106.23	$101.37
Cocaine (16.7% powder)	WAUSID	$135.80	$153.32	$142.34	$119.16	$107.42	$102.08	$97.12
Heroin	0.1g <= AMT <= 1 g @ 0.4g*	$457.24	$430.88	$404.44	$405.46	$417.76	$381.45	$386.86
	1g < AMT <= 10 g @ 2.5g*	$300.68	$271.35	$271.02	$265.91	$298.14	$254.49	$265.57
	WAUSID	$375.66	$393.25	$390.26	$376.33	$394.38	$401.84	$399.20
Meth-amphetamine	0.1g <= AMT <= 10 g @ 2.5g*	$212.25	$211.84	$178.62	$171.86	$164.72	$120.29	$165.99
	10g < AMT <= 100 g @ 27.5g*	$158.68	$128.05	$116.14	$95.04	$85.25	$64.33	$95.42
	WAUSID	$178.50	$161.55	$141.14	$113.98	$109.91	$104.13	$114.03

Source: Fries et al. (2008). To obtain prices for cocaine, crack and powder prices were averaged so that crack received a weight of 0.833 and powder received a weight of 0.167. Weighting was determined by the proportion of crack/powder samples in the file.

Source: ONDCP (2012), Table 3.18

5.2. Current Approach

Using purchase patterns described in ADAM as a guide, we believe both sets of quantities to be considerably higher than the typical purchase and choose to use smaller referent quantities for cocaine, heroin, and methamphetamine. We respect the idea of developing a new approach that estimates pure quantity obtained as a function of amount spent, but we have concerns about whether it works out well in practice, and in the present work went with a simple adaptation of the familiar approach pioneered by Arkes et al. (2004) and used by Fries et al. (2008).

We began by replicating the approach we took in our 2004 study of the STRIDE data (Arkes et al., 2004). We received STRIDE data covering 1982 through mid-June 2012, which included approximately 1.6 million drug seizure and undercover purchase observations. Of these, we limited the data to include only the drugs of interest: 792,000 observations covering cocaine, heroin, and methamphetamine. We next eliminated approximately 30,000 observations that occurred outside the United States.[18] As our focus is on purchases and seizures, another 40,000 observations were discarded for not meeting that criterion based on a previous acknowledgment that other types of acquisitions (e.g., "flashing money") may included spurious or erratic information. We also took several additional cleaning steps, eliminating impossible values for purity and weight, as well as quantities denominated in units other than grams, which tend to be very large and missing key information such as price, primary drug type, and location. These reduced our sample size by another 30,000 observations, to approximately 692,000.

Table 5.2: Criteria for Outlier Deletion Based on Reported Cost (in 2010$ based on 2002 nominal values)

Drug	Nominal Price	Real Price per Gram	Real Price per Gram
Cocaine	< $ 3.64	< $ 2.42	> $ 3,640
Heroin	< $ 3.64	< $ 9.09	> $12,120
d-Meth	< $ 3.64	< $ 2.42	> $ 3,640
Marijuana	< $ 0.12	< $ 0.06	> $ 121

Observations with very small quantities tend to show a lot of volatility in reported price and purity, so we chose to eliminate the 82,000 observations with masses less than 0.1 grams. As we are estimating prices, we also discarded observations without price information, losing 382,000 from our sample. Using CPI to adjust prices to the 2002 levels used in our 2004 analysis, we excluded roughly 1,000 outliers with reported costs that did not meet the criteria listed in Table 5.2.

[18] Observations from Puerto Rico were also excluded from the analysis.

Our final cleaning steps were meant to limit the amount of artificial volatility in the time series. As their markets and resulting prices are not perfectly overlapping, we estimated powder and crack cocaine independently. We discarded the 196 crack cocaine purchases reported between 1982 and 1986, as they are too few to be predictive of crack prices in those years. For each drug model, we also required that a cell, defined as a particular city and year, must have at least five observations to be included in the model. Unlike ONDCP (2012), we did not exclude heroin Domestic Monitoring Program data, as it appears to be very similar to other heroin data in the sample. Unlike Arkes et al. (2004), we included both d-methamphetamine and methamphetamine observations categorized to have an unknown isomer as they look statistically similar in price, purity, and typical purchase quantities. We also estimated this model at the year frequency rather than quarter. Our final analytic dataset contained approximately 206,000 observations (see Table 5.3).

Table 5.3: Representation of Each Drug in the Analytic STRIDE Dataset

Drug	# of Observations	Share of Sample
Crack cocaine	75,988	36.8
Powder cocaine	60,130	29.2
Heroin	41,034	19.9
Methamphetamine	29,102	14.1
Total	206,254	100

Expected Purity Hypothesis Model Specification

We employed a two-stage expected purity model to first generate purity estimates for every observation. Following the Expected Purity Hypothesis, we developed an empirical model of price that was based on the assumption that the price a buyer is willing to pay for the drug is determined by his/her perception of purity at the time of the transaction, not actual purity of the drug. For our first stage, the empirical specification of the random coefficient purity model can be written as:

$$Purity_{ijk} = \alpha_{0k} + \alpha_{1k} time_{ij} + \alpha_{2k} AMT_{ijk} + \varepsilon_{ijk} \qquad (5.1)$$

$$\alpha_{0k} = \gamma_0 + u_{0k}$$

$$\alpha_{1k} = \gamma_1 + u_{1k}$$

$$\alpha_{2k} = \gamma_2 + u_{2k}$$

$$\varepsilon_{ijk} \sim N(0,\sigma^2),$$

where $time_{ij}$ is a vector of dummy variables representing a year (i.e., 31 years, though only half of 2012 is captured in the data)[19] and AMT_{ijk} is the raw weight of the *ith* observation in city k at time j. The coefficient α_{0k} represents the intercept for city k, α_{1k} is a vector for the time coefficient for city k, and α_{2k} is the amount coefficient for city k. The terms, γ_0, γ_1, and γ_2, respectively, are the overall mean estimates for the intercept, time, and amount effects. The random coefficients for the intercept, amount, and time are assumed to be independently and identically distributed as previously specified.

We assigned every observation in the dataset to one of 38 geographic groups, 29 cities, and the other areas in their metropolitan areas, and nine census districts. Almost 50 percent of the 206,000 observations fell into the non-metropolitan areas, and the remainder occurred in or near a major city. Equation (5.1) is estimated as a linear random coefficients model with observations nested in cities (and census districts) using the 'xtmixed' command in Stata 12, and the estimates of expected purity generated from this model are then used in the second-stage price model. Cases with predicted purities not between 0 and 1 were modified to equal either 0.5 or 99.5 percent. To reduce the effect of outliers, we iteratively ran the model to discard expected purity estimates with standardized residuals that were more than 3.09 standard deviations away from the mean. This process excludes extreme outlier observations based on a set of estimated coefficients. The coefficients were re-estimated for each iteration to avoid the influence of these possibly erroneous outliers in the final purity predictions. We typically lost between 3 and 6 percent of observations from this step, with heroin on the high side of loss and cocaine on the low side. As the sample size for each drug is fairly large and the model is highly parameterized, we estimated the model using a simplified random intercepts model when the random coefficients specification failed to converge during the iterative step. As the random intercepts model is nested in the random coefficients model, this is unlikely to affect the accuracy of our estimates. The final model from which expected purity values are estimated always uses random coefficients. Table 5.4 shows our annual counts of observations, expected purity estimates, and actual reported purity values in the data. As can be seen, expected purity values for heroin values are lower than observed and declining over time, suggesting systematic differences between perceived and actual purity while purity erodes.

[19] To identify the model, we assume the excluded year is fixed. As the estimated coefficients on the other year dummies are measured relative the excluded year, this should not affect our estimates significantly. This is the most flexible specification for time.

Table 5.4: Sample Size, Estimated and Actual Purity Based on the Expected Purity Model

	Crack			Powder Cocaine			Methamphetamine			Heroin		
Year	**Count**	**EP**	**Actual**	**Count**	**EP**	**Actual**	**Count**	**EP**	**Actual**	**Count**	**EP**	**Actual**
2000	1,137	.67	.65	160	.61	.60	241	.42	.41	949	.37	.46
2001	969	.64	.63	102	.56	.55	284	.52	.52	959	.31	.40
2002	656	.68	.67	92	.66	.65	288	.57	.56	786	.32	.42
2003	534	.72	.72	117	.67	.67	304	.57	.57	956	.29	.39
2004	573	.76	.75	133	.61	.60	296	.56	.56	826	.25	.34
2005	673	.78	.78	130	.64	.63	541	.69	.69	872	.28	.37
2006	776	.79	.79	104	.72	.71	439	.45	.44	899	.25	.34
2007	884	.75	.74	122	.62	.61	328	.45	.45	949	.28	.37
2008	816	.66	.65	123	.54	.53	287	.50	.49	914	.28	.37
2009	692	.59	.58	131	.43	.43	448	.65	.65	886	.26	.34
2010	558	.55	.54	144	.48	.47	665	.79	.79	788	.16	.25

Note: Our model is based on STRIDE observations from the Q1 1982 to Q2 2012 for powder cocaine, heroin, and meth, and starting with Q1 1986 for crack cocaine. We included only observations with purchase quantities that are similar to small street-level retail buys typical of chronic users. This truncation limited the total number of observations for crack cocaine to 19,297, for powder cocaine to 7,215, for heroin to 20,565 and for meth to 7,121. This table shows only the 11 years relevant to this study.

Price Model Specification

To model prices, we again used a linear random coefficients model, as was originally described in McCullagh and Nelder (1989). We specify:

$$E(ln(real\ price_{ijk})|\gamma_{0k}, \gamma_{1k}, \gamma_{2k}) = \gamma_{0k} + \gamma_{1k}time_j + \gamma_{2k}[ln(AMT_{ijk} + ln(predicted\ purity_{ijk})] \quad (5.2)$$

$$\gamma_{0k} = \beta_0 + \varepsilon_{0k}$$

$$\gamma_{1k} = \beta_1 + \varepsilon_{1k}$$

$$\gamma_{2k} = \beta_2 + \varepsilon_{2k},$$

where the natural logarithm of transformed real price for observation *i* in year *j* in city *k* is estimated as a function of year effects, city effects, and the sum of the natural logarithms of amount and expected purity from our stage 1 estimation. We generated real prices by adjusting price to 2010-equivalent levels using the Consumer Price Index. Note that the sum of log-amount and log-purity is equivalent to the natural log of expected pure grams, so this model can be understood as price being a function of time, location, and expected pure grams. The β terms are the mean effects of each control on expected real price, where β_2 is our estimate for quantity discount. The model is flexible in its specification, as we do not explicitly estimate price per pure gram, but can estimate pure prices by virtue of evaluating the function at a purity value of one for the predicted purity coefficient.

We followed the same steps to estimate this model as we did for expected purity in stage 1. As our model's price predictions depend on the amount of pure grams being theoretically transacted, we then evaluated the models at specific amounts for each drug to replicate previous work. After validating our results against Fries et al. (2008) and our previous report (Arkes et al. 2004), we established new amounts that are closer to the quantity (in pure grams) typically purchased for each drug in a street-level sale, shown in Table 5.5.

Table 5.5: Pure Gram Price Evaluation Levels for Each Drug

	Crack Cocaine	Powder Cocaine	Heroin	Methamphetamine
Evaluation level	0.25 pure grams	0.25 pure grams	0.1 pure grams	0.25 pure grams

Since retail prices are often reported for transactions of one pure gram, these levels might initially seem low. Yet, if anything, the adjustment down to these levels is conservative. At the average prices over the 2000–2010 period, these correspond to purchases of about $50 for crack and powder cocaine and $90–$100 for heroin and meth.

ADAM I respondents described 71,251 past-month purchases of $200 or less including multiple purchases by the same respondent; e.g., a respondent reporting spending $25 on their past purchase and 30 purchases in the past month is considered to spend $750 in the past month ($25 over 30 purchases). (As ONDCP (2012) did in its past report, purchases over $200 were excluded as being unlikely to represent final retail purchases.) The two most common purchase sizes were $20 and $10, each accounting for close to one in five purchases. The median purchase was $20, although the median gram was obtained in a $100 purchase. Hence, a purchase of 0.25 pure grams—or roughly $50—is perhaps the most representative purchase size for powder cocaine. The situation for crack is very similar. The modal purchases are $10 and $20, each accounting for about one in five purchases; $20 is the median purchase, and the median gram was obtained in a $100 purchase.

Meth purchases were slightly larger. The most common purchase size was $20, but $50 was the second-most common purchase size, and the median was $40. However, since meth prices are higher per pure gram, 0.25 pure grams was usefully central to the overall distribution.

Heroin purchases were even a bit smaller, in dollar value, than cocaine purchases. $20 and $10 were the two most common, and $20 was again the median purchase, but the median gram was purchased in a $60 transaction. Cutting the referent quantity from 0.25 to 0.1 pure grams brought it closer to the typical heroin transaction. Arguably, the meth and heroin referent quantities could have been even smaller, perhaps 0.1 and 0.05 pure grams, respectively, but we favored round numbers, consistency across the simulants, and not departing too far from the tradition of treating one pure gram as a typical retail purchase size.

We generated annual price estimates for each of 29 cities (falling into nine census divisions) by evaluating price per gram at each evaluation level, year, and location. However, not all cities have data in all years, so we used a city- and year-fixed effects OLS model to impute estimates for city-years with missing values for each drug. To collapse the city-year series to annual national estimates, we weighted each city-year price estimate by the city's share of the total population of the 29 cities for each year. The census division weights are calculated in the same way, as a share of the nine divisions. We believe price estimates from a single metropolitan area are more informative than those generated from large census divisions. By that logic, we distributed 1/30th of the total model weights to the nine census divisions and 29/30ths to the cities. For each city *k* and year *j*, we estimated price per pure gram as:

$$P_j = \Sigma P_{jk}\omega_{jk} \qquad (5.3)$$

Section 6. Details of Marijuana Analyses

6.1. Previous Marijuana Estimates

The earlier versions of this report employed a straightforward approach based principally on NSDUH data (as in this report). Users were counted as those who reported using marijuana or hashish in the last 30 days. A chronic user was defined as someone who consumes marijuana four or more days per month (e.g., weekly), and an occasional user as someone who consumes three or fewer days per month.

Then, the counts of total, chronic, and occasional users were adjusted upward to "reduce reporting bias" that results from that fact that "respondents to NSDUH frequently understate their drug use" (ONDCP, 2012c). The authors note the size of the required adjustment is unclear, but identifies what is deemed a justifiable adjustment factor based on the study of Harrison et al. (2007), which explores the reliability of self-reported drug use. The adjustment is calculated as the fraction of users divided by fraction that reported using. They assume everyone who tests positive or reports use is a user. Some unknown fraction also used despite a negative test and denial of use; assuming that unknown fraction is zero leads to an adjustment factor of 1.35, and assuming the truthfulness rate for those who test positive is the same as for those who test negative leads to an adjustment factor of 1.27. The adjustment factor is assumed to be 1/0.75 (which is between 1/1.35 and 1/1.27, and where 0.75 represents truthfulness). They indicate this estimate of truthfulness accounts for most of the uncertainty.

The previous version of this report first estimated total expenditures based on NSDUH respondents' self-reports of how much they spent on their last purchase, combined with information about the amount of their last purchase and frequency of purchase during the month. It then estimated consumption by dividing expenditures by estimates of marijuana prices. (Estimates from NSDUH of grams purchased are used only in a secondary role.) These estimates of expenditure were divided by price series information to generate consumption estimates. The price estimates were based on the expected purity analysis output from IDA's work using DEA STRIDE data. To account for quantity discounts, they use log-log regression of price on quantity to calculate separate average prices tranches of 0.1g to 10g, 10g to 100g, and more than 100g. Based on these price series, they calculated a nation-level price per gram.

The previous version of this report also provided an alternative estimate of marijuana users based on data from ADAM, UCR, TEDS, Monitoring the Future (MTF), and NSDUH, similar to the ADAM-based approach they employed for heroin, cocaine, and methamphetamine. It relied upon the assumption that chronic users have an appreciable probability of being arrested (for drug use, drug selling, or a nondrug offense). This estimate was built through a series of steps, starting with a national estimate of chronic adult users from ADAM, UCR, and TEDS data. Chronic juvenile users were determined with MTF data, and occasional users with NSDUH data. They followed an approach similar to that used in the NSDUH-based marijuana analysis to develop expenditure estimates. Again, they adjusted for outliers and rescaled using the regression-based methods previously discussed. But the alternative approach covered only the years from 2000 through 2003, since ADAM I was discontinued after the 2003 round.

Table 6.1 summarizes the baseline marijuana results from the previous version of this report (and shows two additional years of calculations that were provided to us but weren't included in the 2012 report).

Table 6.1: Marijuana Figures from the Previous Version of This Report (ONDCP, 2012)

		2001	2002	2003	2004	2005	2006	2007	2008
# Occasional users, unadjusted (NSDUH)	K	4,017	4,496	4,816	4,957	4,534	4,686	4,611	4,813
# Chronic users, unadjusted (NSDUH)	K	7,617	9,599	9,556	9,331	9,607	9,760	9,361	9,949
# Occasional users, adjusted by 4/3 (NSDUH)	K	5,356	5,995	6,422	6,609	6,046	6,248	6,148	6,418
# Chronic users, adjusted by 4/3 (NSDUH)	K	10,156	12,799	12,742	12,441	12,809	13,013	12,482	13,265
Annual expenditures (unadjusted)	$B	$18.90	$26.80	$27.70	$22.70	$25.10	$25.30	$28.50	$32.10
Annual expenditures (adjusted=unadjusted*1.33)	$B	$25.20	$35.73	$36.93	$30.27	$33.47	$33.73	$38.00	$42.80
Unit price	$/g	$5.48	$8.03	$7.27	$5.84	$7.05	$7.87	$7.14	$7.14
Annual consumption	MT	4,595	4,453	5,082	5,182	4,750	4,285	5,322	5,994
Consumption per person *per year* in grams	g	296.21	236.93	265.18	272.03	251.93	222.48	285.68	304.55
Consumption per person *per month* in grams	g	24.68	19.74	22.10	22.67	20.99	18.54	23.81	25.38

Looking more closely at the alternative (ADAM-based) approach, we see that, in general, this method results in marginally lower use rates than in the NSDUH-based method. Using the ADAM price and quantity series, the previous report generated price and tonnage estimates that are very similar to those estimated using NSDUH (see Table 6.2). This table, however, illustrates why the consumptions estimates are so disparate in the early years of the decade: Their ADAM-estimate shows increased unit price without increased expenditures, leading to a dramatic decrease in consumption, while their NSDUH-estimate shows increasing unit price and expenditures, leading to nearly stable consumption estimates. The current analysis suggests that users, consumption, and expenditures were relatively flat in the early part of the decade (discounting the lower values recorded in 2000 and 2001, which are likely driven by the NHSDA-NSDUH discontinuity).

Table 6.2: Comparison of Baseline (NSDUH) and Alternative (ADAM) Approaches from ONDCP (2012)

			2000	2001	2002	2003
Unit price	$/g	NSDUH		$5.48	$8.03	$7.27
		ADAM	$6.4	$7.06	$9.3	$8.53
Expenditures (adjusted)	$B	NSDUH		$25.20	$35.73	$36.93
		ADAM	$34	$34	$34.5	$35.8
Consumption	MT	NSDUH		4,595	4,453	5,082
		ADAM	5,113	4,824	3,751	4,130

6.2. Random Incidence

Broadly speaking, there are two ways that surveys can support estimates of spending on marijuana: asking directly what people have spent on marijuana, and asking how much marijuana they used. The latter is relevant because multiplying quantity consumed by price per unit quantity is an alternative way to estimate spending.

If surveys routinely asked respondents how much they had spent on marijuana in the last week, understanding of national spending on marijuana would be greatly improved. One would still need to make adjustments to deal with use by populations outside the survey's sampling frame and for under-reporting, but the basic calculation would be trivial; just add across users and multiply by 52.

Since 2002, NSDUH has asked a question about spending on marijuana. Alas, the question inquires about the amount spent only on the most recent purchase, not the amount spent over a particular period of time. One might still hope to recover spending over time, since there is also a question about how many purchases were made in the last month. However, it questionable to assume that multiplying the number of purchases by the amount spent on the most recent purchase necessarily provides a sound estimate of total spending. It would—only if the most recent purchase were representative of all purchases. Yet that is a strong and untested assumption, and there is reason to believe that is not the case. It seems plausible, if not unlikely, that most recent purchases are systematically larger than are typical purchases, so the naïve approach may tend to overestimate spending. The extent of the upward bias could be very large, and is essentially impossible to bound.

There are many reasons why there could be bias. For instance, if surveys tend to be conducted during weekdays or during the daytime, the most recent purchases may underrepresent purchases made on Friday night.

A potentially larger and subtler concern goes under the technical name "random incidence." The common-sense logic is that if unusually large purchases are followed by unusually long intervals before the next purchase, then the survey is more likely to fall into one of those unusually long interpurchase gaps with a most-recent-purchase that is unusually large. Conversely, if very small purchases are soon followed by another purchase, then it is rather unlikely that the person administering the survey will happen to knock on the door at just the right time to catch the respondent between that pair of purchases.

This concept is perhaps most easily grasped with a numerical example, and for reasons to be explained shortly, we illustrate it for estimating total quantity consumed, rather than the amount spent buying marijuana. Imagine a population of users who consume one gram every day (and buy all that they consume). Furthermore, suppose these people make a one-ounce (28 gram) purchase every month, and also two one-gram purchases (e.g., they buy a gram on the first day of the month, an ounce on the second day, and then another gram on the last day of the month). Since an ounce is 28 grams, 1+ 28 + 1 = 30—meaning that they buy exactly enough to cover their use.

These particular numbers are selected to make the example's arithmetic simple, but they are not unrealistic. Those who consume as much as 1 gram (~2 joints) per day might well wish to obtain most of their marijuana an ounce at a time, to take advantage of quantity discounts, but they might also occasionally make smaller purchases, perhaps while traveling or if they want to experiment with a different kind of marijuana.

At any rate, those administering surveys have no idea when the respondent has made a small or a large purchase; they just knock on the respondent's door on some random day during the month. This means there is one chance in 30 that the person administering the survey shows up on the first day, a day on which the most recent purchase was of 1 gram. And there is also one chance in 30 that they show up on the last day, the (only) other day of the month on which the most recent purchase was only one gram. And there are 28 chances out of 30 they administer the survey on some other day of the month, for which the most recent purchase was 28 grams.

So, the expected size of the most recent purchases is (2/30) * 1 + (28/30)*28 = 26.2 grams. Regardless of which day the survey is administered, the number of purchases made within the last 30 days is 3. So the naïve approach would estimate quantity consumed as 26.2 * 3 = 78.6 grams per month, or more than two-and-a-half times the actual consumption rate of 30 grams per month. That is an enormous error. Furthermore, there is nothing that guarantees that the error could not be even greater. If there were five one-gram purchases for every 28-gram purchase, then consumption would be overestimated by almost a factor of five.

Essentially, the same problem can occur when estimating spending by multiplying the number of purchases in the last month or year by the amount spent on the last purchase. The twist is that there is usually some sort of quantity discount on price. That slightly moderates the extent of the error—and complicates the example's arithmetic, which is why we illustrate the principle with consumption rather than spending. But the same principle leads to the same bottom line: Multiplying the number of transactions by the size of the most recent transaction—whether size is measured in grams or dollars—can overestimate aggregate spending or consumption.

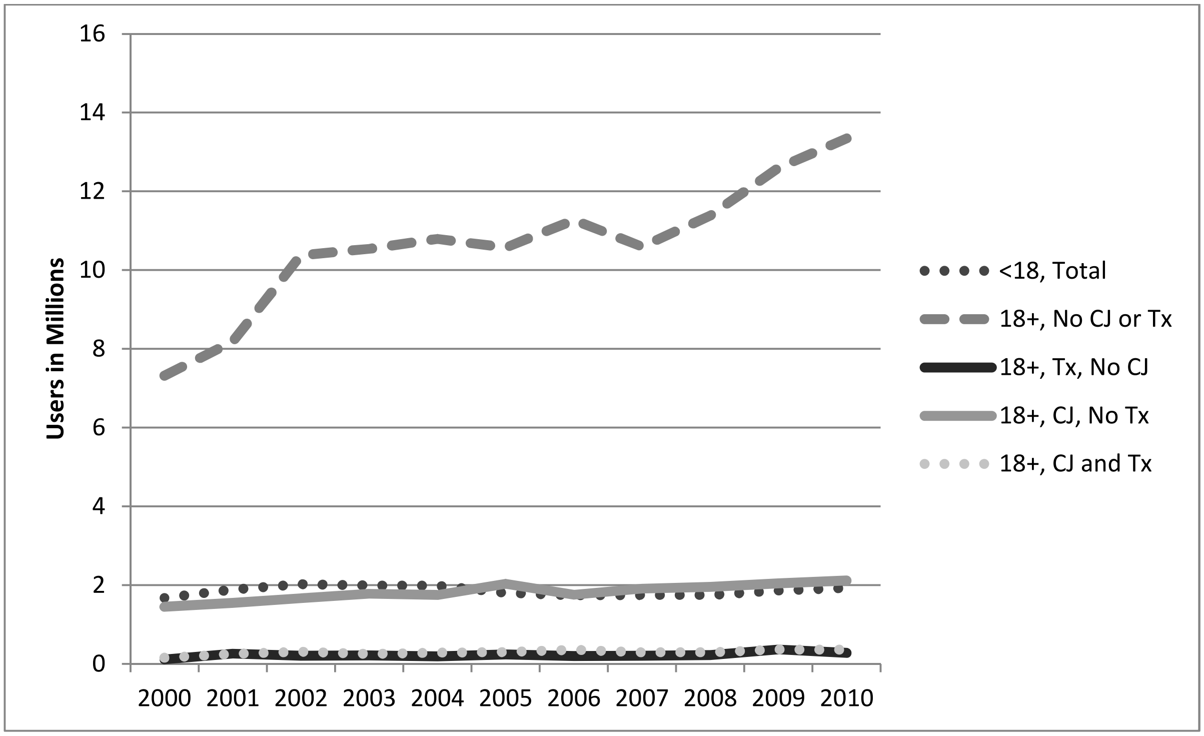
So in this report, we instead took the more laborious approach of starting with a metric whose quantity is asked about for an entire period, not just with respect to the most recent instance; that metric is days of use. NSDUH asks about the number of days on which the respondent used marijuana in the last month or year. Multiplying by quantity consumed per day of use converts that to quantity consumed in the last month or year. Multiplying again by price produces an estimate of spending.

6.3. NSDUH Prevalence Adjustments

We disaggregated the NSDUH data into subcategories that facilitate examination of the extent to which the NSDUH user estimates are biased downward. The subcategories we consider are youth (ages 12–13, 14–15, and 16–17), adult (i.e., 18 or older) respondents who report past-year treatment for illicit substance use, adult respondents who report past-year criminal justice involvement (i.e., arrested and booked, on probation, or on parole in the past year), and other adult respondents. As illustrated in

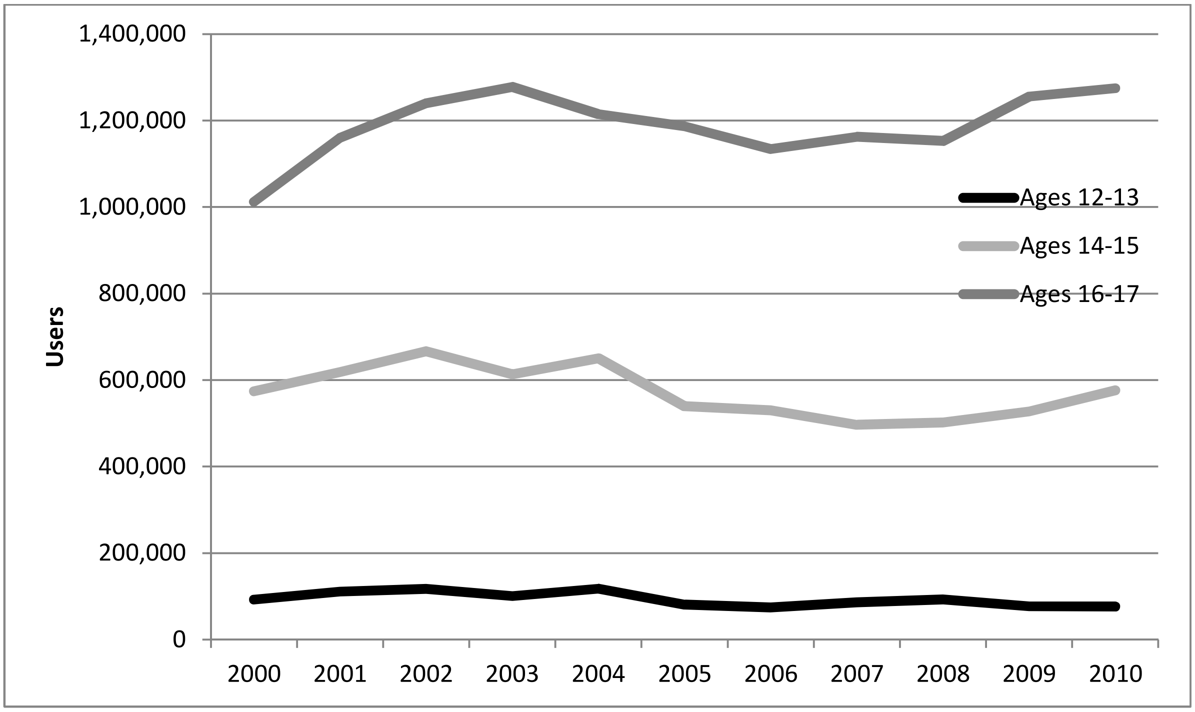
Figure 6.1, the recent increase in the number of users is mostly among adults who are not in treatment or the criminal justice system. Older teens account for most of the increase in youth marijuana users, as shown in Figure 6.2.

Figure 6.1: Marijuana Users by Subpopulations (NSDUH)



Note: The 2000 and 2001 estimates are not perfectly comparable to the later years, as a consequence of the differences between National Household Survey on Drug Abuse (NHSDA) (2001 and earlier) and NSDUH (2002 and later).

Figure 6.2: Number of Marijuana Users Among Youth (NSDUH)



Note: The 2000 and 2001 estimates are not perfectly comparable to the later years, as a consequence of the differences between NHSDA (2001 and earlier) and NSDUH (2002 and later).

Comparing Youth Estimates. Prevalence estimates from MTF and Youth Risk Behavior Surveillance System (YRBSS) are generally higher than NSDUH-based estimates, as explained in an appendix to the 2010 NSDUH Summary of National Findings. Figure 6.3 compares the percentage of past-30-day marijuana users among NSDUH respondents in grades 8, 10, and 12 with the same percentages estimated by the MTF survey. The differences are significant (and may be even more so if the intensity of use is considered). Moreover, the even greater gap between the YRBSS and NSDUH (Figure 6.4) highlights the need for caution in interpreting the NSDUH youth estimates and judging whether they are accurate. Based on methodological differences between the three surveys, it is conceivable that MTF and YRBSS overestimate as much as NSDUH underestimates youth marijuana use, though the authors of the MTF studies argue theirs are more likely to underestimate drug use among youth than overestimate it. However, this evidence suggests an adjustment to the youth prevalence is justified. Our approach replaced the youth prevalence rates derived from NSDUH data with prevalence rates derived from MTF.

Figure 6.3: Comparing NSDUH Youth Prevalence to MTF

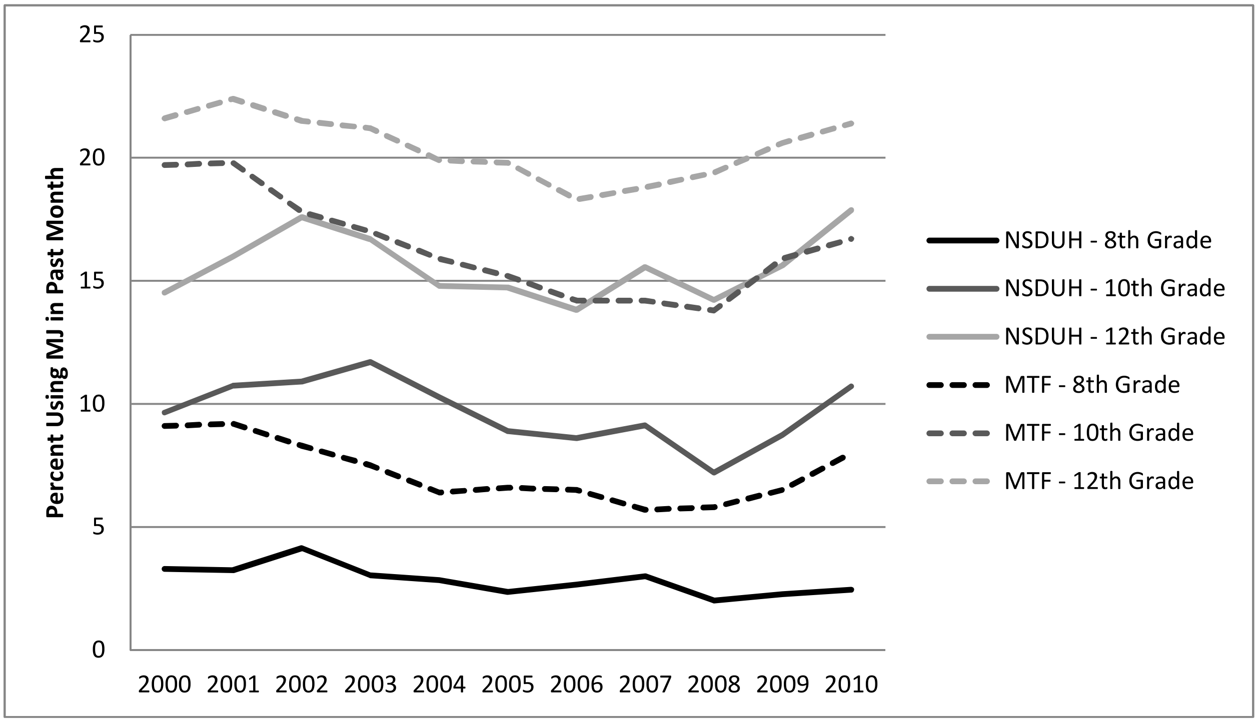
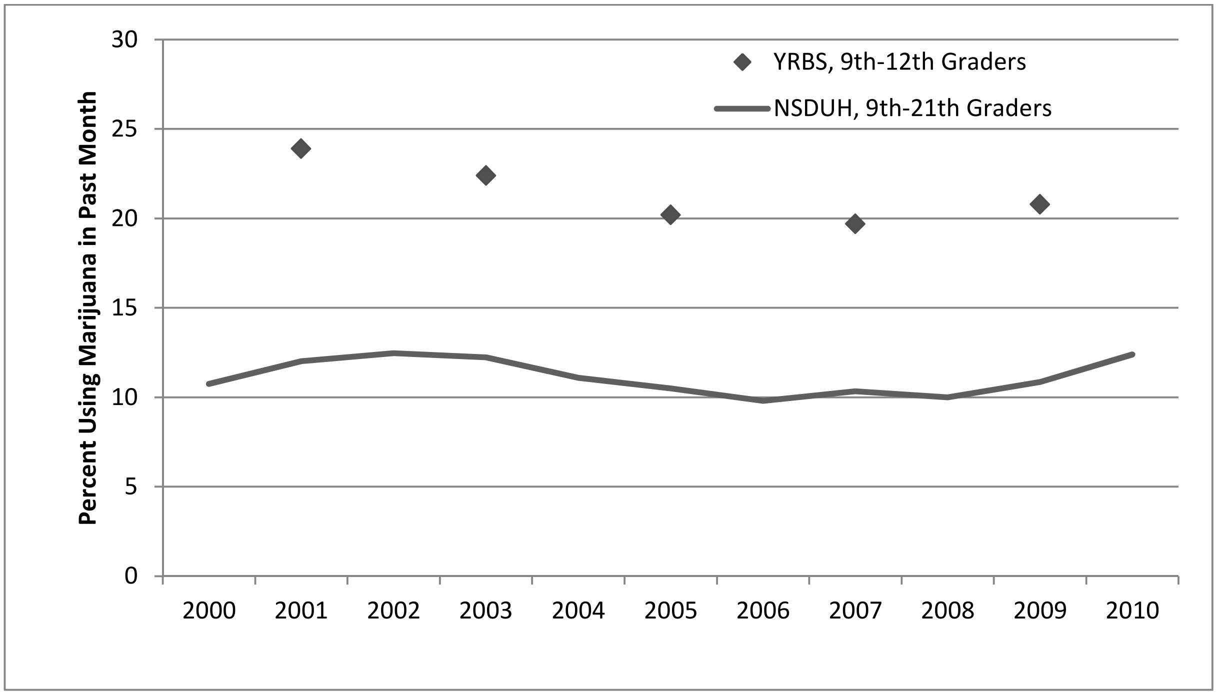


Figure 6.4: Comparing NSDUH Youth Prevalence to YRBS



Comparing NSDUH to ADAM and UCR Arrest Data. There is reason to believe that criminally involved marijuana users may underreport their marijuana use in NSDUH to a greater extent than do other marijuana users. In particular, the prevalence of marijuana use among those who report a past-year arrest in NSDUH varies between 20 percent and 30 percent over the 2000–2010 period. In ADAM II cities for years in which data are available, the percent reporting past-30-day marijuana use ranges from 30 percent to 60 percent, and approximately the same range applies to the percentage of people who test positive for marijuana use. The median across cities of each measure falls in the mid-40-percent range.

Hence, instead of applying the same adjustment for underreporting that we used for the general adult population, we increased the marijuana use rates of adults involved in the criminal justice system to those derived from ADAM data.

6.4. Grams of Marijuana Consumed Per Day

Our estimate of grams consumed per day is the product of joints consumed per day and weight (grams) per joint. We discuss each factor in turn**.**

Joints per Day. Both NSDUH and ADAM ask about days of use, but neither ask about how much was used nor how it was consumed. (NSDUH stopped inquiring about joints per day in 1995). The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) is a nationally representative repeated-panel survey of the noninstitutionalized U.S. population 18 years of age and older. It includes a number of questions about substance use, including "How Often Used Cannabis in the Last 12 Months" (S3BD5Q2C) and "Number Of Cannabis Joints Usually Smoked in a Day in the Last 12 Months" (S3BQ3). The first wave of NESARC was fielded in 2001/2002; the follow-up was fielded in 2004/2005. Table 6.3 presents the mean joints per day reported by wave for four categories of users, distinguished by past-month frequency of use. The follow-up cohort in 2004/2005 is not representative of the entire adult (18 and older) population. At follow-up the respondents are four years older, and so include no 18–21-year-olds. There may also be some attrition bias. So, we use the results from the original cohort (presented in the first column of the table) to inform the estimate of marijuana consumption.

Table 6.3: Mean Joints per Day from NESARC, By User Type and Age (95-percent confidence interval)

Type of Marijuana User	**(1) 2001/2002 Total**	**(2) 2001/2002 23+**	**(3) 2004/2005 Total**
MJ1: 20+ times month (Every day, Nearly every day)	3.87 (3.33-4.39)	3.04 (2.57-3.52)	2.73 (2.42-3.04)
MJ2: Less than 20, more than 3 (1-4 times week)	1.92 (1.63-2.20)	1.67 (1.39-1.95)	1.79 (1.55-2.02)
MJ3: 1 to 3 days months (1-3 times month)	1.68 (1.46-1.89)	1.59 (1.34-1.84)	1.34 (1.21-1.47)
MJ4: Less than 1 day per month (<12 times a year)	1.17 (1.12-1.22)	1.15 (1.10-1.20)	1.14 (1.09-1.20)
TOTAL	2.00 (1.86-2.15)	1.72 (1.58-1.85)	1.62 (1.53-1.71)

Grams per Joint.[20] There is no direct data on how much marijuana is in a typical joint, but the size of a joint can be derived from price data. The ADAM program asks arrestees detailed questions about the price they paid for marijuana at their last transaction. We estimated the grams per joint based on information from 1,613 arrestees in the 2000–2003 ADAM survey who reported purchasing either 1 g of marijuana or one joint of marijuana at the time of their last purchase. We did not focus on other quantities because we preferred to minimize complications that could arise from possible differences in quantity discounts across ADAM jurisdictions. Table 6.4 presents the descriptive statistics about price for these two types of purchases.

Table 6.4: Nominal Price Paid for One Joint or One Gram by Arrestees, ADAM 2000–2003

	1 joint (N=527)	**1 gram (N=1,086)**
Mean	$8.00	$16.74
Median	5	20
25th percentile	5	10
75th percentile	10	20
90th percentile	10	20

Note: Outliers were not removed before calculating these summary statistics.

We constructed a statistical model that estimated the average price per gram for each location from the gram data, adjusting for a year trend. Since some locations had very few observations, we used a random effect to stabilize the estimates and shrink the price per gram estimates from each location toward one another. Our model described each of the joint weights as coming from a common distribution with a shared mean and variance. Then we modeled the price for each joint, as the unobserved joint weight multiplied by the price per gram for that location, multiplied by the year-effect plus noise. We simultaneously computed Bayes estimators of all of the model parameters: each location's price per gram, the year trend, the weights of each of the joints, and their common mean and variance.

More technically, each reported price for a gram of marijuana was modeled as coming from a $N(\beta_j\alpha_k, \sigma^2)$ distribution, where β_j is the average price per gram in location j and α_k is a coefficient to adjust for price changes in year k. The location was treated as a random effect so that $\beta_j \sim N(\beta_0, \tau^2)$. A nonparametric distribution for the β_j produced nearly identical results. For identifiability, α_{2000} was set equal to 1.0 so that for future years the values of α_k represent yearly changes in price relative to 2000. The weight of joint i, w_i, was modeled as coming from a $N(\mu, \sigma_w^2)$. The observed price paid for the joint was modeled

[20] This section is largely reproduced from Kilmer et al. (2010b).

as $N(\beta_j \alpha_k w_i, \sigma^2 w_i^2)$, the same distribution used for the price per gram but rescaled by w_i. All hyper-parameters were given standard uninformative priors. The prior for τ was Uniform(0,10), the upper bound of which is well above plausible values for τ. We reported the posterior mean and posterior 95-percent interval for μ.

This generates an estimate of 0.46 grams/joint with a 95-percent interval of (0.43, 0.50). To address outliers, we dropped those observations with price values greater than the 99th percentile separately for 1 joint and 1 gram, joints reportedly costing more than $20, and grams reportedly costing more than $50. Note that most of these outliers were on the order of $100–$500, far beyond reasonable values and almost certainly errors in reporting or recording.

There are two minor biases in this analysis. Both will lower the estimated price per gram of a joint and, thus, tend to overestimate the weight of a joint. First, the average "one-gram" purchase on the street weighs slightly less than 1.0 gram; drug dealers tend to err on the light side when preparing sales. Second, the price per gram for a gram purchase is estimated at a (marginally) higher “market level" than is the price per gram for a single joint. Price as a function of weight is often modeled as a power function with an exponent in the vicinity of 0.8. In such circumstances, the price per gram will tend to be lower for the larger quantity (e.g., with an exponent of 0.8, doubling the quantity would lead to a price per gram that is 13 percent lower).

Thus, for an estimate of the size of a joint, we choose 0.43 g/joint, the lower bound of the 95-percent confidence interval.

6.5. Marijuana Price

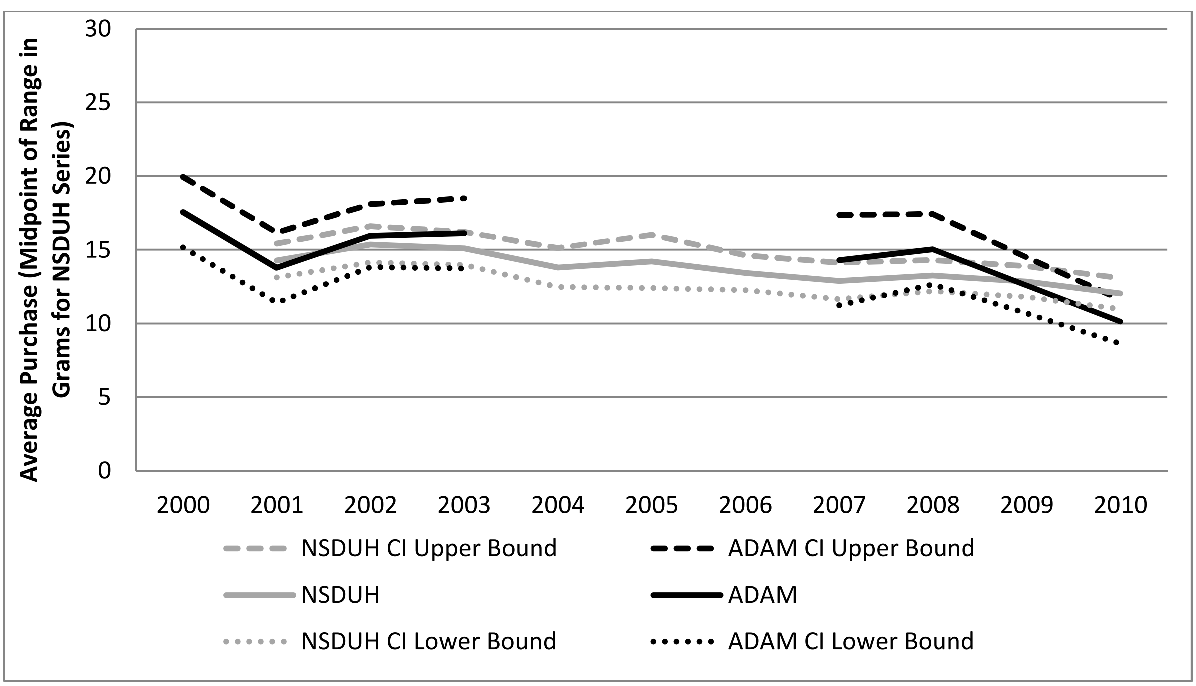
Since 2001, NSDUH has included a series of questions about marijuana acquisition patterns that are given to individuals who report use of marijuana in the past year. Respondents were allowed to answer questions on the prices paid using whatever form of marijuana they purchased (loose or as joints).[21] Those who bought in loose form were then asked if the unit of purchase was in grams, ounces, or pounds, and how many of that unit they purchased. If they responded “grams,” they were asked whether the quantity of their purchase was between one and five grams, between five and ten grams, or ten or more grams. If they responded ounces, they were asked if the quantity was between 1/8 and 1/4 ounce, 1/4 and 1/3 ounce, 1/3 and 1/2 ounce, 1/2 and one ounce, one to five ounces, five to ten

[21] In ADAM, respondents are even allowed to answer in terms of pounds. Price data in ADAM were reported as continuous variables, however, so no adjustment needed to be made.

ounces, or ten to 16 ounces. If they responded pounds, they were asked if the quantity was any of five categories, from one up to two, two up to three, and continuing up to five or more pounds. For estimation convenience, non-gram purchase quantities were translated to gram-equivalents. They were also asked the amount they paid for their last purchase, choosing one of the following: less than $5.00, $5.00 to $10.99, $11.00 to $20.99, $21.00 to $50.99, $51.00 to $100.99, $151.00 to $200.99, $301.00 to $500.99, $501.00 to $1000.99, or more than $1000.99. To facilitate analyses, the midpoint of each range (and the minimum of the top-coded option) was used as a point estimate for the total transaction cost and amount in all of our analyses.

Among reported purchases, the average amount of loose marijuana acquired has remained near one-half ounce over the last decade, which may be in part because of the categorical way in which the data was obtained; however, a similar trend is seen in the continuous ADAM series (see Figure 6.5). But shifts in the number of people reporting amounts in a given range are clearly occurring, as there is visual evidence of a slight downward trend in NSDUH. Importantly, we see the same downward trend in average amount of loose marijuana purchased in the ADAM surveys, but the trend is even more pronounced.

Figure 6.5: Average Marijuana Purchase Amounts over Time (Limited to purchases less than or equal to 5 oz)



As prior research has identified individual and situational factors that are correlated with prices paid for marijuana (Caulkins and Pacula, 2006; Pacula et al., 2010), we did not simply average reported price paid for a specific amount from those reporting purchases in the NSDUH survey data. Instead, we employed a prediction model that allowed us to control for a variety of factors that have been shown to be associated with the actual price paid. Specifically we estimate the following regression for each year *t*:

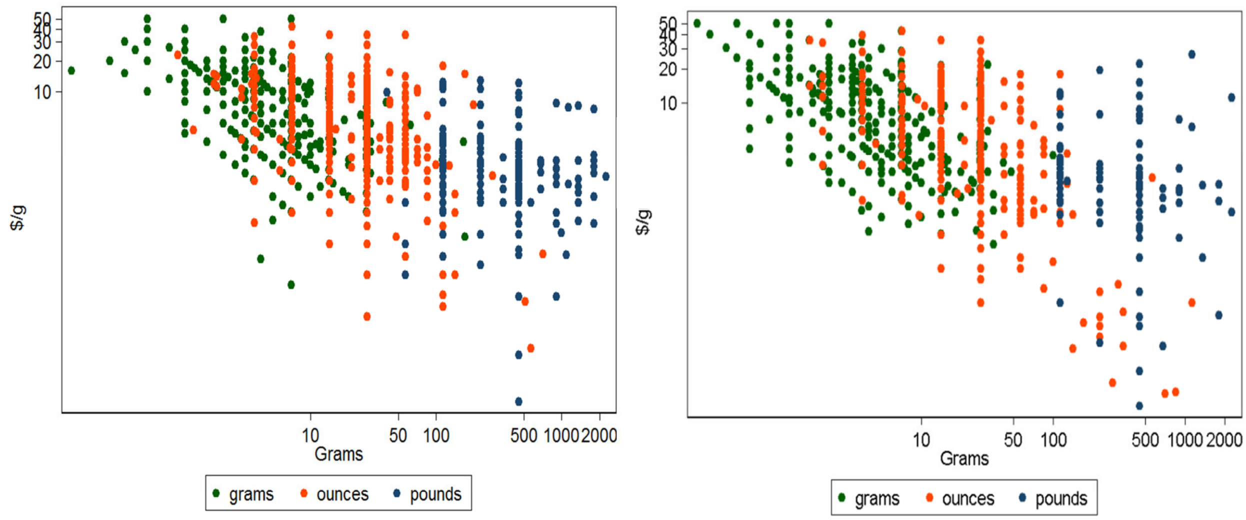
$$ln(price\ per\ gram_{it}) = \alpha + \delta\ (AMT_{it}) + \beta\ X_{it} + \lambda Z_{it} + \gamma(QTR_i) + \varepsilon_{it} \qquad (6.1)$$

Where *price per gram*$_{it}$ is the unadjusted (raw) price per gram paid constructed by simply taking total purchase amount paid and dividing by the amount purchased converted into grams for individual *i* in year *t*, *AMT* is the amount actually purchased in that transaction specified in grams, X_{it} is a vector of characteristics of the individual *i* who engaged in the exchange during time *t*, including age, gender, race/ethnicity, education, and income, and Z_{it} is a vector of exchange characteristics experienced by individual *i* at time *t*, including whether the purchase was made indoors vs outdoors, in a public setting, in the respondent's neighborhood, from a stranger, friend or family member and from his/her regular seller.[22] QTR_i represents the calendar quarter of the year in which the survey was conducted to account for possible factors associated with the time of year (either as seasonality in prices or other effects associated with time-varying factors not captured elsewhere). The regression error term is reflected by the term ε_{it}.

There are three important data-cleaning steps that occur before we estimate the model. First, we removed individuals reporting they were dealers or stating that they resold some or all of their last purchase acquired. As individuals who plan on selling marijuana generally buy more than they intend to consume, they benefit from lower prices associated with buying larger quantities (i.e. "bulk" or "quantity" discounts) than the typical end user. Caulkins and Pacula (2006) demonstrate that quantity discounts exist in the household data. Additional evidence of quantity discounts in the general marijuana markets comes from the ADAM data (see Figure 6.6). Even among arrestees, we see that those reporting purchasing larger amounts of the drug pay a lower price per gram than arrestees who purchase smaller (single consumption) amounts.

[22] Estimated series based on ADAM do not include income or relationship to seller.

Figure 6.6: Relationship Between Price/Gram and Size of Purchase (ADAM data)



ADAM-1 ADAM-2

This implies that, if we retain individuals who sell part of their purchase, it is possible we are underestimating the price that the final end user paid for the marijuana consumed. Furthermore, we risk double-counting the price paid for that gram, as the price that the seller had to pay to acquire the purchase is embedded in the end-price. Thus, we exclude from our analysis individuals who report in a subsequent question that they sold some or all of their purchase.[23]

The second major data-cleaning step is also done to reduce the influence of intermediate purchases. Specifically, we exclude purchases greater than five ounces, because such large volumes are unlikely to all be for self-use.[24] If large, expensive purchases are indeed completely for self-use, their inclusion presents a risk of biasing consumption estimates up due to random incidence. This analysis and others preceding it based consumption estimates on the number of times each respondent reported using and purchasing in the past month, and the respondents' single most recent transaction. This may overstate

---

[23] ADAM-based analyses exclude participants who report consuming less than 50 percent of the quantity they purchased, as no specific question about intent to resell is given. This is not preferred, as amounts given away are still presumed to be consumed and have a known price—that which was paid for the whole amount in the initial sale. However, this factor was only a statistically significant predictor of price in the ADAM estimates in one of the eight years estimated.

[24] Additionally, the analysis was run for purchase of one ounce or less as a more conservative test of price sensitivity. Prices were estimated to be systematically higher in this smaller subset, averaging a 30-cent premium in NSDUH and $1.04 in ADAM to the sets of five ounces or less.

consumption for a person who buys in bulk occasionally as was discussed in Section 6.1. By excluding purchases of greater than five ounces we mitigated this risk of overestimating consumption.

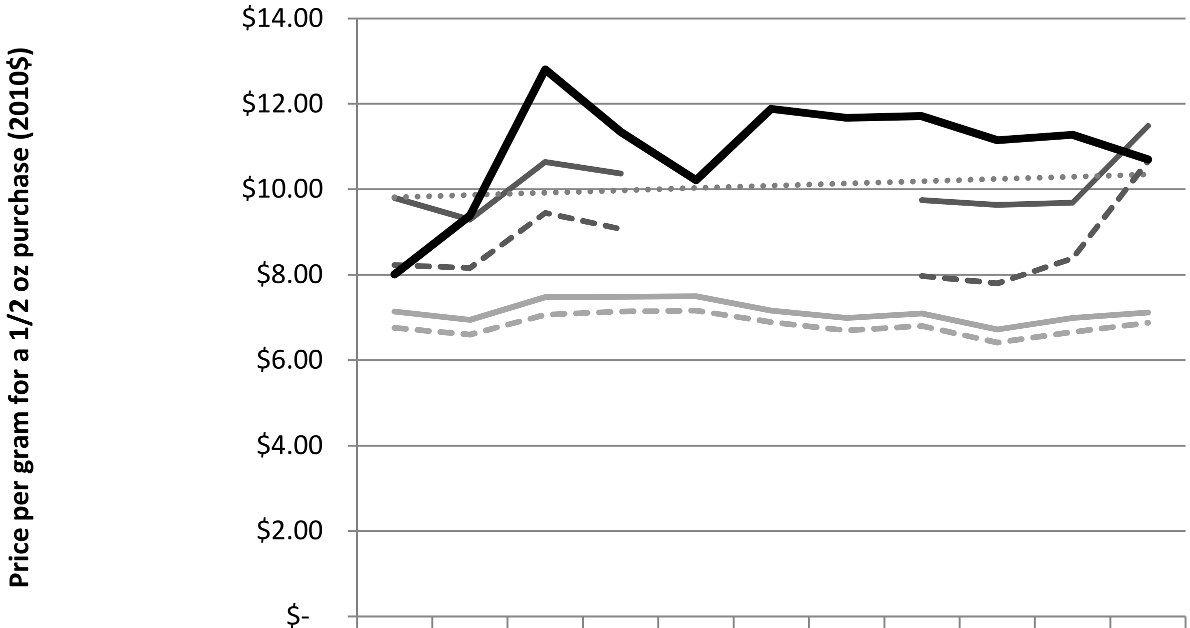
Finally, we excluded extreme price values from the analysis. As noted in the previous version of this report, "very large amounts of marijuana were sometimes purchased for very low prices, and small amounts of marijuana were purchased for very high prices" (ONDCP, 2012). To minimize the risk of erratic participant-reported data, we predict log-prices using log-quantities and exclude observations with prices more than two standard deviations away from their expected value. The resulting data from ADAM are plotted in Figure 6.6.

We estimate dollar per gram using equation (6.1) for each year from 2001 to 2010 individually. Coefficient estimates from these models are retained so that they can be used to generate a predicted average price per half-ounce controlling for the different factors that influence price. To implement this, we use the individual values of the controls in our model, replace the AMT variable with 14.175 grams (the amount equivalent to a half-ounce) and multiply these individual values by the estimated coefficients that were retained from estimating equation (6.1). This provides a predicted log price per half-ounce of marijuana, which was the most common amount purchased in our ≤ 1- and ≤ 5-ounce sample (as shown in Figure 6.7).

As equation (6.1) is log-linear, a retransformation bias results if we simply exponentiate the predicted value from the model. We therefore corrected for the retransformation bias using the Duan smearing estimate, done automatically with the *levpredict* command in Stata 12 once the AMT variable is set to its desired value (here, one half-ounce).

Our estimated price series using the NSDUH data and following a similar strategy with the ADAM data are presented in Figure 6.7, along with the most recent version of the Fries et al. data (which was used in the previous edition of this report). Discounting the rapid increase over 2000–2002 in the series developed by Fries et al., the NSDUH price series follows a path similar to, though more stable than, the Fries et al. series, but at a a roughly 35-percent discount on average. Due to missing data in the years when ADAM surveys were not conducted, it is impossible to infer much from the ADAM data safely, though they also suggest a downward trend in price through the mid-2000s, albeit with an anomalous spike in 2010.

Figure 6.7: Predicted Price per Gram Evaluated at a 0.5-Ounce Purchase Quantity



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
NSDUH Obs ≤ 1 oz	7.14	6.95	7.48	7.49	7.50	7.16	6.98	7.09	6.72	6.99	7.11
ADAM Obs ≤ 1 oz	9.80	9.29	10.64	10.37				9.75	9.64	9.69	11.49
Linear Trend of ADAM	9.81	9.87	9.92	9.98	10.03	10.08	10.14	10.19	10.24	10.30	10.35
NSDUH Obs ≤ 5 oz	6.76	6.60	7.06	7.14	7.16	6.89	6.70	6.80	6.41	6.66	6.87
ADAM Obs ≤ 5 oz	8.23	8.16	9.45	9.07				7.97	7.80	8.38	10.66
Fries et al.	8.01	9.38	12.81	11.35	10.21	11.88	11.67	11.71	11.15	11.27	10.70

Section 7. Sinsemilla Consumption in the United States

It has become standard practice over the last 20 years to work with purity-adjusted prices for cocaine, heroin, and methamphetamine, not just the "raw" price per gram, unadjusted for purity. Marijuana is more complicated, because it contains many different psychoactive chemicals, but tetrahydrocannabinol (THC) is the most important, so there is likewise an argument for working with THC potency-adjusted prices for marijuana. That has not been standard practice in the past, but we attempted to factor potency trends into our estimates here.

Potency affects the amount consumed per unit of intoxication. Roughly speaking, one-third of a gram of sinsemilla containing 15 percent THC can provide as much intoxication as one gram of commercial-grade marijuana that contains 5 percent THC. So when potency goes up, the weight of marijuana consumed per day of use may go down. Since potencies have in fact been trending upward, and more potent variants are more expensive per unit weight, failing to adjust for potency trends might lead to an artificial increase in estimates of marijuana spending.

There is no single data source that tracks how much sinsemilla is being consumed in the United States over time. The national drug use surveys (e.g., NSDUH, MTF) do not ask respondents about the type or potency of marijuana, and even if they did it is not clear that users could or would provide reliable information. Thus, researchers are forced to combine insights from multiple sources to approximate levels and trends.

A number of indicators support the idea that the share of marijuana consumed that was sinsemilla increased in the United States from 2000 to 2010; however, neither the magnitude of the increase nor the current levels are known. Based on our analyses of market transactions from NSDUH and other sources, we assume that sinsemilla's "market share" increased from 10 percent in 2000 to 20 percent in 2010 (We use round figures to avoid suggesting that these figures are precise). The implications of these figures for total marijuana consumption and spending depend on one's beliefs about whether THC consumption per typical use day increased over the decade or remained constant. If weight consumed per day stayed constant, then THC consumed per day would have increased; if THC consumed per day stayed constant, then weight consumed per day would have fallen—necessitating an adjustment to get spending estimates correct. This issue of different types of marijuana could become even more important in the future, if the recent trend toward proliferation of alternative forms of use continues (edibles, hash oil, “dabbing”, etc.).

7.1. Potency Trends by Type of Marijuana

There is no nationally representative sample of marijuana samples’ potencies. Indeed, there are no routinely reported time series describing potency or the market shares of different types of marijuana. Probably the best source of marijuana potency information is the University of Mississippi Marijuana Potency Monitoring Program (MPMP) data set. As far as we know, the MPMP data are highly reliable in terms of their chemical analyses; we do not know of any reason to doubt that when a particular sample is described as being 12 percent THC that it was not, in fact, 12 percent THC. However, the MPMP is not designed or intended to be representative of the overall market; the MPMP samples are the product of enforcement actions, so the sampling is purposeful and varies with enforcement tactics and priorities.[25] There may also be variation around the country and/or over time in policies governing which enforcement samples get sent for analysis.

If the universe of marijuana samples were homogenous, this would not be a problem. However, marijuana is not a single drug; indeed, technically it is not a drug at all. Literally marijuana is a slang term for various products of a type of plant that includes scores of different psychoactive compounds, whereas properly speaking a drug is one a specific chemical. That distinction may become important someday, with prices and potencies handled differently for different types of marijuana characterized by different ratios of the main psychoactive chemicals. However, even when thinking only in terms of THC content, as we do here, there is still such an enormous range in potencies that it can be misleading to lump together forms as different as ditchweed and sinsemilla. Potency also varies from one part of a plant to another, creating distinctions between leaves and buds (or “flowering tops”), as well as between seeded and seedless varieties. Likewise, THC content can vary substantially over the plant’s life, making it important to distinguish between mature and immature samples.

Our main conclusion is that over the time period of interest, the average potency of marijuana consumed in the United States has increased. There have been increases in potency of most types of marijuana (where sinsemilla buds collected during DEA operations might be considered to be one “type”

[25] From ONDCP’s Michael Cala (personal communication): “Different DEA labs send different proportions of specimens. All labs send specimens if over 75 grams, but two labs along Southwest Border only sent a subset of those.”

in this context). In addition, there is evidence suggesting that the "market shares" of the different types of marijuana have shifted. Notably, the relative numbers of different types of marijuana in the MPMP data have changed. In particular, the proportion of samples that are sinsemilla as opposed to general marijuana (which would include what we call "commercial grade") increased dramatically, from less than 10 percent before 2001 to almost 50 percent today. Lumping all types of marijuana together in the MPMP data and asking what has happened to the average THC of that polyglot mix would yield the answer of an even bigger increase in apparent THC content than one sees within any particular type.

There is no good reason to assume that the changes in relative proportions in the MPMP database matches the changes in the market generally; there could well have been changes in the enforcement priorities and/or the process generating UMISS samples, not just changes in the market. Hence, it is important to look at trends in MPMP potency data by type of marijuana, not in aggregate. In doing so, we restricted analysis to the following observations:

- TYPE = C (meaning cannabis as opposed to hashish or hash oil)
- MATURE = M (there are systematic differences between mature and immature samples' potencies)
- NUM_ANAL being a positive integer (no 0 or missing values to avoid averaging in a recorded potency of 0 that really means a missing observation)
- DESC = SM (sinsemilla) or MH (regular marijuana), meaning exclude ditchweed and Thai sticks
- CANSAMP = BD (Bud), KB (Kilobrick), LF (leaf), or LS (loose leaf), meaning exclude CANSAMP = unknown

We looked (separately) at observations stemming from DEA samples (CLASS = PM) and state eradication program samples (CLASS = ST), but focused on the DEA samples, which are more numerous.

As noted, the basic story is modest upward trends in potency within type of marijuana, but suggestions of substantial changes in composition (e.g., proportion of observations that are sinsemilla has increased). That is problematic because composition effects are most vulnerable to the data being enforcement driven, rather than drawn from a representative (let alone random) sample of the market. If DEA decided to increase the priority given to enforcement against domestic sinsemilla production (e.g., indoor grows) relative to other forms of marijuana, then we would observe a composition effect in the MPMP data even if there were none in the actual market. Figures 7.1 and 7.2 give a sense of the magnitude of these potential composition effects.

Figure 7.1: State Series Showing Potency Trends Over Time

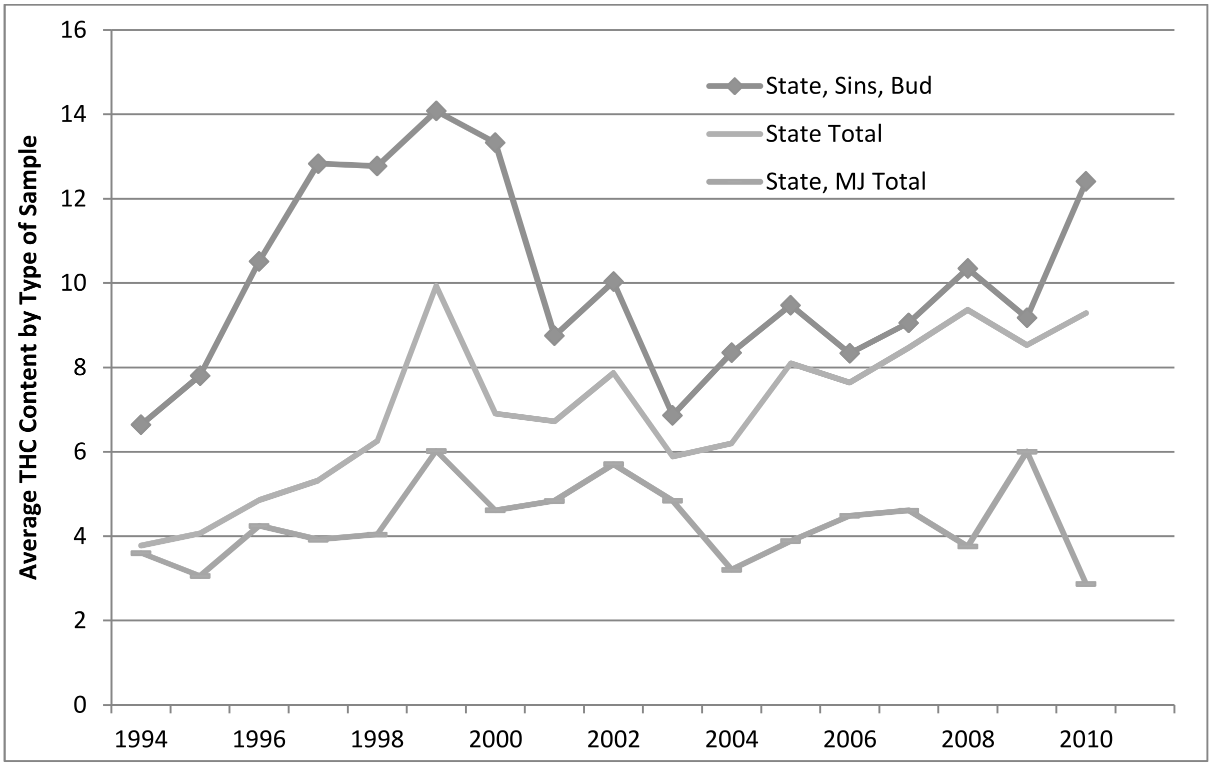
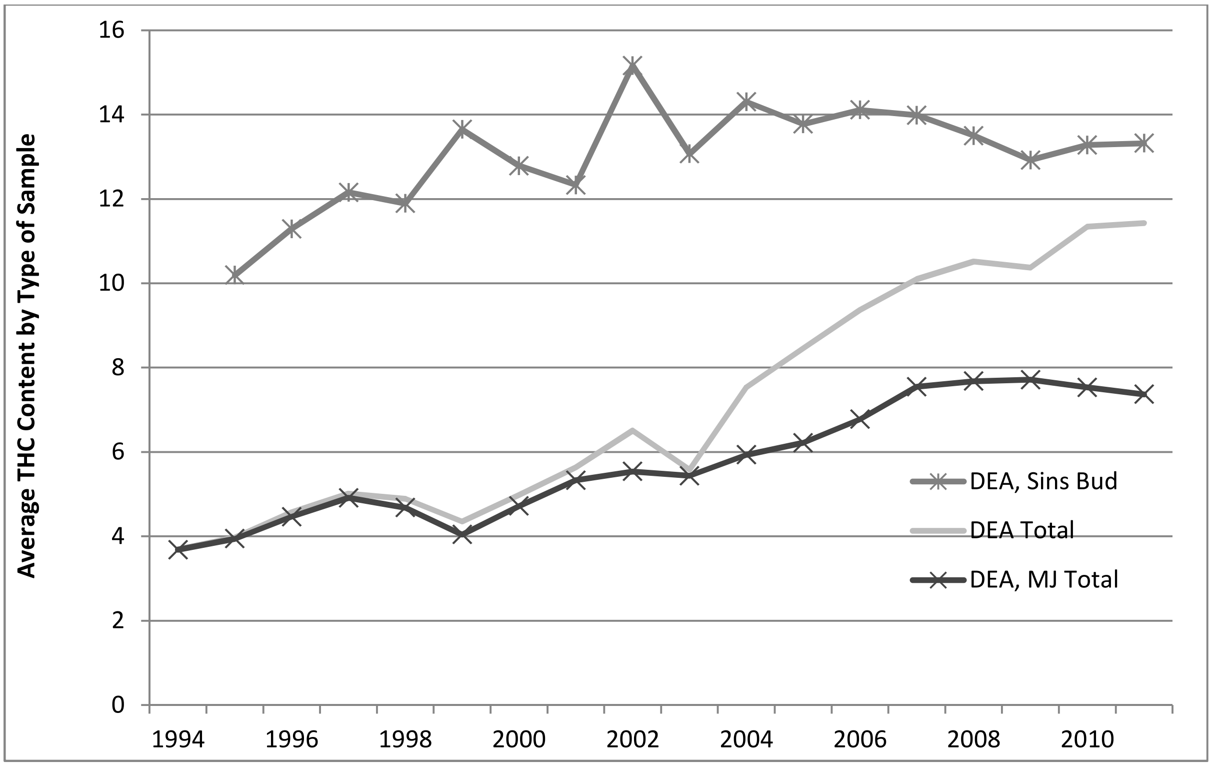


Figure 7.2: DEA Series Also Suggest Composition Effect



7.2. Evidence about Composition of Marijuana Consumption[26]

Again, there are no data series that report directly what proportions of marijuana consumed were of one type or another. However, since sinsemilla prices per unit weight are typically two to three times higher than commercial grade prices (Kilmer et al., 2010b, Appendix), the prices that users report paying for a given weight provide a proxy for the type of marijuana that was purchased. For any given cut-off price that lies between the typical prices for commercial grade and sinsemilla, one could guess that most purchases below that price were of commercial grade marijuana, and most with higher prices were sinsemilla. This proxy measure can be used to estimate market composition using data from any survey that asks respondents to describe the price paid and quantity obtained, as do both NSDUH and ADAM.

This exercise is not precise. There are not just two distinct types of marijuana or two marijuana prices. Rather, there is a continuum. So we varied the assumed cut-off price as a sensitivity analysis. Likewise the distribution of prices paid for the last purchase may differ from the distribution of prices paid for

[26] The setup for this subsection is from the Appendix of Kilmer et al. (2010b).

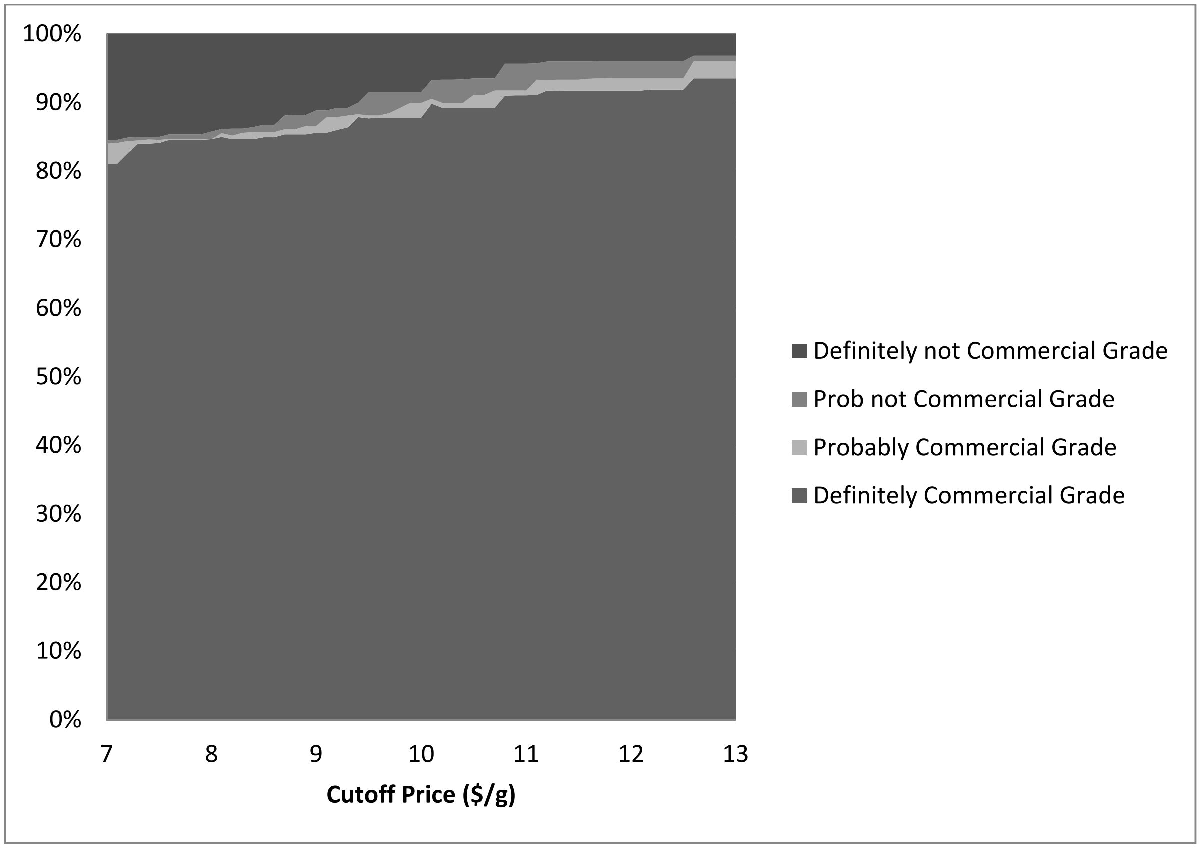
consumption overall, and survey responses are often given as ranges, not specific numbers. Nevertheless, the overall finding comes through rather forcefully. Marijuana that is as expensive as prices typically cited for sinsemilla accounts for a rather modest share of all self-reported purchases. That was true in 2000 and remained true in 2010, albeit to a lesser extent.

Since NSDUH captures the majority of marijuana consumption, NSDUH data are of the greatest interest in this regard. The details of the method for NSDUH were developed and described by Kilmer et al. (2010b, Appendix). The important points are that they apply for respondents who account for about 95 percent of past-year days of use and are reasonably decisive for respondents who account for about 60 percent of that use. In particular, they can be applied to respondents who describe their most recent purchase in terms of grams or ounces, but not pounds or joints. The indecisiveness primarily concerns those who describe their most recent purchase as being of one to five grams. If someone reported spending \$11–\$20.99 to buy one to five grams, the per-gram price could be as high as \$20.99 (which would clearly indicate high-grade marijuana) or as low as \$11/5 = \$2.20 (which is a very low price, strongly suggesting that the type purchased was not sinsemilla). Someone describing their most recent purchase in those terms would be impossible to categorize.

Because some individuals are hard to categorize, the method divides respondents into four bins: (1) those whose purchases were definitely not made at commercial-grade prices, (2) those whose purchases were probably not commercial grade, (3) those who probably purchased commercial grade, and (4) those whose purchases were definitely made at prices associated with commercial-grade marijuana.

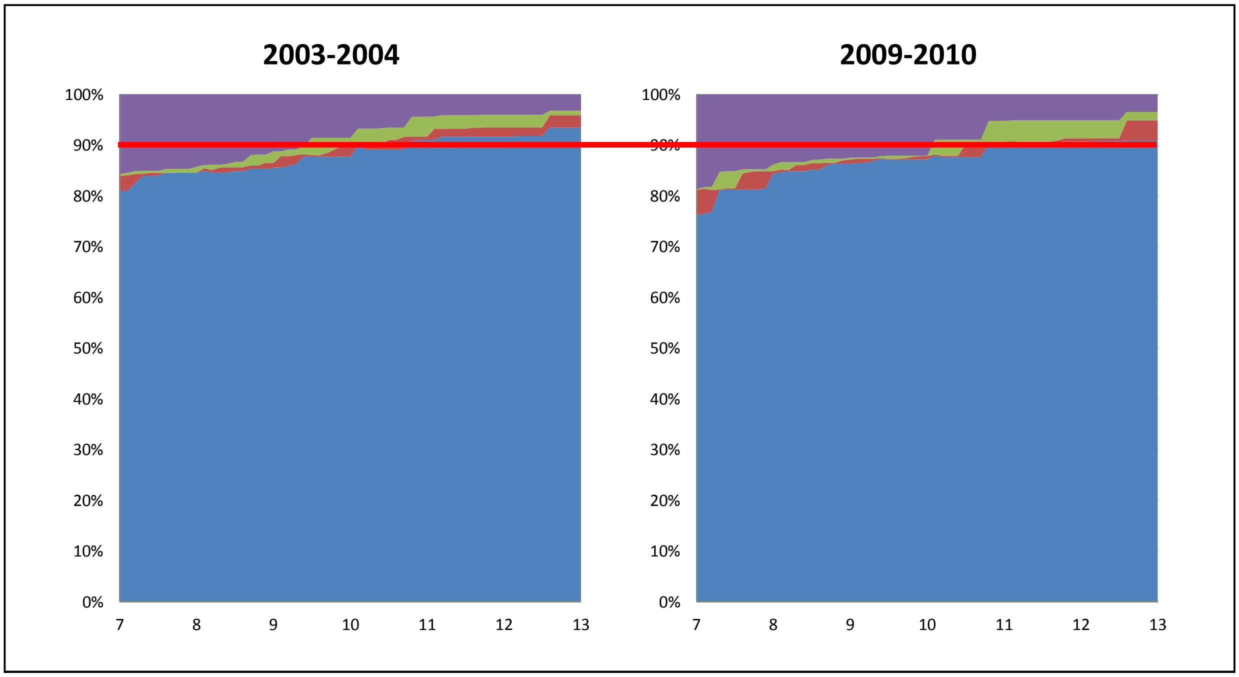
Consider, first, \$10 per gram as a lower bound on the price of marijuana that is not commercial grade. Using that cut-off for purchase data from the 2003–2004 NSDUH surveys produces the results in Figure 6.3, in which 85 percent of respondents' purchases were definitely commercial grade vs. 10 percent that were definitely not—meaning, in this two-tiered simplification, they were definitely sinsemilla. Figure 7.3 also shows how these proportions change if we vary the cut-off price, considering values anywhere within the \$7–\$13 per-gram range, not just \$10 per gram. Naturally, the higher the cut-off, the greater the proportion of respondents' purchases that are classified as being commercial grade. Even with a conservative threshold of \$7 per gram of sinsemilla, 81 percent of those purchases had prices that were definitely below that cut-off in 2003–2004.

Figure 7.3: Deduced Grade of Marijuana (10+ grams purchased, 2003–2004)



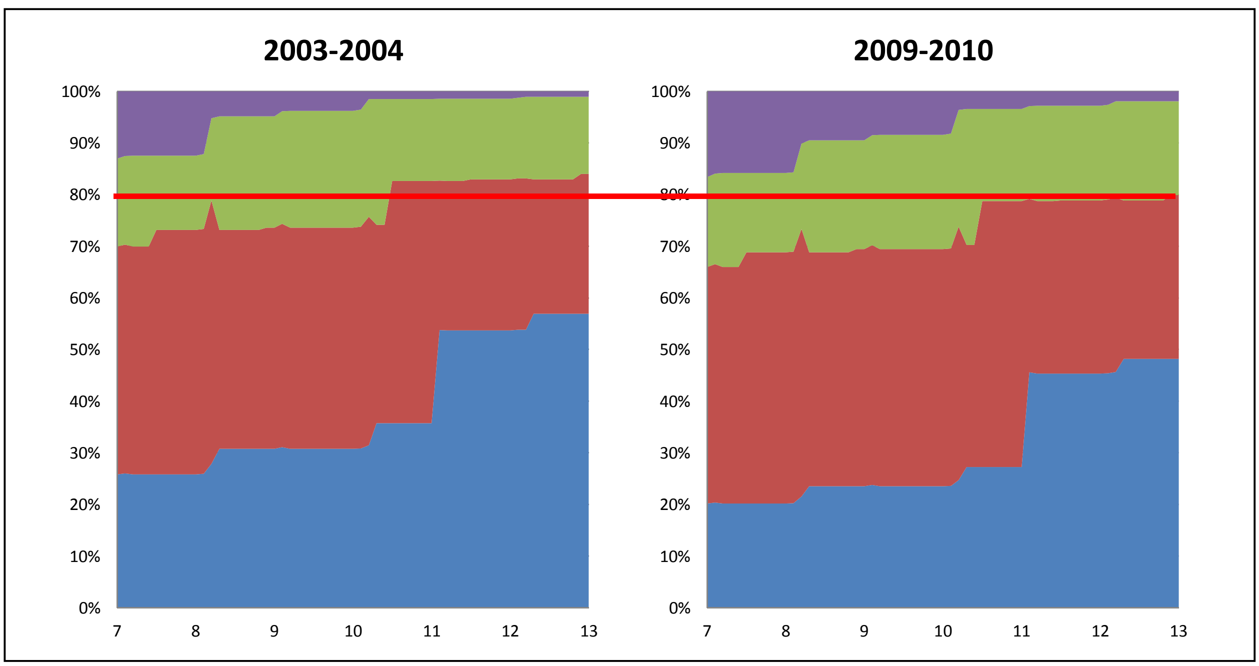
To help assess whether there have been changes in the share of marijuana purchases that are commercial grade, Figure 7.4 displays the 2003–2004 chart (same as Figure 7.3) next to a similar chart based on data from 2009–2010. (For this comparison, we use 2003–2004 instead of 200–2002 to avoid confounding time trends with the changes to the NSDUH instrument in 2002.) The reduction in the blue "definitely commercial grade" is consistent with an increase in sinsemilla market. We see a similar story in Figure 7.5, which is based on purchases of one to five grams. Over time, not only does the blue area get smaller, but the purple "definitely not commercial grade" area gets larger.

Figure 7.4: Comparison of Deduced Grade of Marijuana, 10 or More Grams Purchased



Note: Blue = definitely commercial grade; red = probably commercial grade; green = probably not commercial grade; purple = definitely not commercial grade.

Figure 7.5: Comparison of Deduced Grade of Marijuana, 1 to 5 Grams Purchased



Note: Blue = definitely commercial grade; red = probably commercial grade; green = probably not commercial grade; purple = definitely not commercial grade.

In like manner, Table 7.1 presents the share of marijuana purchases that are sinsemilla for the entire period 2001–2010, based on various price thresholds. For our purposes, we consider a purchase to be sinsemilla if it is "definitely not commercial grade" or "probably not commercial grade." Previous price analyses (Kilmer et al., 2010; Bond & Caulkins, 2012) suggest that $200–$250 is a reasonable threshold for determining whether someone purchased an ounce of sinsemilla or an ounce of commercial-grade marijuana circa 2010. However, this price threshold may not have been stable over the decade.

Table 7.1: Share of Marijuana Purchased that is Sinsemilla, 2000–2010

Per Gram	**$7**	**$8**	**$9**	**$10**	**$11**
Per Ounce	**$198**	**$227**	**$255**	**$284**	**$312**
2001/2	17.3%	16.5%	9.0%	8.0%	5.5%
2003/4	20.4%	19.8%	12.8%	11.0%	7.0%
2005/6	24.6%	23.8%	13.8%	12.7%	8.5%
2007/8	19.7%	19.0%	11.9%	10.7%	7.2%
2009/10	22.0%	21.4%	14.8%	13.2%	9.0%

There are no published national price series for sinsemilla in particular, but to get some inkling as to possible trends, we can examine two imperfect sources: wholesale price information collected by a medical marijuana testing lab, and self-reported purchase information submitted to High Times magazine.

SteepHill Labs in Oakland, California, (SteepHill, 2010) provides testing services for a number of clients, many of which are in Northern California, an area known for sinsemilla production and export. The lab surveyed growers, brokers, and dispensaries about wholesale prices for high-grade "medical cannabis" (which is typically sinsemilla) sold in one- or two-pound increments from the last half of the decade. Their data suggests that per-pound wholesale pieces decreased by roughly 15 percent to 30 percent from 2005 to 2010, depending on whether the marijuana was grown indoors and in which part of the state. Stronger statements are not possible given the nature and limitations of the data. Note also that we are interested in retail prices, not wholesale prices.

Second, a recent study examined the effect of medical marijuana laws on marijuana prices (Anderson, Hanson, & Rees, 2013). For their analysis, Anderson et al. collapse the price information reported by purchasers to High Times (n>8,000; 1990-2011) into median values for each state and year. They do this for both “high-quality” and “low-quality” marijuana.[27] Averaging these median values over the period from 1990 to 2011 generates a high-quality price ($313, in 2000 dollars) that is 2.4 times larger than the low-quality price ($129, in 2000 dollars).[28] Converting these mean values to 2010 dollars generates figures of $396 and $163, respectively.

Anderson et al. find that states with allowances for medical marijuana saw a 26-percent decrease in the price of high-quality marijuana.[29] The effect size fell to 10 percent once state-specific linear time trends were added to the model. While we cannot back out the effect for specific states or time periods, this average decrease in the retail price of “high-quality” marijuana is consistent with a decrease in sinsemilla prices over the 2000s.

These bits of circumstantial evidence suggest we probably should not apply the same “commercial-grade” price threshold to the entire decade, but they do not provide clear guidance about what the different thresholds should be for each year. Since we believe the threshold has decreased over time, we choose a value suggested by the *right*-side of Table 7.4 for 2001/2002 and a value consistent with the *left*-side of the table for 2009/2010. In particular, we assume the values to be 10 percent for 2001–2002 and 20 percent for 2009–2010. We use round numbers to highlight the lack of precision, and the implicit assumption that the share of marijuana consumed that was sinsemilla doubled over the decade does not seem unreasonable, especially given the compositional changes in MPMP samples.

7.4. Overall Potency Trends

We choose to focus on two series: DEA sinsemilla bud observations as our measure of trends at the high end of the market, and DEA kilobrick observations as our measure of trends in potency in the bulk of the market that is not sinsemilla (i.e., the commercial-grade market). Generally, there are two categories of data sources for potency information: seizure and undercover buy information from law enforcement agencies and information volunteered by survey respondents, and both categories have advantages and caveats. Enforcement-driven databases alone are difficult to derive much information from regarding

[27] Anderson et al.: “Distinguished between high-quality (a category that included Californian and Hawaiian sinsemilla) and low-quality (a category that included commercial grade Colombian and Mexican weed) marijuana.”
[28] The annual estimates are not published in the paper.
[29] The phrase “high quality” is often used interchangeably with “high potency” or “high THC”; however, it is possible to obtain marijuana with a lot of THC that may not be considered good quality (e.g., could have a lot of pesticides, could be dry). Conversely, there are now high-CBD low-THC strains sold in medical dispensaries that could be considered high quality for certain purposes, even if they are not highly intoxicating.

composition. It would be risky to assume that the proportion of STRIDE observations that are cocaine, heroin, or meth can be used to estimate the relative sizes of the markets for those three drugs in the country because enforcement prioritization directly affects those shares.

At present, the questions on national surveys asking about marijuana use do not differentiate between types of marijuana. So our quantity estimates are built up from reported days of use of some marijuana product, not from separate estimates for different types of marijuana. We must derive differences from price and quantity information provided by respondents.

While we believe there is a trend toward higher potency over time, the trends we estimate are an effort to smooth idiosyncrasies in the NSDUH-derived series. We are not confident enough in the higher frequency movements of the series to want to let them introduce jumps and fluctuations in the final series.

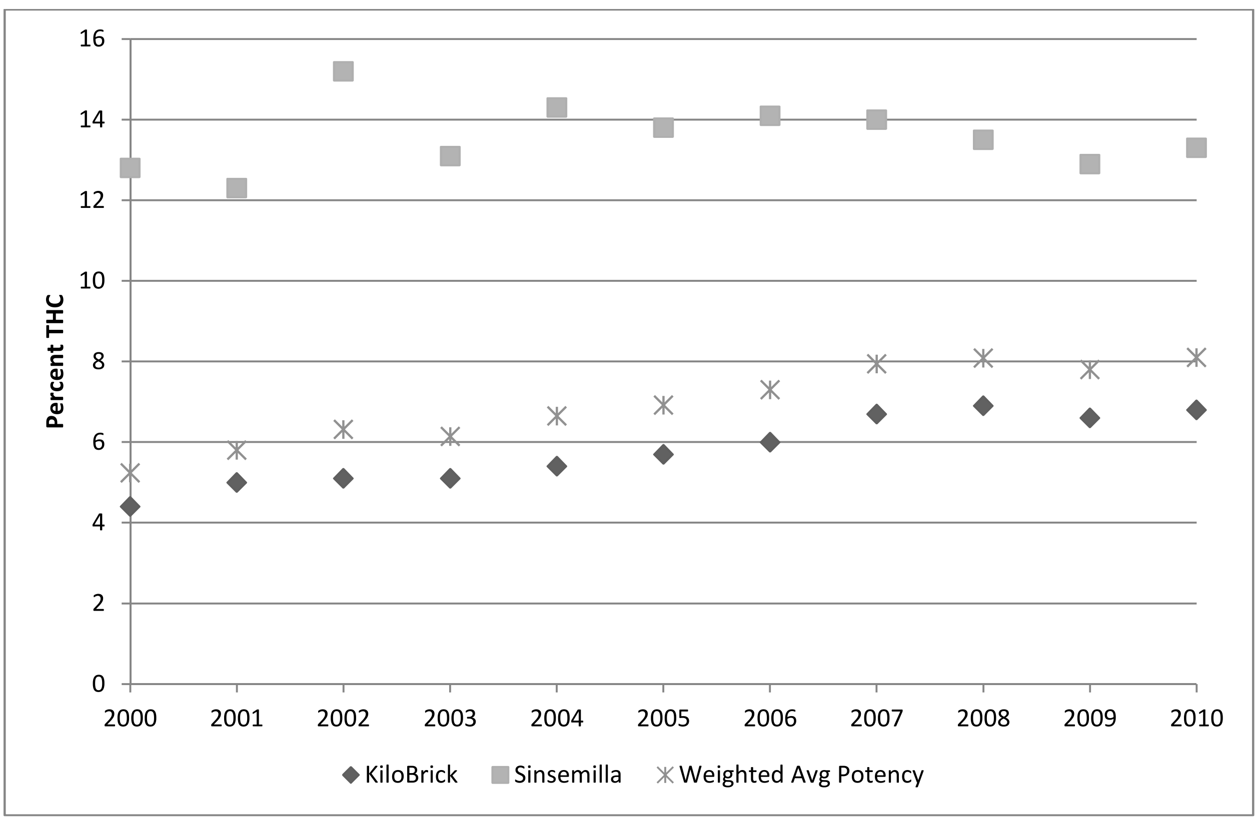
We ultimately want to find some weighted average of the two potency series—one for kilobricks (interpreted as commercial-grade marijuana) and one for sinsemilla—that best captures the overall trend in average marijuana potency. A strong finding is that composition effects can matter, and by using just one “type” of observation instead of a polyglot mixture, we find actual trends, not just composition effects. For example, the proportion of mature versus immature changes over time, as does the proportion of DEA versus state seizures. We believe that kilobricks in large quantities near the Mexican border are the most homogenous “type” that had a relatively large sample size and seemed related to the Mexican imports. These kilobricks are the bulk of what we attempt to capture for the non-sinsemilla part of the series. Empirically, kilobrick observations everywhere and of all sizes seem to be from the same distribution, so we use kilobricks throughout. We estimate Mexican kilobrick imports to be roughly 60 percent or more by weight, so using a consistent measure of the bulk of the non-sinsemilla market seems better than using a polyglot mix.

Table 7.2 presents estimates of the weighted average potency based on insights from MPMP (Columns 1 and 2) and NSDUH (Column 3). To generate the commercial-grade market share (Column 4), we generate a sinsemilla market share linear trend and subtract it from 1. The weighted average presented in Column 5 is generated via the following formula: (Col 4)*(Col 1) + (1-Col 4)*(Col 2). The various potency series, along with the calculated weighted average, are displayed in Figure 7.6.

Table 7.2: Generating a Weighted Average Potency

	(1) Average THC—Kilobrick	(2) Average THC—Sinsemilla	(3) Sinsemilla Market Share*	(4) Commercial Grade Market Share (Column 1 – Column 3)	(5) Weighted Average THC
2000	4.4	12.8	10%	90%	5.2
2001	5.0	12.3	11%	89%	5.8
2002	5.1	15.2	12%	88%	6.3
2003	5.1	13.1	13%	87%	6.2
2004	5.4	14.3	14%	86%	6.7
2005	5.7	13.8	15%	85%	6.9
2006	6.0	14.1	16%	84%	7.3
2007	6.7	14.0	17%	83%	7.9
2008	6.9	13.5	18%	82%	8.1
2009	6.6	12.9	19%	81%	7.8
2010	6.8	13.3	20%	80%	8.1

Figure 7.6: Various Potency Measures



Section 8. Contents of the Spreadsheets

The Excel workbook "RAND WAUSID Calculations (Delivered September 16 2013).xlsx" contains 30 spreadsheets used to generate the results, figures, and tables in the Main Report and the previous sections of this Technical Report. Tabs pertaining to marijuana are shaded green.

The first four sheets ("Cocaine", "Heroin", "Meth", and "Marijuana") present the main calculations for our estimates of CDUs, expenditures, and consumption. The first three begin with the total number of adult male arrest events involving a positive test for each respective substance (calculations described in Section 2). A series of adjustments are made to convert these figures to total CDUs, and then expenditure and consumption estimates (which we will describe later). The "Marijuana" tab starts with self-report information about marijuana use days from NSDUH.

The next three sheets ("CDU comparison with ONDCP 2012", "$ comparison with ONDCP 2012", "Q comparison with ONDCP 2012") compare our result for cocaine, heroin, and meth with what was reported in the previous version of this report (ONDCP 2012). The eighth tab ("Marijuana comparison with ONDCP 2012") compares the CDU, expenditure, and consumption estimates for marijuana.

The next five sheets ("Adj #1", "Dist by type of CDU", "Adj #2 & #3", "Adj #4", "Adj #5 & #6") present the data and calculations needed to adjust the number of adult male arrests events involving a positive test for cocaine, heroin, and meth to our final estimates of CDUs. Table 8.1 describes these adjustment factors and the data used to inform these six factors. The next sheet ("Adj #7") uses data from NSDUH, TEDS, and ADAM to understand the ratio of chronic to occasional users of cocaine, heroin, and meth.

Table 8.1 Adjustment Factors for Estimating the Number of Chronic Hard-Drug Users (4+ days in the past month)

Factor	Adjustments	Data	Years covered	For which user categories?
	Start: Number of adult male arrest events with a positive urinalysis test			
F1	1. For adult male arrest events with a positive test, percent using four or more days in past month[a]	ADAM-II (based on urinalysis and self-report information)	Average across annual estimates, 2000–03, 2007–10	Calculate for three groups: four to ten use days in past month, 11–20 days, 21 or more days
	Multiplying by F1 yields the number of adult male CDU arrest events with a positive urinalysis test			
F2	2. Number of arrests with positive test per person arrested and testing positive[a]	ADAM-I-II (based on self-report information about arrests in the past year, excluding warrants)	Estimate from pooled ADAM-I-II (2000–03, 2007–10)	Calculate for two groups: four to ten use days in past month, 11 or more days[b]
	Dividing by F2 yields the number of adult male CDUs who are arrested and have a positive urinalysis test			
F3	3. Proportion of adult male criminally active CDUs who get arrested each year[a]	Take arrests per arrestee from #2, assumes criminally active CDUs get arrested according to a Poisson distribution[c]	Estimate from pooled ADAM-I-II (2000–03, 2007–10)	Calculate for two groups four to ten use days in past month, 11 or more days[b]
	Dividing by F3 yields the number of criminally active adult male CDUs			
F4	4. Adult male CDUs who are not criminally active	Number of adult male CDUs who report never having been arrested in NSDUH, multiplied by 4[d]	Estimate from pooled NSDUH (2000–10)	Calculate for three groups: four to ten use days in past month, 11–20 days, 21 or more days
	Adding F4 gives the number of adult male CDUs			
F5	5. Ratio of adult CDUs (male + female) to just adult male CDUs	Drug-specific ratios from (1) NSDUH Days of Use, (2) NSDUH CDUs Days of Use, (3) NSDUH number of CDUs, (4) TEDS Users in Treatment, (5) TEDS CDUs in Treatment, (6) Drug Abuse Warning Network (DAWN), (7) Vital Stats overdoses	Generate 2000–2010 average for each of these seven factors, take simple average of these seven values	Calculated for one group: four or more days in past month
	Multiplying by F5 gives the number of adult CDUs			
F6	6. Ratio of all CDUs (adult + juvenile) to just adult CDUs	Drug-specific ratios from 1) NSDUH Days of Use, 2) NSDUH CDUs Days of Use, 3) NSDUH number of CDUs, 4) TEDS Users in Treatment, 5) TEDS CDUs in Treatment	Generate annual average across these five factors, impose linear trend for cocaine and meth (heroin constant 0.03)	Calculated for one group: four or more days in past month
	Multiplying by F6 gives the number of CDUs			

[a] ADAM-I-II: No weights, do not account for those who refuse urinalysis test, no data for 2004–2006.

[b] To boost sample, combine those who used on 11–20 days with those who used 21 or more days. Dropping those brought in on warrants has relatively little effect.

[c] Different from ONDCP (2012c) because we use Poisson assumption only to extrapolate to criminally active CDUs who did not get arrested, not to all CDU who do not get arrested.

[d] Based on ONDCP (2012c) assumption that occasional users of cocaine, heroin, and meth in NSDUH should be multiplied by four because of underreporting.

The next three sheets provide additional data for the hard-drug user estimates. "NHSDA-NSDUH Non-CDU" displays the number of chronic, occasional, and past-year—but not past-month—users of cocaine, heroin, and meth from NSDUH, 2000–2010. "Arrest Events" displays the 95-percent confidence intervals for our estimates of the number of national adult male arrests events involving a positive test, by drug and year. "Price Estimates" includes the prices at the appropriate referent quantities that are used to generate the consumption estimates.

The next three sheets ("Cocaine Indicators", "Heroin Indicators", "Meth Indicators") compare the CDU estimates for cocaine, heroin, and meth with other series related to consumption (e.g., TEDS admissions, DAWN emergency department visits). The final non-green tab ("Group Names") allows us to change how we label our frequency groups (e.g., those who used 21 or more days in past month = daily/near-daily users).

The eleven following spreadsheets all support the marijuana calculations in the "Marijuana" sheet. "NSDUH by Intensity" presents annual estimates of users, past-year use days, and past-month use days by year and frequency group. These figures are from NSDUH and have not been adjusted for undercounting. The "MJ Youth Adj" tab adjusts the NSDUH figures for youth based on the prevalence estimates presented in the "NSDUH, MTF, and YRBS" sheet. Similarly, "MJ Adult Adj" adjusts NSDUH figures based on calculations in the "NSDUH, UCR, and ADAM" sheet. The "NSDUH by Subpopulation" sheet provides additional information about how the number of past-month marijuana users is broken down by age and criminal justice status. The "NSDUH Raw data" present the raw data that underlie all of these calculations. The "Potency" sheet plots UMISS potency data by kilo-brick and sinsemilla and generates a weighted average based on our estimate of sinsemilla market share.

the two final marijuana sheets deal with marijuana expenditures and prices. The "NSDUH & ADAM prices" sheet presents the price per gram based on a half-ounce purchase from various datasets. The "Expenditure Sensitivity" sheet display how the marijuana expenditure estimates vary based on alternative assumptions of THC consumption and the price elasticity of demand.

Finally, the "Ch6 Polydrug" sheet shows the number of chronic users of cocaine, heroin, and meth in NSDUH, plotting their levels, sum, and the union of polydrug users. "Ch7 Supply Charts" includes the underlying data for the charts presented in Chapter 7.