Technical Report For

The Price and Purity of Illicit Drugs:

1981 Through the Second Quarter of 2003

NOVEMBER 2004

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF NATIONAL DRUG CONTROL POLICY

Technical Report
For the Price and Purity of Illicit Drugs:
1981 Through the Second Quarter of 2003

Executive Office of the President
Office of National Drug Control Policy
Washington, D.C. 20503

November 2004

Acknowledgements

This publication was sponsored by the Executive Office of the President, Office of National Drug Control policy (ONDCP). Jeremy Arkes, Rosalie Liccardo Pacula, Susan Paddock, Jonathan P. Caulkins, and Peter Reuter produced this publication at RAND's Drug Policy Research Center and Public Safety and Justice Division under a contract to ONDCP under DHHS prime contract number 282-00-0016, Task Order number 8. Michael Cala served as Project Manager for this publication.

Disclaimer

Public Domain Notice

Electronic Access to Publication

This document can be accessed electronically through the following World Wide Web address: http://www.whitehousedrugpolicy.gov/

Originating Office

Executive Office of the President
Office of National Drug Control Policy
Washington DC 20503

Publication Number NCJ 207769
Printed in November 2004

NCJ 207769

TABLE OF CONTENTS

TABLES

FIGURES

I. Introduction

The purpose of this technical report is to provide detailed documentation of all programming and statistical activities involved in the construction of the price and purity series presented in *Price and Purity of Illicit Drugs: 1981 Through Second Quarter of 2003*, a companion report prepared for the Office of National Drug Control Policy (henceforth referred to as the main report). This technical report contains information on the precise steps taken to (a) acquire the data, (b) assess the quality of the data, (c) clean the data, (d) use the data to generate city-specific price and purity series, and (e) build national indices of price and purity from city-specific estimates. It serves as a roadmap for anyone interested in replicating the results presented in the main report. Because decisions had to be made at various stages about which data to use and how to use them, this document also provides a justification for the decisions made throughout the development of the main report. Many of these decisions were made on the basis of additional analyses not presented in the main report and follow-up discussions with key Drug Enforcement Administration (DEA) personnel who are responsible for maintaining the System to Retrieve Information from Drug Evidence (STRIDE) database and who are knowledgeable about how these drug transactions occur. We would like to thank the various DEA staff for all of their time and assistance in educating us about how things get recorded in STRIDE.

II. Data

Since the 1970s, the Intelligence Division of DEA has been managing STRIDE, a forensic database containing detailed information on the amount, type, potency, and source of drugs acquired by DEA through their various law enforcement activities. The primary purposes of the database are to control the inventory of drug acquisitions and to assay the characteristics of the drugs collected so as to keep law enforcement informed about the drugs being traded and to assist prosecutors in the prosecution of federal drug offenders. The STRIDE database contains information only on drug acquisitions that are sent to a DEA laboratory for analysis. Thus, the vast majority of the data reflect drug acquisitions obtained by DEA agents or through their informants. To the extent that other agencies rely on DEA laboratories for analysis of their own drug acquisitions (e.g., the Washington DC Metropolitan Police Department), such acquisitions are also included in STRIDE. However, most local and state police agencies do not submit information on their undercover acquisitions to DEA laboratories.

The STRIDE data are not collected for the purpose of conducting statistical analyses of drug transactions. Thus, observations are not obtained through the construction of a probabilistic sampling frame of all drug market transactions in a given geographic area. Instead they represent specific transactions that were targeted by law enforcement agencies. Further, the timing and location of encounters varies considerably from quarter to quarter and from year to year. Therefore, the information about drug transactions provided in the data is not representative of all drug transactions that occur in these areas. That does not mean, however, that these data do not provide information. Instead, it means that the distribution of price and purity values observed in the STRIDE data does not mirror the actual distribution of price and purity observations in the United States, and standard statistical assumptions regarding the asymptotic properties of the distribution of data in STRIDE will not hold. Numerous statistical methods have been developed in recent years to deal with drawing inferences from non-representative data and convenience samples. The current report uses some of these new methods and proposes the adoption of additional methods that could be pursued in future efforts.

Of course, these statistical methods cannot address possible systematic differences between prices paid by law enforcement agents and their informants as recorded in STRIDE, on the one hand, and prices paid by criminal participants in the drug trade, on the other. For example, if law enforcement agents and their informants systematically paid 10 percent more to acquire drugs than did criminal participants, then every observation in STRIDE would be inflated by 10 percent relative to what criminal participants pay, and no analysis based on the STRIDE data alone could detect or adjust for that 10 percent differential. It is not uncommon to assume that law enforcement agents and their informants must pay very nearly the market price; otherwise, the other parties to the transaction would realize who they are. A similar but milder assumption is that law enforcement agents and their confidential informants may pay systematically more (or perhaps less) than do typical market participants, but that any differential is likely to be stable over time and, hence, will affect only estimated price levels, not trends in those levels. Since the present analysis is based only on STRIDE data, we cannot comment on the reasonableness of either of those assumptions.

Acquiring the STRIDE Data

Historically, DEA has been willing to share nonconfidential information in STRIDE with other government agencies and the public. However, no formal codebook of the database exists. Hence, persons interested in using these data have not always known exactly what to ask for or how to ask for it. Equally problematic, non-DEA users may not fully understand what it is that they have requested. For example, early in this project we received a download of STRIDE data that did not contain the FORM field. This field identifies the units in which the transaction amount is measured (e.g., grams, capsules, milliliters). While most observations are measured in grams, about 4 percent of the observations in STRIDE are measured in other units, with specific drugs being affected differently. For example, 16 percent of the methamphetamine observations are measured in units other than grams, while less than 1 percent of the cocaine observations are measured in grams. It is our understanding that some past users of STRIDE data received similar files and believed that all of the observations were measured in grams. Clearly, mistakenly assuming that an observation describes the amount paid for 2 grams when in fact it was the price paid for 2 capsules can affect price estimates, even when the proportion of observations measured in units other than grams is small.

As another example, updated downloads of STRIDE (as opposed to new, complete downloads) can contain nearly but not exactly identical records concerning the same transaction. The problem with simple updates of STRIDE is that cases that are seized but not yet analyzed could show up twice, once with incomplete information for various fields and again with more complete information after the data have been analyzed in the lab. Cleaning the merged (original plus updated) data by eliminating only exact duplicate records would not eliminate such double counting. Double counting of an observation, particularly one that happens to have an unusually high or low price can clearly affect estimates of average prices.

An implication of these kinds of complications is that an important first step in documenting the work done for this project is to describe how the data were requested so that other researchers and analysts can replicate the original starting sample of data. The Office of National Drug Control Policy (ONDCP) requested that a colon-delimited ASCII data file be constructed that satisfied the following three main criteria[1]:

(1) The *date analyzed* (not the date the observation was acquired/seized) must be between January 1, 1981, and "the present" (the letter was dated June 27, 2003).

(2) The following Primary Drug Categories must be included: Cocaine (620), Heroin (610), Cannabis (531), Methamphetamine (111), MDMA & other hallucinogens (560), Heroin signature program (904), Domestic monitoring program (905), CHEMCON (906) and Cocaine signature program (912).

(3) All open and closed cases must be included.

[1] Even the colon delimitation is important. Some past requests have been for semicolon delimited files, but it turns out that (at least) two of the STRIDE fields contain semicolons within the field for some records. Hence, all fields to the right of those fields containing semicolons will report incorrect information.

Requesting information based on the date analyzed instead of the date seized helps reduce the likelihood of getting duplicate observations. Further, information on purity is recorded only after the data have been analyzed. Prior to analysis, the purity field may have missing information or a zero. By limiting the data to only those observations that have been analyzed (i.e., using *date analyzed*), it is possible to interpret zero and missing purity information as real information about the transaction.

It is important to request both open and closed cases to obtain all observations that have been analyzed by the labs. Some fields included in STRIDE are identified by DEA as “sensitive” because they contain information that is pertinent to a case currently being processed by the judicial system (i.e., an open case). If these fields are unwittingly requested, DEA will send information only on closed cases, to protect the integrity of cases currently under investigation. Since the average case is open for about three years, conducting analyses on only closed cases can dramatically limit the number of observations available for evaluating recent trends in prices and/or purity. The typical user is unlikely to know which fields are sensitive and which are not. However, by making a request that both open and closed cases are desired, it is possible for DEA to follow up with the requester to determine whether the sensitive field is truly necessary or not, so that all the data can be sent.

Specific fields requested for this project included the following: city, state, country, date received in lab, date collected, date analyzed, drug name, exhibit number, federal number, lab number, method of acquisition, domestic monitoring program flag, net collected, net weight pure, non-DEA case number, number of packages, office code, collecting office name and code, potency, price per pure gram, primary drug category, primary drug code, received amount, related inventory number, signature, status, STRIDE identifier, submitted amount, total purchase cost, agency submitting, enforcement group, form, suspected drug category, suspected drug code, secondary drug code, dosage, number of doses, number of packages, and package description. Inclusion of the related inventory number is particularly important for cocaine and heroin observations because it enables the user to identify those observations that appear twice in the dataset. Double entries occur for some cases obtained through the Heroin Signature Program (HSP) and the Cocaine Signature Program (CSP) because acquisitions obtained through these two initiatives can be sent to the laboratory twice. When the samples from these initiatives are first sent to the lab, they are entered into the forensic database with special drugcodes indicating that they were obtained through one of the signature programs. They are not given a standard STRIDE ID because the sample is analyzed differently, with the goal of obtaining information about the country or region of origin (i.e., the “signature”) of the plant material in the sample. The sample may then later be sent to another lab to examine the purity of the drug it contains, at which point it is given a STRIDE ID and entered into the database a second time. The two entries are linked through the related inventory number.

The file we received from DEA contained 782,031 records, including domestic and foreign data from January 1, 1981, through May 31, 2003.

Constructing the Sample

We began with the raw data file sent to ONDCP and then imposed several restrictions on the data to limit the sample to only the data needed for this analysis. Table 1 summarizes how specific restrictions incrementally affected the sample. For ease of presentation, we use the term *MJ* to refer to marijuana observations, *Meth* to refer to d-methamphetamine observations, and *DMP* to refer to observations from the Domestic Monitoring Program.

(1) Basic Steps to Reduce the Sample to Relevant Observations

We started by identifying, by drug category, all cocaine (drug category 620), heroin (610 and 905), methamphetamine (111), and marijuana (531) observations. The decision to include DMP observations (primary drug category 905) in the sample was made after hearing from DEA personnel that these observations are acquired in a manner consistent with the bulk of the other STRIDE heroin observations. This is not generally true about the observations acquired through the CSP and the HSP, which is why observations from these drug codes were dropped. This restriction reduced the sample to 753,845, as seen in the second row of Table 1. Next, we checked for duplicate records and found none. We then restricted the sample to include only observations from within the United States, which eliminated a little less than 3 percent of the overall sample. An additional 93 observations were dropped because of missing information on the state in which the acquisition occurred. These deletions impacted primarily the sample for heroin.

The next major restriction imposed on the data was to limit the sample to observations where the method of acquisition was either a purchase (P), a seizure (S), or a lab seizure (L). This restriction was based on a recommendation made by DEA, who explained that information obtained from other types of acquisitions (e.g., "flashing money") may be less reliable than that acquired through purchase attempts and seizures, because such transactions may not have been completed. This restriction reduced the total sample by 5.2 percent, to 693,648, with heroin again having the greatest relative decrease in sample size.

Table 1. Impact of Sample Restrictions on Number of Observations Included in Analysis

	Number of observations remaining in sample							
	Total	**MJ**	**Meth**	**Heroin**	**DMP**	**Cocaine**	**Powder Cocaine**	**Crack Cocaine**
Starting number	782,031							
Restrict to 4 main drug classes	753,845	212,621	88,647	105,021	8,894	338,662		
Only U.S.	731,437	210,897	87,130	97,251	8,514	327,645		
Non-missing state	731,344	210,882	87,130	97,179	8,514	327,639		
Purchases and seizures only	693,648	204,057	81,531	89,123	8,465	310,472		
Raw weight > 0	690,749	203,625	81,233	88,622	8,455	308,814		
Measured in grams	662,168	197,009	68,055	86,542	8,294	302,268		
Purity is non-missing and purity ≤ 100	662,114	197,007	68,052	86,536	8,294	302,225		
Narrowing drug codes	600,139	193,641	60,926	60,558	8,294	276,720	154,155	122,565
Reassigning heroin DMP	599,425	193,641	60,926	68,138			154,155	122,565
Weight ≥ 0.1 gram	554,815	186,637	50,002	63,482			145,353	109,341
Final sample for purity analysis: purchases and seizures	**368,178**	**---**	**50,002**	**63,482**			**145,353**	**109,341**
Price > 0 and non-missing	137,222	4,695	12,313	27,953			45,618	46,643
Remove other gross outliers	136,505	4,597	12,232	27,797			45,423	46,456
Delete crack if year < 1986	136,268	4,597	12,232	27,797			45,423	46,219
Delete obs in city-quarters with < 5 obs.	136,213	4,597	12,181	27,797			45,419	46,219
Stage I: sample for purity models	128,283	---	11,976	27,262			44,913	44,132
Stage II: final sample for price models	131,184	4,359	11,682	26,594			43,953	44,596

* Note that marijuana is not evaluated in the purity analyses with purchases and seizures because data on purity of marijuana are not available in STRIDE.

(2) Preliminary Data Cleaning

The next steps focused on the two primary descriptive variables in the data: amount and potency. First, we deleted observations with missing or zero amounts, where amount represents the raw weight of the purchased or seized package. This had a very small effect on the sample. We then further restricted the sample to observations that were measured in grams, because the remaining forms could not be easily converted into grams. For example, 16.2 percent of the methamphetamine observations were measured in either milliliters (MLS) or capsules (CAP). The precise conversion of these units into grams depends on a number of factors we cannot observe, and hence we decided to delete these observations from the data. This restriction reduced the overall sample available for analysis by 4.1 percent overall. Next, we deleted 54 observations in which either purity was missing or the purity measure was greater than 100 percent. Most of these were cocaine observations.

The next data-cleaning step involved the identification of more homogenous drug codes within specific drug categories. A major criticism raised by the National Research Council regarding past price indices constructed from STRIDE data concerned the aggregation of prices across different forms of a drug that could represent different drug products to consumers.[2] To address this concern, we examined the drug codes represented under specific drug categories and consulted with representatives from DEA to identify those drug forms that were and were not likely to be physically distinguishable to a buyer. Using this information, we aggregated within a drug category drug forms that were physically similar and, hence, likely to be indistinguishable to a buyer.

In the case of cocaine, more than 95 percent of the observations fell under three primary drug code categories: 9041L000 (crack), 9041L005 (powder), and 9041L900 (cocaine, salt undetermined), with the first two categories representing nearly 92 percent of all observations. Crack and powder cocaine are easy to differentiate upon physical inspection, so observations from these two drug categories were separated, as shown in the "Narrowing drugcodes" row of Table 1. DEA informed us that the "salt undetermined" category generally reflects cocaine observations that are too small to analyze chemically. They may or may not have been sold to a buyer as a specific form of cocaine (powder or crack), but the lab technicians were unable to determine the salt attached to the drug, given the time and resources available at the time of analysis. Therefore, the observation was labeled "undetermined." Because this category represents a heterogeneous mix of unidentified cocaine types, we exclude it from further analysis. Excluding the "salt-undetermined" and other cocaine drug forms reduced the total cocaine sample by 8.8 percent.

Most of the heroin observations have one of four drugcodes: heroin hydrochloride (9200.005), heroin base (9200.000), Domestic Monitor Program (9DMP.000), and salt undetermined (9200.900). The breakdown of these heroin observations is as follows:

Type	Frequency	Percent
Heroin base	2,415	2.79
Heroin HCl	58,190	61.33
Heroin, salt undet	25,898	27.30
DMP	8,294	8.74

Again, DEA assisted us in the identification of which drug codes to merge. They informed us that heroin base cannot easily be physically distinguished from heroin hydrochloride (HCl) without chemical analysis, and thus it makes sense to group these two products together. Mexican "black tar" heroin is very easy to physically distinguish from heroin powder, but its physical form does not have a one-to-one correspondence to a distinct chemical form of heroin. The specific chemical form of a black tar sample can be determined by the DEA chemists, but the process can be much more time- and resource-intensive than the process for powder forms of

[2] Manski, Charles F., John V. Pepper, and Carol V. Petrie (eds.), 2001. *Informing America's Policy on Illegal Drugs: What We Don't Know Keeps Hurting Us,* National Academy Press, Washington, DC.

heroin. Hence, it would not be uncommon for the DEA labs to simply ascertain that a black tar sample was heroin but not to further identify its specific type unless there was a specific need for that more precise determination. Many black tar observations are therefore placed in the *heroin, salt-undetermined* drug form. However, not all salt-undetermined observations are black tar heroin. An observation could get classified as salt-undetermined for other reasons. For example, if the heroin is mixed with many other diluents and adulterants containing salt bases, it could again be prohibitively time-consuming for a chemist to distinguish the salt attached to the heroin from the salts attached to the diluents and adulterants. Alternatively, the sample might be too small. Hence, the salt-undetermined category represents a heterogeneous mix of different forms of heroin. As such, it was decided that this form of heroin would remain separate from the other two forms (heroin base and heroin hydrochloride). Further, we decided not to construct a formal price/purity series for this form, since changes in the series could reflect changes in the makeup of the form included in the salt-undetermined sample rather than real trends in price or purity for a particular form of heroin. We verified, however, that a series for the heroin salt-undetermined category follows a different general trend from that observed for the main heroin series (see Section VIII, Supplemental Analyses). For example, in recent years, the purity of salt-undetermined heroin is generally lower than that of heroin base or heroin hydrochloride.

The DMP category consists of purchases made through the Domestic Monitoring Program, a program in which law enforcement goes into specific cities and makes small buys (usually $100) of whatever type of heroin is available on the streets. Thus, these DMP observations are a heterogeneous mix of different forms of heroin. However, the specific forms can be identified through the secondary drug code, and thus it is possible to identify which DMP observations are heroin hydrochloride and which are heroin base. Examination of the secondary drug code for the sample of DMP observations in this dataset revealed that 7,580 observations could be included in the main heroin sample because they were either heroin hydrochloride or base. All other heroin observations were dropped.

Three main types of methamphetamine are marketed: d-methamphetamine, dl-methamphetamine, and l-methamphetamine. These types differ in the form of the isomer, something that is not immediately apparent to the buyer at the time of the transaction. However, according to DEA personnel, the three types of methamphetamine differ significantly in their quality, so sellers usually make the type known to the buyer as a way of indicating the quality of the drug. In the STRIDE data, the great majority of observations measured in grams were of d-methamphetamine. The other two forms of methamphetamine together made up only 10.5 percent of the total methamphetamine sample. Hence, it was possible to develop a price series only for the d-methamphetamine type.[3]

More than 97 percent of the marijuana observations fell into one of two drugcode categories: 7600.000 (no plant material detected) and 7360.4 (all plant material).[4] The next two largest forms, intact plants (7360.5) and cannabis seeds (7360.0), would be very easy to physically differentiate from general plant material and were too small to generate their own price series, so

[3] STRIDE contains numerous drug codes that represent different physical forms of d-methamphetamine. According to information from DEA personnel, the physical form is not nearly as important to a buyer as the type. Thus, for the construction of this methamphetamine price series, all forms of d-methamphetamine were included.
[4] These two forms were grouped together on the basis of a recommendation by DEA.

they were dropped. Observations from the drug code labeled Tetrahydrocannabinol-Organic (7371.000) were retained, however, because this category indicated a plant-like material, which would be difficult to physically distinguish from the other forms. Limiting the marijuana sample to these three drug codes decreased the sample by only 1.7 percent.

(3) Gross Outlier Deletions

Next, we deleted observations with a weight of less than 0.1 grams. This data-cleaning step has been used in previous reports and also by other researchers using the data.[5] The primary justification for the deletion is that purity data are unreliable for observations weighing less than 0.1 grams. As evidence, there is a disproportionate number of low and zero-purity observations at these very small quantities, presumably because it is difficult for lab technicians to chemically determine the potency and specific drug forms involved.[6] Approximately 7.0 percent of the overall sample was lost after imposing this restriction, the largest losses occurring for the d-methamphetamine sample (18 percent) and the crack cocaine sample (11 percent).

Excluding the marijuana observations because they contain no information on purity, the resulting sample of 368,178 observations was the primary sample used to evaluate purity with both the seizure and purchase data. Additional data-cleaning steps were taken to arrive at the sample for price and purity models that employed primarily purchase data.

The first step taken to generate a sample for estimating price, after bringing marijuana observations back into the sample, involved deleting those observations that were missing information on cost. The vast majority of observations with missing price information are seizures, although a few purchase observations are also missing price information.[7] Deleting observations in which price was missing significantly reduced the sample, to 137,222, a 75.3 percent decrease overall. The largest percentage decrease was for marijuana, the sample for which was decreased by 97.5 percent, to 4,695 observations.[8]

Next, observations that were outside the distribution of realistic prices for 1 gram not adjusted for purity were deleted. Criteria for deleting specific data points generally followed those employed in previous reports (see Table 2). For example, observations in which the nominal price (i.e., price not adjusted for inflation) was too low or the inflation-adjusted (or real) price per gram was too high were dropped. However, in contrast to previous reports, observations in which the inflation-adjusted (real) price per gram was too low were also dropped. This additional restriction was placed on the data following close examination of the distribution of

---

[5] Caulkins, J.P. (1994), *Developing Price Series for Cocaine*, MR-317-DPRC, RAND, Santa Monica, CA; Saffer H. and F.J. Chaloupka (1995). “The Demand for Illicit Drugs” National Bureau of Economic Research Working Paper #5238, August 1995; Grossman M., F.J.Chaloupka, and C.C. Brown (1998), “The Demand for Cocaine by Young Adults: A Rational Addiction Approach,” *Journal of Health Economics*, Vol. 17, No. 4, pp. 427–474.

[6] In previous reports, observations for marijuana in which the amount, not adjusted for purity, was less than 0.2 grams were deleted. The methods for marijuana employed here are therefore less restrictive than those used in previous reports.

[7] Crack cocaine purchases had the most cases of missing price information, with 1944 purchase observations missing price. The number of purchase observations missing information on price for powder, heroin, methamphetamine, and marijuana were 1645, 598, 365 and 775, respectively.

[8] There were a few seizure observations that included a non-zero price (0.3% of our sample). Given that price was reported for these observations, they were retained under the presumption that a coding error had occurred labeling these observations as seizures when they were, in fact, purchases.

real prices per gram (not adjusted for purity) for each drug. In total, these conditions reduced the total sample by only 0.5 percent overall. Drug-specific samples were reduced by less than 1 percent, with the exception of marijuana, which had a 2.1 percent reduction.

Table 2. Criteria for Deleting Data Points, by Drug

Drug	Nominal Price	Real Price per Gram	Real Price per Gram
Cocaine	< $ 3.00	< $ 2.00	> $ 3,000
Heroin	< $ 3.00	< $ 7.50	> $10,000
d-Methamphetamine	< $ 3.00	< $ 2.00	> $ 3,000
Marijuana	< $ 0.10	< $ 0.05	> $ 100

(4) Additional Data Cleaning Related to Proper Model Estimation

Because data for crack cocaine were sparse prior to 1986, all crack cocaine observations between 1981 and 1985 were deleted. There were only 137 crack observations for the entire 1981–1985 time period, so runs that were done including these observations generated extremely volatile trends. The volatility was attributed to the enormous sampling error and did not represent true volatility in crack cocaine prices.

Similar artificial volatility was possible in all of the drug models because of the geographic and temporal variability of law enforcement activities. Thus, to limit the amount of artificial volatility in the time series, we required, for each drug model, that a cell, defined as a particular quarter and year, had to have at least five observations to be included in the model. For the marijuana models, at least five observations had to exist for a given year, because the models are estimated on the basis of years. The justification for this restriction was that it would reduce the sensitivity of the model to unusual or outrageously large/small prices observed in specific locations that might otherwise heavily influence the prediction for that quarter-year. Requiring at least five observations in a given quarter increases the probability that deleting observations with extreme residuals (described below) will delete the unrealistically priced observations. This restriction deleted only 51 methamphetamine and four powder cocaine observations.

The resulting sample of 136,213 observations was the starting sample for the estimation of the price/purity models.

As will be described in Chapter IV, the first stage of the price/purity model—the purity equation—excluded observations with a purity of zero. The reason for this is that a zero purity in STRIDE may indicate a true zero purity or it may indicate that the purity was missing (not yet determined). Thus, to avoid incorrectly assigning zero to the purity of many observations, we just deleted these observations. This reduced the samples for d-methamphetamine, heroin, and cocaine powder by less than 2 percent, but it reduced the crack cocaine sample by 4.5 percent. Once the purity model is run, however, it is possible to generate a predicted purity for all of these zero-purity observations, allowing them to be brought back into the analysis for estimating the final price model. Thus, the omission of zero-purity observations influenced only the sample of observations predicting the purity model.

The final step to reach the end sample for the price model was the exclusion of extreme residuals. The exact process for deleting extreme residuals (described in Chapter IV) basically involved an iterative loop where extreme residual observations were dropped from the model until no more could be identified. Comparing the last row in Table 1 to the third-from-last row, one can see the impact of deleting the extreme residuals from the samples. Overall, the total sample decreased by 3.7 percent, with the sample for marijuana being the most impacted. The final total data sample contained 131,184 observations.

(5) Other Modifications to the Data

In addition to the above general modifications, some additional issues needed to be addressed. First, while this was not an issue for the last data extract from DEA, a prior extract from STRIDE had duplicate records, as indicated by the STRIDE ID variable. Duplicate records can occur if users request updates of the STRIDE database instead of complete downloads starting at the same base year, because cases in STRIDE that have not yet been analyzed at the lab can be analyzed during the updating period. Both observations would have the same STRIDE ID but would contain different information for potency and other variables of interest. In addition, duplicate records may result from the inclusion of observations from specific operations (e.g., DMP or CSP). These can be identified through a non-missing inventory number, which will specify an original STRIDE ID if one was previously assigned to that observation. Finally, some records in STRIDE appear to look like duplicates, although different STRIDE IDs have been assigned or specific variables (e.g., lab number) differ slightly. Although these “nearly duplicate” records are maintained in the current analysis, future work should evaluate whether they are indeed unique.

A relatively minor data-cleaning issue was that of the correction of erroneous date codes. A very small number of observations (21) had seizure or purchase dates that were just not possible (e.g., February 29 in a non-leap year or 31 in months with only 30 days). It was presumed that these erroneous dates were coding errors specific to the day of the month, so these observations were back-coded to the closest earlier logical date, following an algorithm employed in earlier reports.

The data-cleaning steps outlined above generally followed those undertaken in previous reports, with three primary differences. First, previous reports aggregated all the drug codes within specific drug categories, so, for example, crack and powder cocaine (as well as the other, less frequent drug codes under cocaine) were estimated together as a single series. Likewise, all drug codes for heroin, d-methamphetamine, and marijuana were included in the samples employed in previous reports. Hence, the series presented in the current main report do not represent the same drugs as the series in previous reports. Second, previous reports did not delete observations in which the real price per gram (not adjusted for purity) was too low. Instead, those reports imposed restrictions that required potency to be above specific thresholds for each drug. The present report does not include this restriction because low-potency observations are viewed as valid “rip offs” in the data, and the model is modified to accommodate the information revealed from these transactions. Third, previous reports did not exclude quarters with too few observations. This additional restriction was imposed in order to reduce artificial volatility in the price and purity trends.

Defining Key Variables

The Dependent Variables

The primary dependent variable in the price model is the inflation-adjusted price (or "real price") of a particular drug. We adjust for inflation using the quarterly Consumer Price Index for All Urban Consumers (1982–1984 base year).[9] All prices are reported in constant 2002 dollars because 2002 is the last year for which the STRIDE data are fully reported. Thus, the prices for observations up to the second quarter of 2002 are adjusted upward for inflation, while those starting in the third quarter of 2002 are adjusted downward.

For the powder cocaine, crack, heroin, and d-methamphetamine price models, a two-stage estimation technique is employed in which the purity of the drug is estimated in a first stage regression and the predicted value of those purity observations, or the expected purity, is put into the right-hand side of the price model. Thus, the purity of each drug is also a dependent variable. Because the unit for potency in the raw data from DEA is *percent* (so something that is 95 percent pure has a value of 95), this variable is converted into a fraction by multiplying each value by 0.01 before entering it into the model. Hence, all of the potencies reported in the paper are presented as fractions, where 1.00 = 100 percent pure.

The Independent Variables

One of the key dependent variables in both the price and purity equations is the amount of the drug involved in the transaction. DEA reports the weight involved in a transaction under its AMOUNT variable. Note that the measure AMOUNT does not adjust for purity; it simply reflects the weight of the transaction in grams.

All of the price and purity models also account for the quarter and year in which a transaction took place. Information on the date of the transaction is obtained from the variable DATE SEIZED. From this we construct a series of quarter/year interaction terms by constructing 90 dichotomous indicators (T1 through T90) representing specific quarter/years from the first quarter of 1981 through the second quarter of 2003. A transaction that took place on March 24, 1990, for example, would have a value of one for the time indicator T37 and zero for all the other time indicators (T1 through T90).

Finally, the models include information on the location where the transaction took place. Previous models estimating price and purity indices using these data included dichotomous indicators for 29 metropolitan areas (called *cities*) in the data and an "other" category that encompassed the rest of the country outside those 29 cities. The cities that have unique identifiers are Atlanta, Baltimore, Boston, Buffalo, Chicago, Cleveland, Dallas, Denver, Detroit, Houston, Kansas City, Los Angeles, Miami, Milwaukee, Minneapolis-St. Paul, New Orleans, New York, Newark, Philadelphia, Phoenix, Pittsburgh, Portland, San Antonio, San Diego, San Francisco, Seattle, Saint Louis, Tampa, and Washington DC.[10]

---

[9] Historical information on the annual Consumer Price Index for All Urban Consumers (seasonally adjusted) can be obtained from the Bureau of Labor Statistics at http://www.bls.gov/cpi/home.htm.

[10] Other cities in the United States are also represented in the data, and unique identifiers could be included for those with sufficiently large numbers of observations over time.

The current specification of the price model includes dichotomous indicators for the same 29 locations as in the previous report, but the remaining “other” category is subdivided into nine separate Census divisions: Pacific, Mountain, North West Central, East North Central, West South Central, East South Central, South Atlantic, Mid Atlantic, and New England. While the earlier approach would group, say, rural Florida and Montana into the same “other” category, this approach separates these locations by assigning them to their specific region, thus reducing the amount of unexplained variation remaining in prices. The current model identifies 38 geographically distinct areas, whereas previous methods identified only 30.

The nine Census divisions used in this analysis are defined as follows:

- Pacific: Alaska, Hawaii, Washington, Oregon, California
- Mountain: Arizona, Idaho, Montana, Colorado, New Mexico, Utah, Nevada, Wyoming
- West North Central: North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa, Missouri
- East North Central: Wisconsin, Illinois, Michigan, Indiana, Ohio
- West South Central: Texas, Oklahoma, Louisiana, Arkansas
- East South Central: Mississippi, Alabama, Kentucky, Tennessee
- South Atlantic: Florida, Georgia, South Carolina, North Carolina, Virginia, District of Columbia, Maryland, Delaware, West Virginia,
- Mid Atlantic: New Jersey, New York, Pennsylvania
- New England: Connecticut, Rhode Island, Massachusetts, Vermont, New Hampshire, Maine

It is also important to clarify the way the cities are defined. Previous reports tried to aggregate cities into metropolitan statistical areas (MSAs). Definitions for these MSAs, however, have changed over time as new cities and areas have become incorporated. We decided that the current project would employ the most recent definitions from the Office of Management and Budget (OMB) in 2003 to deal appropriately with new locations not previously included in earlier analyses. Using lists of all towns and cities included in specific counties from the Office of Social and Economic Data Analysis (<http://oseda.missouri.edu/plue/>) and MSA definitions based on counties and cities, we constructed SAS code that sorted all of the city names provided in STRIDE into one of the 29 MSAs or the nine Census regions.[11]

[11] A few cities within an MSA did not properly get categorized into that MSA based on the data from the Office of Social and Economic Data Analysis. In these cases, code from the previous report was used to assign these cities to an MSA.

III. Redefining Distribution Levels

It is widely believed that drug markets offer quantity discounts, or that the price paid per gram of a substance falls as the quantity purchased rises.[12] Indeed, our previous reports support this hypothesis. However, little is known about the levels at which these quantity discounts kick in or what transaction sizes are involved for defining different market levels. It is extremely difficult to identify market distribution levels for specific illicit drugs because it is impossible to perfectly classify all observations. For example, the term *retail* typically refers to transactions in which the buyer is the end-user, but there is no way of identifying end-users from low-end sellers in the STRIDE data. All that is available in STRIDE is the amount traded. Some people may buy large quantities for their personal consumption over a long time period, while others might buy small quantities with the intent of further dividing the substance into individual-size packages to sell separately.

In previous reports, different distribution levels were defined on the basis of the number of pure grams involved in the transaction. This report deviates from that procedure in two ways. First, it no longer specifies quantity ranges based on pure grams, but rather bases quantity ranges on amounts unadjusted for purity. It is more natural to think of distribution levels in this fashion, and doing so reduces the likelihood that valid rip-offs get misclassified. Under the old classification scheme, a transaction involving 100 grams of heroin that was 0.1 percent pure would be analyzed as a *retail* transaction because it involved 0.1 pure grams of heroin. Under the present scheme, such a transaction would be grouped with other transactions involving similarly large amounts.[13]

---

[12] Caulkins, J.P. (1995), "Domestic Geographic Variation in Illicit Drug Prices," *The Journal of Urban Economics*, Vol. 37, No. 1, pp. 38–56.

[13] An alternative approach would be to base the distribution level on the cost of the transaction. While this approach has many desirable features, it could create problems. First and foremost, it could cause a misrepresentation of price differences across distribution levels because higher-price-per-pure-gram transactions would get pushed to higher distribution levels. For example, consider a case in which there are two heroin observations of 1 gram at 50 percent purity, with one transaction being for $300 and the other for $150. If $200 were the threshold between the first and second distribution levels, then these two observations would be classified in different categories even though they involve transactions of the same quantity and purity. A second potential problem that could result from defining distribution levels based on the cost of the transaction is that a large price change over time could affect the classification of a transaction.

Second, this report does not refer to these different levels as distribution levels, but instead refers to them as quantity levels. Although some transaction sizes are relatively more common than others in the data, it is not entirely clear which transaction sizes clearly distinguish end-users from low- and mid-level sellers. In an effort to avoid problems associated with specific definitions used to describe precise distribution levels, observations for each drug are simply separated into three (or, in the case of powder cocaine, four) bins, which we refer to as *quantity levels*, based on the amount involved in the transactions. Specific cutoff points are determined on the basis of two objectives: (1) trying to find reasonably round transaction amounts that appear relevant in the data, and (2) trying to retain a large number of observations in each bin (to assist with estimation of the empirical models). Table 3 identifies the cutoff points for each drug and the number of observations included at each quantity level. Plots showing the frequency of specific amounts used to examine the reasonableness of these definitions are presented in Appendix A.

Table 3. Market Quantity Levels, by Drug, for Price/Purity Model Sample

Quantity Level	Amount in Grams	Number of Observations	Percent of Observations
Powder cocaine			
1	AMOUNT<=2	6,345	14.0
2	2< AMOUNT<=10	7,807	17.2
3	10<AMOUNT<=50	18,979	41.8
4	AMOUNT>50	12,292	27.1
Crack cocaine			
1	AMOUNT<=1	13,844	29.8
2	1< AMOUNT<=15	17,006	36.6
3	AMOUNT>15	15,606	33.6
Heroin			
1	AMOUNT<=1	13,294	47.8
2	1< AMOUNT<=10	7,552	27.2
3	AMOUNT>10	6,951	25.0
d-Methamphetamine			
1	AMOUNT<=10	3,565	29.1
2	10< AMOUNT<=100	5,487	44.9
3	AMOUNT>100	3,180	26.0
Marijuana			
1	AMOUNT<=10	2,281	49.6
2	10< AMOUNT<=100	846	18.4
3	AMOUNT>100	1,470	32.0

Source: System to Retrieve Information on Drug Evidence (STRIDE).

IV. Price and Purity Models

Methodological Changes

A number of methodological changes have been made for the price and purity econometric models. The most significant modification is the adoption of the Expected Purity Hypothesis (EPH), which develops an empirical model of price based on the assumption that it is the buyer's perception of purity at the time of the transaction, not actual purity of the drug, that determines the price he or she is willing to pay for the drug. Illicit drugs are what economists refer to as "experience goods"; purchasers often cannot readily assay the quality of the drug purchased until it is consumed, which generally occurs after a price is negotiated and the deal is completed. Hence, it is typically not the actual purity of the drug that governs the negotiated price at the time of the transaction, but rather the supposed or expected purity of the drug. For example, one might observe that most transactions of a particular drug at a particular time, place, and transaction size are 60 to 80 percent pure, but that a minority have very low or even zero purity, even though the price paid for these very-low-purity drugs is not noticeably lower. The view implicitly adopted by past statistical methods was that purchasers of these low-purity observations were knowingly paying much more, sometimes ten or more times as much, per pure gram than were most customers. The view implicit in the EPH models is that these customers were ripped off; they paid a price typical of 60 to 80 percent pure transactions because they thought or expected that they were buying drugs that were 60 to 80 percent pure. In the EPH model, these low-purity transactions are not discarded; they represent a real cost to customers. Therefore, they are incorporated into expectations of the pure quantity contained in purchases, on average, rather than assuming that they represent fully informed purchases.[14]

The adoption of the EPH has two important implications for the way the data get analyzed. First, observations with low purity, but not zero purity, are retained in the analysis, provided they meet other general criteria for inclusion. Second, price is estimated through a two-step procedure where expected purity rather than actual purity is included in the price regression model. Expected purity is the predicted value obtained from a first-stage regression where actual purity is estimated as a function of all other observable information available to the buyer and included in the data (amount, city, quarter, year). Because expected purity is far less volatile than actual purity, the EPH model generally produces smoother price series, even when relatively fewer data points are available (e.g., when estimating prices for a specific city). Failing to use the EPH model can either inflate or suppress the estimated price level somewhat, depending on the details of the distribution of purities observed and whether and how many low-purity observations are discarded. Thus, it is not appropriate to compare the level of prices produced by an EPH model with that produced using a non-EPH method.

[14] This framework for estimating price series from STRIDE has been widely adopted in the economics literature. See, for example, Caulkins, J.P. (1994), *Developing Price Series for Cocaine*, MR-317-DPRC, RAND, Santa Monica, CA; Saffer H., and F.J. Chaloupka (1995), "The Demand for Illicit Drugs," National Bureau of Economic Research Working Paper #5238, August; Grossman M., F.J. Chaloupka, and C.C. Brown (1998), "The Demand for Cocaine by Young Adults: A Rational Addiction Approach," *Journal of Health Economics*, Vol 17, No. 4, pp. 427–474; DeSimone, J. (2002), "Illegal Drug Use and Employment," *Journal of Labor Economics*, 20(4), pp. 952–977.

A related change is the use of inflation-adjusted price, not inflation-adjusted price per pure gram, as the dependent variable in the statistical regression models. Because quantity levels are defined in terms of grams unadjusted for purity, it no longer makes sense to estimate the models in terms of the inflation-adjusted price per pure gram. Further, including actual (or expected) purity in the denominator of the dependent variable (i.e., estimating price per pure gram) causes the coefficient on the actual (or expected) purity variable on the right-hand side of the equation to be biased and hence will generate biased predictions.

A third methodological improvement is that both the purity and price models are estimated using hierarchical modeling (HM), which offers at least three principal advances over the previous methodology employed. First, it directly accounts for the nested nature of the data being used and adjusts standard errors and variance-covariance matrices to account for the fact that specific clusters of observations (in this case, observations from the same city) are not entirely independent. The error terms across observations from within a city are allowed to be correlated to account for city-specific unmeasured components of price (or purity). Second, it adjusts the variance-covariance matrix to account for unequal variances in error terms across different cities, which could result because of different unobservables that exist across cities and different sample sizes across cities. Finally, HM is highly flexible and allows each city to have unique relationships between price and the other independent variables. The methodology employed in previous reports allowed the price levels to differ from city to city (through a city-specific intercept term), but the relationship between other variables, such as the amount of the transaction, and price was assumed to be constant across all cities. The interpretation of this restriction is that within a specific market level, quantity discounts across cities are all the same. This assumption is likely to be overly restrictive. With HM, the relationship between price and amount (or any other independent variable) can vary across cities and over time. The specific form of the HM model employed here is a random coefficients model.

Price/Purity Model Specifications

We estimate the price/purity model for each distribution level for each drug. As described above, the model is estimated using a two-step procedure that involves first estimating purity. The purity model is implemented without zero-potency observations. The previous report estimated the purity model as a logistic model because the dependent variable should be constrained to be between 0 and 1. However, using a simple linear specification of the random coefficients model, very few cases have an estimated purity that exceeds those bounds. Thus, a linear specification of the model is retained here. In addition, in contrast to the previous report, the amount of the transaction (measured as weight in grams) is included as an additional regressor.

The empirical specification of the random coefficient purity model can be written as

$$Purity_{ijk} = \alpha_{0k} + \alpha_{1k} time_{ij} + \alpha_{2k} AMT_{ijk} + \varepsilon_{ijk} \qquad (1)$$
$$\alpha_{0k} = \gamma_0 + u_{0k}$$
$$\alpha_{1k} = \gamma_1 + u_{1k}$$
$$\alpha_{2k} = \gamma_2 + u_{2k}$$
$$\varepsilon_{ijk} \sim N(0,\sigma^2)$$

$$\begin{pmatrix} u_{0k} \\ u_{1k} \\ u_{2k} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{pmatrix} \right)$$

where $time_{ij}$ is a vector of dummy variables representing a year-quarter (i.e., 90 quarters over 22.5 years) and AMT_{ijk} is the raw weight of the i^{th} 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 specified above. The interpretation of the slope coefficients in a random coefficient model is similar to that in other OLS models; α_{1k} is a vector of coefficients identifying time (year-quarter) effects for each city *k*. Test statistics used to evaluate the appropriateness of this functional form versus a simpler random-intercept model are provided in Appendix B.

Equation (1) is estimated as a general linear random effects model, and the estimates of expected purity generated from this model are then used in the second stage price model. For the few cases in which the predicted purity was below 0 percent or above 100 percent, we modified the prediction to equal either 0.5 or 99.5 percent, respectively.

Price is modeled using another general linear random effects model. This model was described in McCullagh and Nelder (1989) and used in ONDCP (2001).[15] The equation is the following:

$$E(\text{real price}_{ijk}|\gamma_{0k}, \gamma_{1k}, \gamma_{2k}) = \exp[[\gamma_{0k} + \gamma_{1k}\text{time}_j + \gamma_{2k}[\ln(\text{AMT}_{ijk})+\ln(\text{predicted purity}_{ijk})] \quad (2)$$

$$\gamma_{0k} = \beta_0 + \varepsilon_{0k}$$
$$\gamma_{1k} = \beta_1 + \varepsilon_{1k}$$
$$\gamma_{2k} = \beta_2 + \varepsilon_{2k}$$

$$\begin{pmatrix} \varepsilon_{0k} \\ \varepsilon_{1k} \\ \varepsilon_{2k} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{pmatrix} \right)$$

$$\text{Var}(\text{real price}_{ijk}|\gamma_{0k}, \gamma_{1k}, \gamma_{2k}) = \lambda^2 [E(\text{real price}_{ijk}|\gamma_{0k}, \gamma_{1k}, \gamma_{2k})]^2 \quad (3)$$

In Equation (2), the real price for observation *i* in period *j* in city *k* is estimated as a function of time effects, city effects, and the sum of the natural logarithm of amount and the natural logarithm of expected purity. Note that this last term is just a long-form way of specifying the price in terms of the natural logarithm of expected pure grams.[16] Average mean effects of these control variables on price are captured through the coefficients β_0, β_1, and β_2. The coefficients γ_{0k}, γ_{1k}, and γ_{2k} are the city-specific intercept, time, and ln(expected pure grams) coefficients, respectively, each of which comes from a common normal distribution centered at 0. We estimate Equation (2) as a hierarchical generalized linear model with a log link function, a gamma error structure, and a constant coefficient of variation. This model transforms the dependent variable to the desirable log form during estimation, which essentially puts the coefficient estimates in percentage terms. Equation (3) shows the conditional variance function of this model, where λ is the coefficient of variation of the real price.[17] Additional test statistics that provide information on the appropriateness of the functional form employed here versus a simpler random intercept model are provided in Appendix C.

Summing the natural logarithms of amount and predicted purity gives us identification in the model, which is necessary because all of the explanatory variables in the purity model of Equation (1) are in the price model of Equation (2). Furthermore, such summing provides a specification of price in which price is a function of the expected pure grams involved in the transaction.

---

[15] (1) McCullagh, P., and J.A. Nelder (1989), *Generalized Linear Models*, Second Edition, Chapman and Hall, London, Chapter 8; Office of National Drug Control Policy (2001), "The Price of Illicit Drugs: 1981 through the Second Quarter of 2000," Washington, DC; Littell, Ramon C. , George A. Milliken, Walter W. Stroup, and Russell D. Wolfinger (1996), *SAS System for Mixed Models*', SAS Institute, Inc., Cary, NC, Chapter 11.

[16] In other words, [(ln AMT) + ln (expected purity)] = ln (AMT*expected purity). When we multiply the weight by expected purity we are generating an estimate of the weight in expected pure grams.

[17] McCullagh P. and Nelder J.A. (1989). Generalized Linear Models, Second Edition. Chapman and Hall, London, p. 285.

As discussed previously, transactions that deviate significantly from the norm could have large effects on the coefficient estimates and the predicted prices. These deviations could occur, for example, from miscoded data or poor bargaining on the part of a DEA agent. While the gross-outlier restrictions eliminate many of these observations from the sample as a whole, some still survive because they are gross outliers for a time period or a specific quantity level (even if they are not gross outliers for the full period or all quantities combined). To reduce the potential influence of these outliers, extreme-residual analysis is performed following criteria employed in previous reports. Operationally, the model is estimated and the residuals from the model are kept, standardized, and plotted. Using the same criterion as that employed in previous reports, observations that fell beyond 3.09 were deleted (so the probability of deleting good data is set to 0.002). This process of reestimating the model and deleting residuals is continued until no extreme residuals exist in the sample. Most of the models lost between 3 and 5 percent of the sample, with the smallest loss being the third quantity level for d-methamphetamine (2.2 percent) and the largest loss being the first quantity level for marijuana (7.4 percent).

Because these models are highly parameterized and separate models are run for specific quantity levels that are determined after gross outliers have been deleted, it is difficult in some cases to achieve convergence. Problems with convergence could be due to the data not fitting well in our imposed functional form, or they could be caused by outliers in a city with few observations, which would make it difficult to identify a city-specific slope coefficient. To assist in convergence, a simpler second stage price model is estimated to identify the first round of extreme residuals for each quantity level. These simpler models allow for only a random intercept, so slope coefficients are forced to be constant across cities. With the reduced parameterization of the model, the models quickly converge, and it is possible to identify many extreme residuals. Once these first round extreme residuals are dropped, the full random coefficients model expressed in Equations (2) and (3) is estimated and convergence is obtained for all quantity levels. Future work should explore the feasibility of model alternatives to accommodate or downweight the influence of extreme residuals within a particular quantity level.

Marijuana Price Model Specification

A modified model is developed for marijuana because no information on marijuana purity is available and because the smaller number of observations makes it difficult to fit the highly parameterized model developed for the other drugs.

The marijuana model specification is as follows:

$$E(\text{real price}_{ijk}|\gamma_{0k},\gamma_{1k},\gamma_{2k}) = \exp[\gamma_{0k} + \beta_1 \text{quarter}_{ij} + \gamma_{1k}\text{year}_{ij} + \gamma_{2k}\ln(\text{AMT}_{ijk})] \qquad (4)$$

$$\gamma_{0k} = \alpha_0 + \varepsilon_{0k}$$
$$\gamma_{1k} = \alpha_1 + \varepsilon_{1k}$$
$$\gamma_{2k} = \alpha_2 + \varepsilon_{2k}$$

$$\begin{pmatrix} \varepsilon_{0k} \\ \varepsilon_{1k} \\ \varepsilon_{2k} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{pmatrix} \right)$$

$$\text{Var}(\text{real price}_{ijk}|\gamma_{0k},\gamma_{1k},\gamma_{2k}) = \sigma^2\, [E(\text{real price}_{ijk}|\gamma_{0k},\gamma_{1k},\gamma_{2k})]^2 \qquad (5)$$

There are two primary differences between this model and the model described Equation (2). First, because no information is available on the purity of marijuana, price is estimated in terms of actual amount and not expected pure weight (i.e., the log(AMT)). Second, due to the significantly smaller sample size, it is not possible to estimate individual city-specific interacting time trends for quarters and years as we did in the previous models. Instead, the model includes year trends (captured through a series of dichotomous indicators for each year) and quarter trends separately. Further, only the year effect varies randomly across cities; the influence of specific quarters is fixed across all years and cities. Finally, for the first quantity level of marijuana, the specification further restricts year to be constant across cities, so γ_{1k} is set equal to just α_1 in that model only. The strategy of deleting extreme residuals employed in the previous models is also used here.

V. Empirical Results

The regression models for each distribution level for each drug produce thousands of coefficient estimates. In the following tables, we present only the group average effect of amount (AMT) and, in the case of the price model, expected pure grams as well as their corresponding covariance estimates. When the covariance estimates are statistically significant, it means that there is significant variation across cities in the relationship between the dependent variable (price or purity) and the specific variable of interest.

Purity Equation Results

Table 4 presents coefficient estimates and covariance estimates from the purity equations. The first column identifies the specific drug and distribution level being evaluated. The second column indicates the number of observations used to generate the estimate. The third column shows the group average effect of amount (i.e., weight in grams) on purity and thus indicates the relationship between purchase weight and purity within the given distribution level. The fourth column shows the covariance of the random effect of amount on purity across cities and thus indicates the extent to which the relationship between amount and purity varies across locations. Note that this final column does not pertain to variation in the average purity across cities, which is often considerable, nor does it pertain to variation across cities in the extent to which purity differs across market levels. Rather, it shows variation across cities in the extent to which purity varies with purchase weight within the given market level.

Column three of Table 4 shows that higher amounts of powder cocaine and heroin are associated with greater purity at the two highest quantity levels. This would be consistent with dilution/adulteration taking place within these levels. There is no significant relationship between amount and potency at the lower quantity levels for these two drugs. For crack cocaine and d-methamphetamine, there is generally a negative and statistically significant relationship between amount and potency. This is consistent with the observed overall relationship between purchase weight and purity across quantity levels for these drugs, particularly in recent years.

For most substances, the coefficient estimates are smaller for higher quantity levels. At the same time, the range in amounts is greater at these higher levels. For example, amounts of powder cocaine vary by only 8 grams within quantity level 2 (specifically, between 2 and 10 grams), whereas they vary by 40 grams within quantity level 3 (between 10 and 50 grams). The product of the coefficient estimate on amount and the range on amount within the quantity level is indicative of the model's predicted purity variation within that quantity level. In nearly all cases, it is less than 10 percentage points. Thus, the relationship between transaction weight and purity is often statistically significant primarily because there are so many data points, not because the magnitude of the relationship is so extremely large. There do not appear to be "cliffs" marking abrupt changes in purity within any of the quantity levels defined for these substances.

Furthermore, the covariance estimates provided in column four of Table 4 indicate that many of the purity models have insufficient variation in amount across cities to estimate random effects, particularly at the highest quantity levels. In all but two cases (crack cocaine, quantity level 2, and heroin, quantity level 2), the covariance estimates are statistically insignificant. Even in those cases where the covariance estimate is statistically significant, the covariance estimates are fairly small. This suggests that the relationship between amount and purity across cities is fairly stable.

Price Equation Results

Table 5 shows selected results from the price equations, the second stage of the price/purity model. For each drug and each quantity level, the average mean effect of the log of expected pure grams on price is positive and statistically significant. Furthermore, the average mean effect of expected pure grams on price is relatively stable, falling between 0.7 and 0.8 for all but a few drugs' quantity levels. In the case of powder cocaine and heroin, the same percentage change in expected pure grams generates an even larger increase in price at larger amounts, as shown by the larger average mean effects at higher quantity levels. The results are less consistent for crack cocaine and d-methamphetamine.

Because there are price markups with distance down the distribution chain, repackaging a certain quantity into smaller bundles and selling those smaller bundles increases the market value of the original quantity. For example, if one could buy a gram of a particular drug for \$100, divide it into eight packages that each contain 1/8 gram, and sell those eight packages for \$20 each, that repackaging and resale would increase the market value by 60 percent, from \$100 to 8 × \$20 = \$160. The regression coefficients in Table 4 quantify the increase in market value that results from repackaging and resale and allow comparisons to be made across market levels and drugs. For simplicity, these quantifications assume that there is no change in any of the other independent variables (e.g., the smaller packages are sold in the same city and time period as the original quantity would have been) and that the change in quantity and value occurs within a given market level.

Table 4. Selected Results from the First Stage Purity Regressions

Drug and Quantity Level	Number of Observations	Coefficient Estimate on AMOUNT	Covariance Estimate on AMOUNT
Powder Cocaine			
1	6,056	-0.0092	0.0002
		(0.0067)	(0.0003)
2	7,765	0.0022	1.2E-5
		(0.0014)	(1.3E-5)
3	18,894	0.0011***	0
		(0.0002)	(--)
4	12,198	5.5E-5***	0
		(0.9E-5)	(--)
Crack Cocaine			
1	12,609	-0.0290***	0
		(0.0064)	(--)
2	16,376	-0.0080***	3.2E-6**
		(0.0005)	(1.8E-6)
3	15,147	-2.3E-4***	0
		(0.4E-4)	(--)
Heroin			
1	12,865	-0.0223	0.0079***
		(0.0177)	(0.0027)
2	7,482	0.0154***	1.7E-4***
		(0.0026)	(0.5E-4)
3	6,915	2.3E-4***	0
		(0.4E-4)	(--)
D-Methamphetamine			
1	3,429	-0.0108***	5.7E-5
		(0.0029)	(5.7E-5)
2	5,446	-0.0007**	0 (--)
		(0.0003)	
3	3,101	-8.1E-6	0 (--)
		(9.1E-6)	

Note: Standard errors are in parentheses.
*** indicates statistical significance at the 1 percent level
** indicates statistical significance at the 5 percent level
* indicates statistical significance at the 10 percent level

Table 5. Selected Results from the Second Stage Price Regressions

Quantity Level	Number of Observations	Coefficient Estimate on Log(Expected Pure Grams)	Covariance Estimate on Log(Expected Pure Grams)
Powder Cocaine			
1	6,122	0.716***	0.0033**
		(0.014)	(0.0018)
2	7,543	0.751***	0.0012*
		(0.010)	(0.0007)
3	18,399	0.787***	0.0012***
		(0.007)	(0.0004)
4	11,889	0.813***	0.0005***
		(0.005)	(0.0002)
Crack Cocaine			
1	13,165	0.731***	0.0052**
		(0.017)	(0.0024)
2	16,393	0.661***	0.0041***
		(0.012)	(0.0012)
3	15,038	0.833***	0.0007***
		(0.006)	(0.0002)
Heroin			
1	12,711	0.531***	0.0281***
		(0.029)	(0.0074)
2	7,219	0.718***	0.0228***
		(0.027)	(0.0060)
3	6,664	0.764***	0.0026***
		(0.012)	(0.0010)
D-methamphetamine			
1	3,426	0.707***	0.0050**
		(0.018)	(0.0029)
2	5,196	0.796***	0.0074***
		(0.021)	(0.0029)
3	3,060	0.663***	0.0115***
		(0.026)	(0.0040)
Marijuana			
1	2,112	0.573***	0.1162***
		(0.066)	(0.0401)
2	815	0.802***	N/A
		(0.025)	
3	1,432	0.783***	0.0012**
		(0.013)	(0.0006)

Note: Standard errors are in parentheses.
*** indicates statistical significance at the 1 percent level
** indicates statistical significance at the 5 percent level
* indicates statistical significance at the 10 percent level

The model suggests that repackaging and reselling a quantity of powder cocaine into K = 8 equal-size smaller packages increases the market value by 44 to 51 percent when done at quantity level 4, 51 to 60 percent at quantity level 3, 61 to 75 percent at quantity level 2, and 70to 91 percent at quantity level 1.[18] Thus, for a given size step down the distribution chain, the markup is greater further down the distribution chain. Conversely, the size (in percent) of the quantity discount for buying K times as much is greatest at the lower market levels. This same basic pattern holds for the other substances, with certain exceptions, specifically, larger than expected quantity discounts at crack cocaine, quantity level 2, and d-methamphetamine, quantity level 3. Markups are also much larger for heroin and marijuana at quantity level 1 than at quantity level 2 (the expected direction but greater difference than for other substances). Apart from these specific exceptions, the price markups are all roughly similar at comparable quantity levels, with point estimates ranging from 42 to 63 percent at the highest quantity level and 75 to 84 percent at the lowest quantity level. Again, these specific figures are for K = 8; for different values of K, the percentages will be different but will still display the same relationship across drugs and quantity levels.

The covariance estimates of the group average effect of expected pure grams on price are all positive and statistically significant except in the case of marijuana, quantity level 2, for which there was either insufficient sample size or insufficient variation in the small sample to calculate random effects. The clear conclusion from these estimates, however, is that cities vary substantially in the quantity discounts offered at every quantity level. This suggests that it is important to estimate quantity discounts by city, rather than arbitrarily assuming the rates are the same around the country either directly by not allowing the relevant regression coefficients to vary by location or indirectly by pooling data from many cities in a single regression. Further, the differences across cities appear to be larger for some drugs (e.g., heroin) than for others (e.g., powder cocaine or crack cocaine).

Estimates from these models are used to generate city-specific predicted price and purity estimates for each drug and quantity level in every quarter possible from 1981 through the second quarter of 2003. To generate each prediction, the model requires that potency and amount be specified at a certain value. In all instances, potency was set to 100 percent. Amounts varied depending on the quantity level being estimated. In all cases, a round number close to the median value within a specific quantity level was used (see Table 6). Estimated annual price and purity values for specific cities shown in the main report (Tables B.1 through B.8) are reported in Appendix D. We estimate these city-specific price and purity series for the lowest quantity level for all drugs except d-methamphetamine, for which we use the middle quantity level, due to low geographic distribution at the lowest quantity level.

One important step taken *ex post* the regressions is the deletion of the predicted annual price for d-methamphetamine in the third quantity level for 1987. In 1987, there were only five observations on which to base this estimate, all of which were in the first quarter and four of which were from Baltimore on February 2. The observations from Baltimore all share the same extremely large nominal price ($17,200) and have ranges of amounts and purities that are amazingly close (amount ranged between 102 and 105 grams, and purity ranged between 5.4 and 8.2 percent). It was concluded that these four records were likely to be duplicates that were not

[18] Ranges based on the point estimates +/– 2 standard errors.

caught previously because they were not exact duplicates. If they are treated as duplicates, then there are only two observations for 1987, an insufficient amount to estimate the model. Thus, we deleted all five observations for that quarter from the sample. While these nearly duplicate records were easy to identify because of the huge price spike they generated in the model, there are likely to be many other probable duplicates in the data that remain unidentified. Future work should attempt to identify and delete likely duplicates.

Table 6. Evaluation "Amount" in Grams for Each Quantity Level

Drug	Quantity Level 1	Quantity Level 2	Quantity Level 3	Quantity Level 4
Powder cocaine	0.75	5	27	108
Crack cocaine	0.3	5	38	
Heroin	0.4	2.5	27.5	
d-Methamphetamine	2.5	27.5	225	
Marijuana	2.5	26	443	

VI. Constructing the National Price and Purity Indices

Once quarterly estimates for the price of a given drug and quantity level in each city are calculated, the national indices for the drug-quantity level are constructed.[19] There are a number of different methods that could be used to generate this index, depending on the assumptions one is comfortable making. The current report uses a methodology employed in previous reports and constructs the index as a weighted average of the prices predicted for each city. However, new weights are used.

The ideal statistic to weight each city-specific average would perhaps be estimates of the number of drug transactions in each city. Unfortunately, the geographical distribution of transactions in the STRIDE data is not representative of the geographical distribution of actual transactions across the country. And there are no existing data that provide accurate estimates of the extent of drug use in the cities of interest for the current analysis. Thus, other alternatives had to be considered.

The weights used in the previous report were based on emergency-room (ER) mentions of the particular drugs in the Drug Abuse Warning Network (DAWN) data. There are several drawbacks to using these data for weights. First, eight of the 29 cities are not represented in DAWN, so their weights must be based on what the analyst considers to be similar cities. For example, Seattle was used as a proxy for Portland, adjusted for population. Second, there are limitations in the DAWN data that raise some doubt about their appropriateness. For example, the number of drug-specific ER mentions within particular cities is very volatile, causing the implied weight constructed from them to be volatile. For cocaine, the weights for New York range from 6.8 percent to 19.2 percent over the 1981–2000 time period. For the “other” category (i.e., those locations not falling in one of the 29 cities), the weight ranged from 2.9 to 47.0 percent. There are similar examples of extraordinary volatility for the other drugs as well. Finally, it is not entirely clear that cross-sectional differences in the number of hospital-related drug episodes are truly reflective of differences in drug use across cities. For example, based on only DAWN data, cocaine use in St. Louis in 2002 would appear to be greater than that in Washington DC. State-level estimates of cocaine use from the 2001 National Household Survey on Drug Abuse suggest that this is unlikely to be the case.

This analysis uses the population of cities and regions as the weights because population data are readily available for all cities, not just those in which DAWN is estimated, and because population estimates are reliable and universally understood. The estimates come from the Census estimates for 1980, 1990, and 2000 reported in *Statistical Abstract.*[20] *Statistical Abstract* population estimates for small cities (e.g., Buffalo) tend to include only the MSA, while larger cities (e.g., Boston) usually include the PMSA and/or the CMSA. Because the CMSAs

[19] We recognize that developing a national average from these relatively sparse and unrepresentative data is not advisable for a number of reasons. We do not, therefore, interpret the numbers calculated as true national averages. Instead, we view them as potential indices, the validity of which will be tested by examining their correlation to external drug-related data obtained through more reliable sources.

[20] Department of Treasury (2002), *The Statistical Abstract of the United States*, Government Printing Office, Washington D.C.

incorporate several PMSAs, this analysis uses only estimates reported for the PMSA. The one exception is San Francisco, where the population estimates include both the San Francisco and Oakland PMSAs because we group drug transactions from these PMSAs together in the San Francisco city group. Populations for the Census divisions are calculated by summing the populations for the states in each division and subtracting the populations of the MSAs in a Census division that are already included in the model. This will not be a perfect measure because some MSAs include counties in other states. However, it does represent a reasonable approximation, given that the error is consistently applied in each year.

Using data from the decennial population estimates for each city (and Census division), a linear (constant) growth trend is presumed for all years between 1980 and 1990 and then again for 1990 through 2000 so that city (or region) population estimates can be obtained for every year. The constant growth rate identified for 1990–2000 is also applied to 2001–2003 to generate population estimates for these years. Weights for the cities are constructed by determining in each year the fraction each city's population constitutes of the total 29 cities' population. The 29 cities sum up to 29/30 of the total weight in the analysis. The other 1/30 is partitioned across the 9 Census divisions, whose weights were determined by their share of the total population across regions in a given year. More weight is given to the cities than to the regions for several reasons. First, this weighting is consistent with what was done in previous reports and is based on the assumption that drug transactions are more concentrated in larger cities than in small cities or rural areas. Second, it seems reasonable to presume that more useful information can be obtained from multiple observations within the same small geographic location (a city) than from observations obtained from a geographically large location (e.g., a Census region). Actual weights used to weight city-specific averages are given in Table E.1 in Appendix E.

VII. Accounting for Uncertainty

The point estimates of price and purity reported in the main report were determined through a modeling process that is based on numerous assumptions that are not known with certainty. For example, the point estimates reflected in the main report may be biased because the STRIDE data may not be representative of the true distribution of price and purity. Thus, it is not known whether the model used here provides a good fit to the actual distribution of prices and purities in the country. All we can do is determine whether the model provides a relatively good fit of the distribution of these quantities observed in STRIDE, as was done in Appendices B and C. And, as is the case with any model, the model could lead to biased estimates if there are omitted variables or if there is model misspecification that is not easily identified from the STRIDE data. It is not possible for us to evaluate whether these sources of bias exist in our model estimates without using other data sources, a task that was outside the scope of the current project. However, it is strongly recommended that such an evaluation be conducted as a next step for future analyses using these data.

Even if the point estimates from the model could be shown to be asymptotically unbiased, a second source of uncertainty is the precision with which the coefficient estimates are determined, in light of the small samples. Every coefficient estimate has an associated standard error describing the precision with which it is being estimated, and this precision is not currently being accounted for in our estimates of the range of predicted prices. Estimation of this variability is a nontrivial task, given the complexity of the models being estimated, unless one is willing to impose strong distributional assumptions that cannot be empirically validated. There are computationally intensive methods, such as bootstrapping or Markov Chain Monte Carlo methods, that could be employed to construct more-precise intervals for the standardized predicted prices as well as the national index estimates themselves.[21] An example of these methods is presented below.

A third source of uncertainty is the estimate of the national price (purity) index. As discussed previously, the models identify independent variation in the distribution of prices (purities) within cities. This unique variation and the limited observation of prices (purities) across all cities in all periods imply that there is uncertainty in the national indices that are calculated. To try to represent some of the variation in the underlying distribution of weighted average indices, the main report presents estimates of the interquartile range of the city-specific prices (or purities) that constitute the national index. The interquartile range is used because it is easily constructed without assumptions regarding the underlying distribution of values and it is fairly easy to understand. The interquartile range simply reflects the highest and lowest values between which half of all the possible values fall. One-quarter of the remaining values are lower than the estimates provided by the interquartile range, and the other quarter are higher.

To generate estimates of the interquartile range, we randomly sample a city based on the city's population weights for a given year. For the sampled city, we take the standardized predictions

[21] For bootstrapping techniques, see: Efron, B., and R. Tibshirani (1993), "An Introduction to the Bootstrap," Chapman & Hall Ltd., London, New York. For Markov Chain Monte Carlo techniques, see Gelfand, A.E., and A.F.M. Smith (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85 , pp. 398–409.

for price and purity for the year. Based on sampling 5,000 city-standardized price and purity measures for each year for a given distribution level and drug, we identify the 25th, 50th (median), and 75th percentile price and purity observations. The 25th and 75th percentile observations provide the boundary of the interquartile range, which is reported for all the price and purity indices in the main report.

It should be emphasized that this interquartile range captures just a small part of the overall uncertainty in the estimates. It relies on the assumption that the price and purity measures for each city and each quarter and the model generating them are unbiased and precise. Future work should build upon this by estimating uncertainty in the price and purity values themselves, which means evaluating the other potential sources of uncertainty as well. To illustrate, we fit a Bayesian version of the hierarchical model using Markov Chain Monte Carlo (MCMC) to obtain the posterior distribution of all the model parameters as well as the national price (or purity) index based on very minimal assumptions of the prior distribution of these estimates. This approach involves first specifying a likelihood function for price (or purity), as was done before (Equations (1) through (5)) and then specifying prior distribution functions for all parameters in the model. The prior and the likelihood functions are multiplied to yield a posterior distribution for the model parameters; all statistical inferences are drawn from this posterior distribution. MCMC is a numerical integration technique for obtaining posterior distributions of interest by simulation. Bayesian hierarchical models were fit for the first quantity level for heroin and powder cocaine.[22] The models were estimated using WinBUGS Version 1.3 software to implement MCMC.[23] A proper but uninformative prior distribution was specified for all of the model parameters[24] and MCMC diagnostics were employed to ensure that the Markov chain had converged to the posterior distribution from its starting value.

Tables F.1 and F.2 in Appendix F compare, for the lowest quantity level for powder cocaine and heroin, the coefficient estimates resulting from the Bayesian hierarchical model with those obtained from a similar model estimated by the method used previously, restricted maximum likelihood (REML). These models differ from those discussed earlier in that they condition on year rather than on quarter, but other than that the likelihood functions are the same. The parameter estimates and their standard errors are remarkably similar across the two methods; the largest visible difference is for powder cocaine purity regression, but the estimates do not differ greatly from each other. The similarity in results despite fundamentally different modeling approaches is reassuring in light of the other problems associated with the data. National price indices were constructed using the predictions generated from the Bayesian hierarchical model data, and the interquartile range for the posterior distribution of the national price index was calculated, as well as the 95 percent posterior probability interval. Figures F.1 and F.2 in

---

[22] At this point, the MCMC for the price models was implemented with predicted purity constructed using the standard techniques in this report instead of the MCMC approach. In addition, this was done with the final sample after having deleted extreme residuals rather than deleting extreme residuals with the MCMC technique.
[23] Spiegelhalter, D., A. Thomas, and N. Best (2000), *WinBUGS Version 1.3 User Manual*, MRC Biostatistics Unit, Cambridge, UK.
[24] Regression parameters had normal prior distributions centered at 0 in all models, with prior variances of 100 in the purity model and 10 in the price model (the difference is due to the fact that purity ranges from 0 to 100, while price is modeled on the log scale); precision (reciprocal of the variance) parameters were assumed to follow Gamma distributions with mean 1 and variances 10 in the price model (e.g., Gamma(0.1,0.1) and 100 in the purity model (e.g., Gamma(0.01,0.01)).

Appendix F show the interquartile range of the posterior distribution, indicated by the 25 percent and 75 percent bands in the figures. These ranges are substantially smaller than those reported using the non-Bayesian methods. The trends in the national price estimates derived from this model, however, are again consistent with those reported previously. The 95 percent posterior probability interval for the national price index is also shown on each figure and is indicated by the lower and upper bounds at 2.5 percent and 97.5 percent, respectively. The advantage of using this type of method to compute an uncertainty estimate for the national price index is that the resulting posterior distributions of the national price indices could be used to test whether the national price index significantly differs at any two given time points.

This estimation merely serves as an example of the type of analysis that could be done to identify some of the uncertainty in the current method if greater time and resources were available. The fact that the model coefficients and predicted trend lines were consistent with what was reported previously is reassuring, but future work should be done to determine whether similar findings hold for the other substances and to test some of the remaining assumptions in the model. For example, the current price model includes a single estimate of expected purity and ignores any uncertainty that might exist with this single estimate. In future work, it would be interesting to test the sensitivity of the results to this assumption by estimating a joint model of price and purity that allows for uncertainty in expected purity.

VIII. Supplemental Analyses

Several supplemental analyses were conducted to test the reasonableness of assumptions made throughout the modeling process. Two analyses that may be of particular interest to users of these data are an investigation of the heroin, salt-undetermined series and an investigation of the importance of observations from the Washington DC Metropolitan Police.

An Examination of Heroin, Salt-Undetermined

As explained previously, there was some concern about constructing a price series for heroin, salt undetermined, given the heterogeneous mix of forms that are included in this category. However, nearly 30 percent of the heroin observations fell into this category, so one might wonder whether the average price for this mix would behave differently from the series constructed for heroin base and hydrochloride. To consider this, we constructed a separate price and purity index for the heroin, salt-undetermined category.

Appendix G contains four charts comparing the price (Figures G.1 and G.2) and purity (Figure G.3 and G.4) of heroin, salt-undetermined with the heroin base sample. The prices for the heroin, salt-undetermined sample (Figure G.2) are generally higher than those for the heroin base sample (Figure G.1), especially prior to 1990. The purity, meanwhile, is much lower for the heroin, salt-undetermined sample (Figure G.4), especially for the highest quantity level (amounts greater than 10 grams). These results suggest that the heroin, salt-undetermined sample follows different trends from those observed for the main series and reinforces the decision to keep them separate.

The Influence of Observations from the Washington D.C. Metropolitan Police Force

Previous research suggests that observations obtained by the Washington DC Metropolitan Police Force (DCMP) follow a different pattern from those obtained by the DEA.[25] Because the DCMP observations represent a large percentage of observations from Washington DC, and in some cases all the observations, there may be concern that their inclusion biases the estimates of the national price and purity indices.

The revised methodology employed in this report should be less susceptible to these sorts of city-specific data differences, as the inclusion or exclusion of observations from DCMP would directly influence only predicted prices in Washington DC and would have an indirect influence on those for other cities through their effect on the DC city-specific estimates. However, if there is a large difference in the DC mean when the DCMP observations are excluded, it is still possible for the overall results to look substantially different, because the random coefficient model borrows explanatory power from other cities when it estimates a city estimate and the global mean.

To explore this, additional models were run excluding observations from the DCMP. For

[25] Horowitz, J. 2001. "Should the DEA's STRIDE Data Be Used for Economic Analysis of Markets for Illegal Drugs?" *JASA*, Vol. 96, No. 456, pp. 1254–1271.

powder cocaine, crack cocaine, heroin, and marijuana, price and purity models were run for only the lowest two quantity levels because these levels accounted for the vast majority of DCMP observations. There was no noticeable difference in the overall level of the national price and purity index at the second quantity level for any of the drugs evaluated except marijuana. However, minor differences did exist for the first quantity level for these drugs; these differences are presented in Appendix H. First, as shown in Figure H.1, only minor differences exist in the level of the national price index for powder cocaine when the DCMP observations are excluded. The similarity in trends is truly surprising in light of the more-measurable differences in the national purity index for powder cocaine (see Figure H.2). When the DCMP observations are excluded from the models, the national purity index is higher between 1985 and 1988 and lower in 1991–1994 than when these observations are included. However, the same broad trends for cocaine powder prices and purities remain, regardless of whether the DCMP observations are included or excluded.

In the case of crack cocaine, the inclusion of the DCMP observations influences the level of the national price index (Figure H.3) but not the purity index (Figure H.4). When the DCMP observations are excluded, the national price index is lower in almost every year, suggesting that the DCMP observations raise the average price per pure gram of crack cocaine in Washington DC in every year. There is virtually no difference in the national purity index, however, when these observations are excluded.

The exclusion of the DCMP observations does not substantially influence either the national price or the purity index of heroin (Figures H.5 and H.6, respectively). There are very small differences in particular years, but these differences do not significantly change the price and purity series. The DCMP observations appear to have the largest impact on the estimated price series for marijuana, which is not particularly surprising, in light of the significantly smaller overall sample size. When DCMP observations are excluded (Figure H.7), the national price index for marijuana, first quantity level, is much more volatile, with large spikes in 1991 and 2002. The national price index for the second quantity level is also sensitive to the exclusion of these observations, but the differences are not nearly as pronounced as those in the first quantity level.

Overall, the supplemental analyses suggest that the inclusion of the DCMP observations in the main analyses does not substantially influence either the level or the trends of the national price and purity indices presented in the main report for powder cocaine, crack cocaine, or heroin. They do, however, influence the national price indices for marijuana, suggesting that the price series for marijuana at the first quantity level is even more volatile than is indicated in the main report.

Appendix A

Distribution of Observations in Terms of Weight in Grams

Figure A.1. Frequency of Powder Cocaine Purchase Observations by Weight, All Quantity Levels

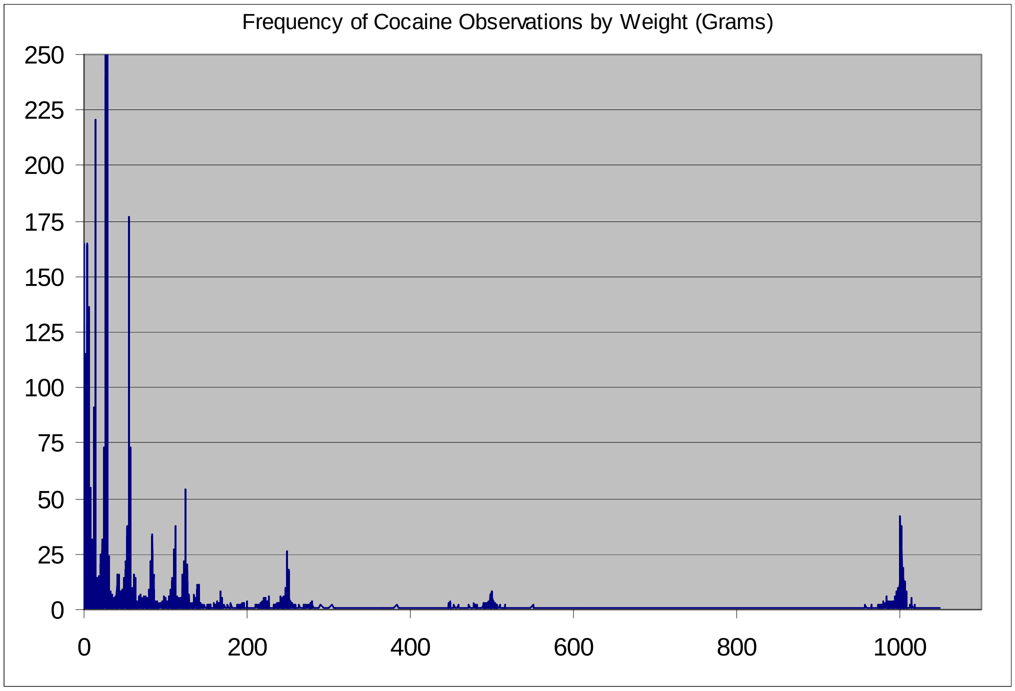


Figure A.2. Frequency of Powder Cocaine Purchase Observations by Weight, Amounts Less than 60 Grams

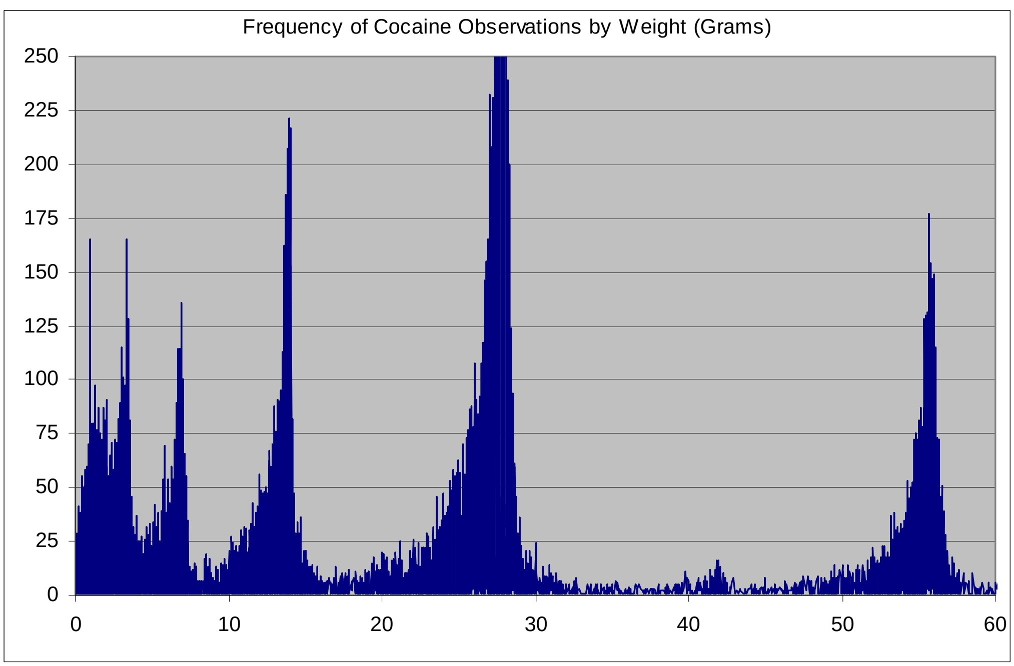


Figure A.3. Frequency of Heroin Purchase Observations by Weight, All Quantity Levels

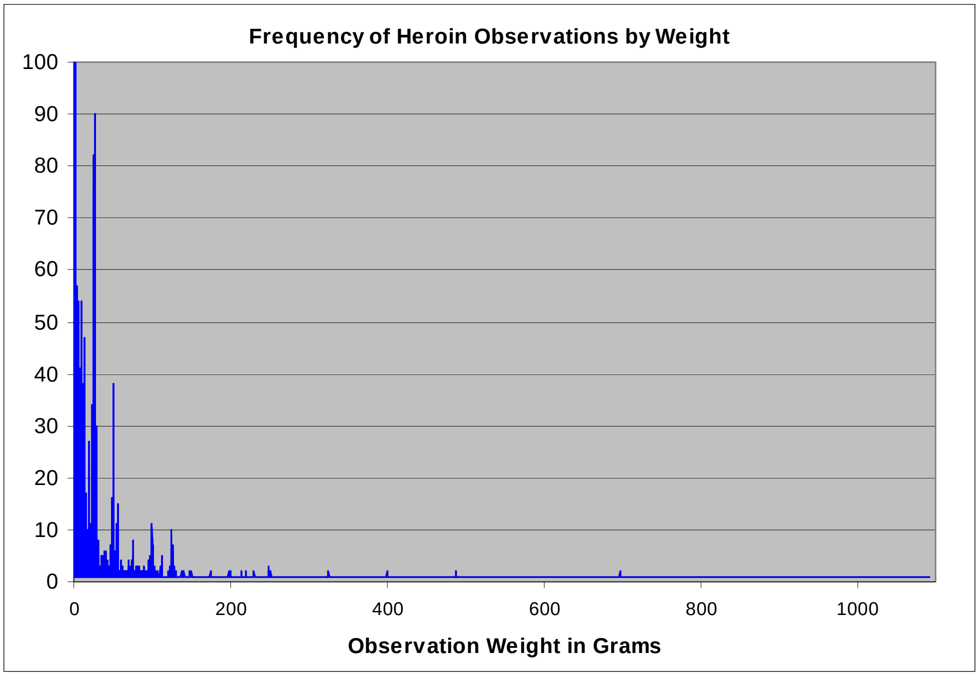


Figure A.4. Frequency of Heroin Purchase Observations by Weight, Amounts Less than 200 Grams

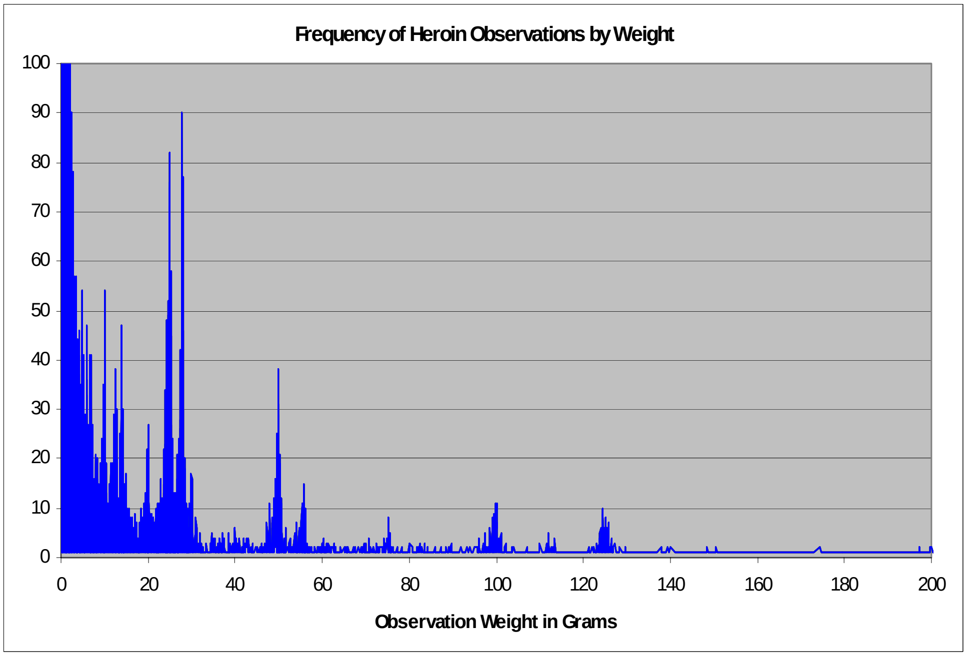


Figure A.5. Frequency of Heroin Purchase Observations by Weight, Amounts Less than 60 Grams

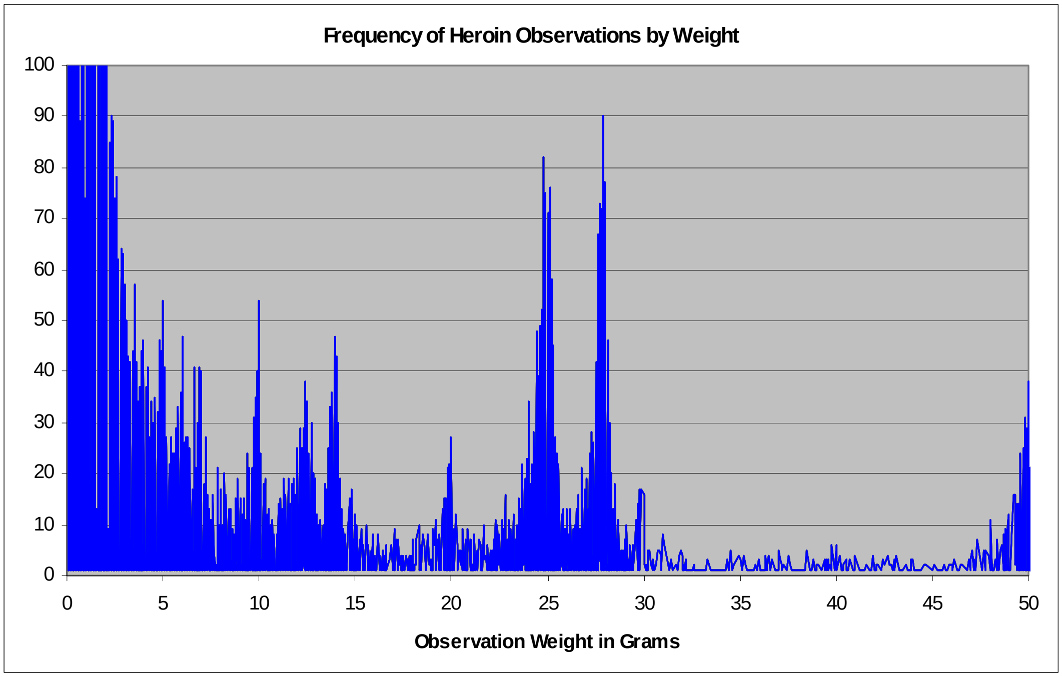


Figure A.6. Frequency of d-Methamphetamine Purchase Observations by Weight, All Quantity Levels

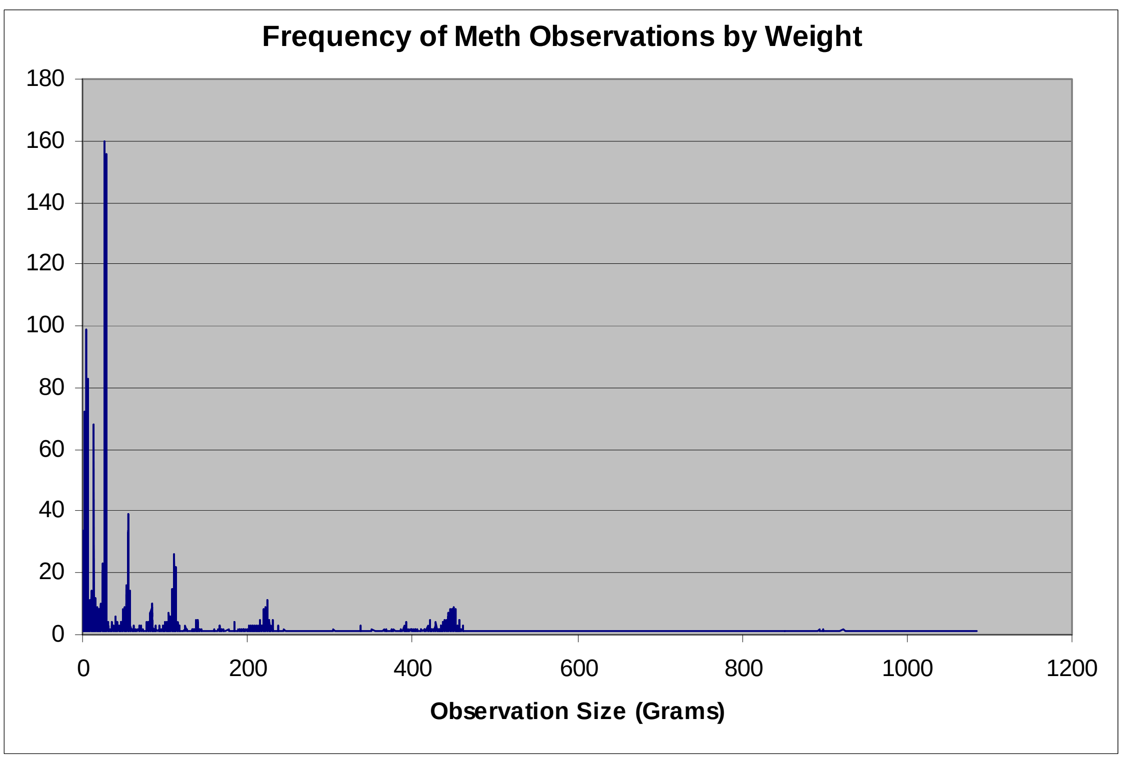


Figure A.7. Frequency of d-Methamphetamine Purchase Observations by Weight, Amounts Less than 60 Grams

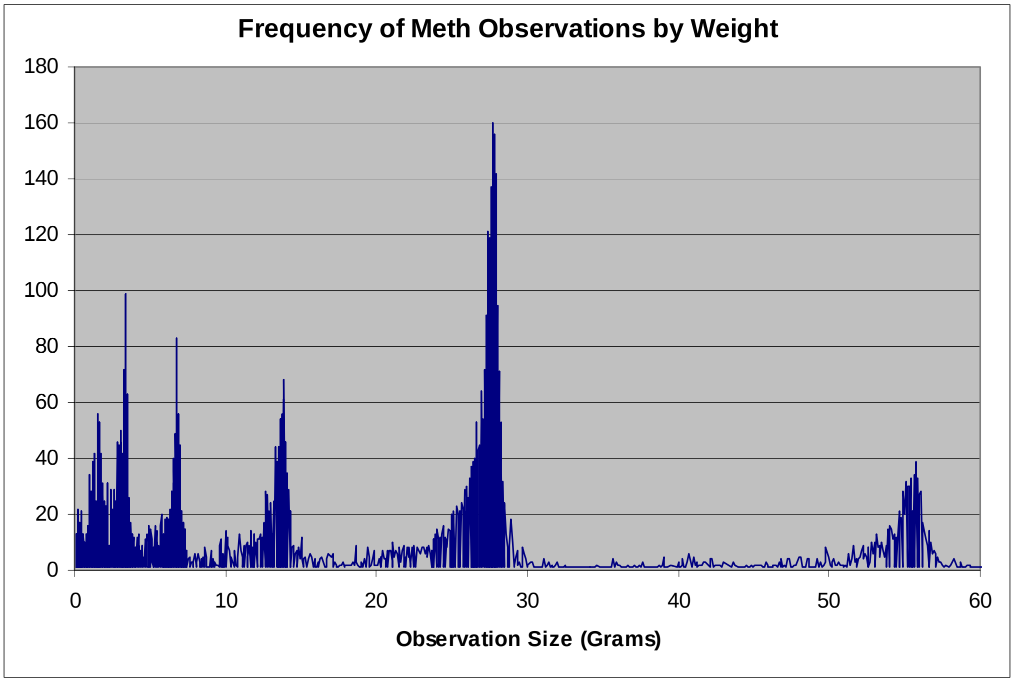


Figure A.8. Frequency of Marijuana Purchase Observations by Weight, All Quantity Levels

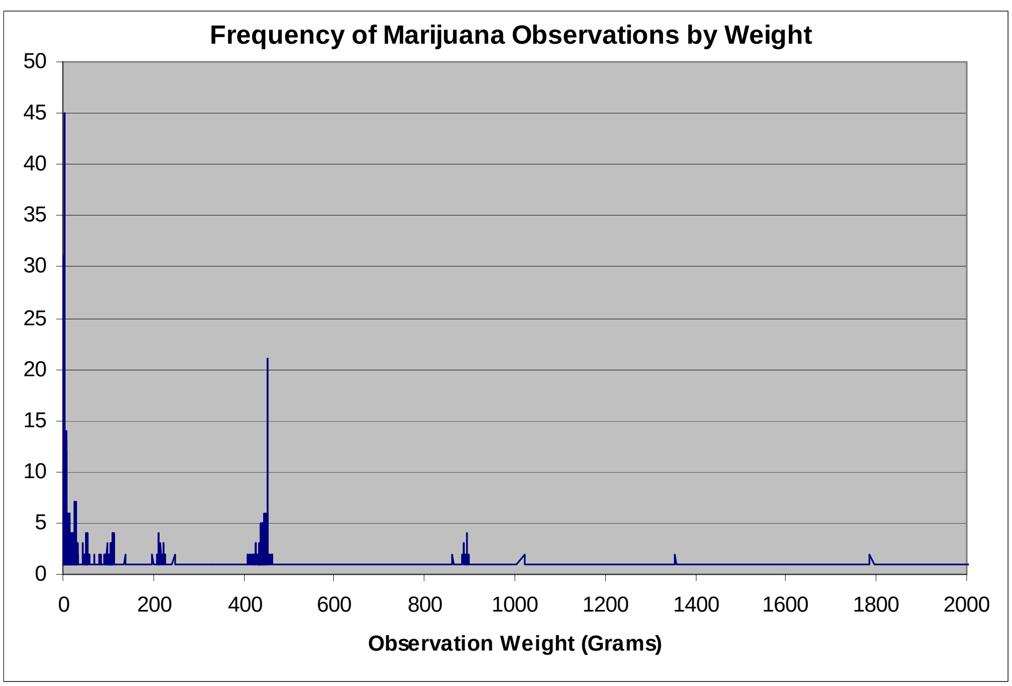
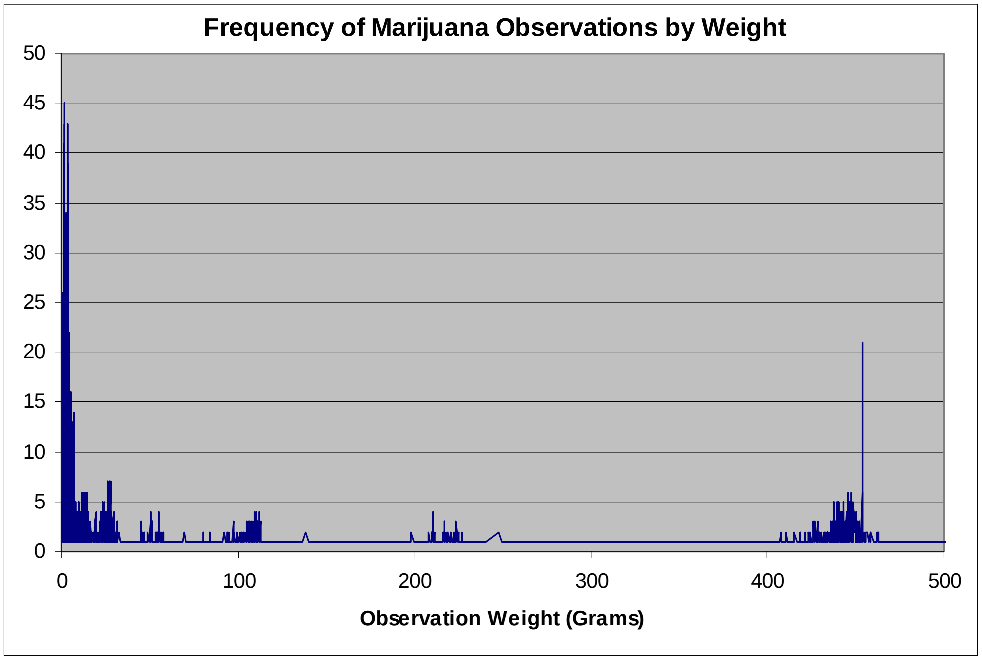


Figure A.9. Frequency of Marijuana Purchase Observations by Weight, Amounts less than 500 Grams



Appendix B

Evaluating the Functional Form of the Purity Models

The econometric models used in this report are highly parameterized, with many random slope coefficients (which in essence represent city-specific slope coefficients). Given the erratic way in which the data are collected and the fact that some models for specific quantity levels are based on a relatively small number of observations, we wanted to assess the goodness of fit of these highly parameterized models. To do so, we compared our most complex version of these models to two simpler forms. The simplest version of the purity model is one that allows only for a random intercept (or individual city-specific effect). The next version of the model additionally allows the relationship between purity and amount of the drug to vary across cities. These simpler models are both nested within our complex version of the purity model, which also allows for unique time effects on purity across cities. Given that the models are all nested within our primary model, we can compare the appropriateness of each, using the Akaike Information Criterion (AIC). A smaller AIC value is indicative of a better-fitting model. AIC statistics for each of the price models are presented in Table B.1.

Table B.1. Akaike Information Criterion Test Statistics Evaluating Goodness of Fit of Alternative Models for the Purity Equation

Quantity Level	Column A Random intercept (no random slopes on amount or time)	Column B Random intercepts and random slope on amount only	Column C Primary Model: Random intercept and random slopes on amount and time
Powder Cocaine			
1	-812.3	-811.7	**-1,198.6**
2	-1,642.3	-1,643.1	**-1,791.2**
3	-6,957.1	-6,956.2	**-7,641.4**
4	-5,323.0	-5,442.6[a]	**-5,729.2**
Crack Cocaine			
1	-11,247.4	-11,247.4	**-11,580.0**
2	-14,767.4	-14,786.5	**-15,599.7**
3	-12,151.1	-12,187.0	**-12,763.4**
Heroin			
1	-5,022.6	-5,172.5	**-6622.4**
2	-1,178.1	-1,340.4	**-2223.9**
3	113.5	102.2	**-303.9**
d-Methamphetamine			
1	986.6	983.7	**773.7**
2	638.6	611.7	**403.9**
3	-491.5	-490.1	**-553.9**

[a] The model did not fully converge but still produced estimates.

The results in Column C represent the AIC value for the primary model presented in the report. We indicate in bold the specification of the model with the smallest AIC value. In each model, the AIC test for the primary model (Column C) is smaller than that for the simpler models. This suggests that the additional parameterization improves the performance of the purity model.

Table B.2 shows the covariance parameter estimates from the primary purity models for the time and amount variables. While most of the purity models did not have enough variation across cities in the relationship between purity and amount to produce random effects across cities, a few did have significant variation in this relationship. Furthermore, the time effects had statistically significant variation across cities in all models. This offers further support for the improvement of using the primary model in this report rather than the simpler models. In the next appendix, we perform a similar analysis for the price models.

Table B.2. Covariance Estimates Generated for the Purity Equation from Our Primary Model, Including Random Slope Coefficients

	Covariance Parameter Estimate for Purity & Time				Covariance Parameter Estimate for Purity & Amount			
Model	**Estimate**	**Std Error**	**Z-Value**	**Prob Z**	**Estimate**	**Std Error**	**Z-Value**	**Prob Z**
Powder								
Quantity Level 1	0.0089	0.0009	10.47	<.0001	1.960E-04	2.980E-04	0.66	0.2551
Quantity Level 2	0.0047	0.0005	8.65	<.0001	1.200E-05	1.300E-05	0.91	0.182
Quantity Level 3	0.0048	0.0003	15.47	<.0001	2.164E-07	0	.	.
Quantity Level 4	0.0039	0.0003	11.04	<.0001	1.664E-09	0	.	.
Crack								
Quantity Level 1	0.0033	0.0004	9.01	<.0001	0	.	.	.
Quantity Level 2	0.0032	0.0003	12.49	<.0001	3.198E-06	1.769E-06	1.81	0.0353
Quantity Level 3	0.0040	0.0003	12.13	<.0001	2.793E-08	0	.	.
Heroin								
Quantity Level 1	0.0123	0.0007	17.07	<.0001	7.897E-03	2.668E-03	2.96	0.0015
Quantity Level 2	0.0174	0.0011	15.8	<.0001	1.690E-04	5.300E-05	3.18	0.0007
Quantity Level 3	0.0126	0.0011	11.89	<.0001	2.156E-08	0	.	.
D-Methamphetamine								
Quantity Level 1	0.0152	0.0019	8.12	<.0001	5.700E-05	5.700E-05	1	0.158
Quantity Level 2	0.0101	0.0012	8.29	<.0001	1.234E-06	0	.	.
Quantity Level 3	0.0033	0.0008	4.44	<.0001	0	.	.	.

Appendix C

Evaluating the Functional Form of the Price Models

We present here the results of our analysis of the price model, which is similar to the analysis of the performance of the purity model reported in Appendix B,. The primary price model employed in these reports includes city-specific slope coefficients on the expected pure amount of the drugs involved in the transactions as well as city-specific slope coefficients on the time variables. Again, we compared our most complex version of these models to two simpler forms, with the simplest version allowing for only a random intercept (or individual city-specific effect) and the next version allowing for random intercepts and the relationship between price and expected pure amount of the drug to vary across cities. We present the AIC statistics for each of the price models in Table C.1, indicating the model with the smallest AIC statistic in bold type.

Table C.1. Akaike Information Criterion Test Statistics Evaluating Goodness of Fit of Alternative Models for the Price Equation

Quantity Level	Column A Random intercept (no random slopes on amount or time)	Column B Random intercepts and random slope on amount only	Column C Primary Model: Random intercept and random slopes on amount and time
Powder Cocaine			
1	6,236.5	**6,158.9**	6,431.8
2	**2,902.2**	2,906.2	2,977.2
3	-3,494.3	-3,576.5	**-5,069.9**
4	-2,310.1	-2,440.2	**-2,650.6**
Crack Cocaine			
1	8,771.9	8,728.7	**8,398.4**
2	13,585.4	13,492.3	**13,089.5**
3	-1,620.8	-1,770.3	**-2,239.0**
Heroin			
1	12,212.9	11,093.3	**9,559.5**
2	21,253.8	15,307.4	**9,398.4**
3	**7,729.7**	8,412.1	7,906.8
D-Methamphetamine			
1	**2,932.6**	3,175.3	3,747.3
2	4,498.5	4,327.6	**4,298.2**
3	6,089.7	5,108.7	**2,474.6**
Marijuana			
1	**1,503.9**	1,818.2	
2	1,167.2	1167.2	**1,165.0**
3	1,500.8	**1,475.5**	1,770.2

The results in Column C represent the AIC value for the primary model presented in the report for all substances except marijuana, lowest quantity level. In that case, the model is estimated using only random intercept and slope coefficients because of the relatively small sample size reducing statistical power. In general, the AIC test for the primary model (Column C) is smaller than that for the simpler models, with a few notable exceptions. In the case of powder cocaine and heroin, where the sample sizes are relatively large, yet the simpler models generated lower AIC values, the difference between the AIC value in our primary model and the winning model is not very large, suggesting that the additional parameterization does not dramatically reduce the overall performance of the model. Further, as can be seen in Table C.2, in most of these cases, supplemental tests suggest that the additional parameterization in terms of random slope coefficients is warranted, based on the statistical significance of the covariance estimates for specific variables.

Table C.2. Covariance Estimates Generated for the Price Equation from Our Primary Model, Including Random Slope Coefficients

	Covariance Parameter Estimate for Price & Time				Covariance Parameter Estimate for Price & Ln(Expected Pure Gram)			
Model	**Estimate**	**Std Error**	**Z-Value**	**Prob Z**	**Estimate**	**Std Error**	**Z-Value**	**Prob Z**
Powder								
Quantity Level 1	0.0864	0.0057	15.15	<.0001	0.0033	0.0018	1.83	0.034
Quantity Level 2	0.0381	0.0022	17.35	<.0001	0.0012	0.0007	1.59	0.057
Quantity Level 3	0.0176	0.0007	23.97	<.0001	0.0012	0.0004	2.58	0.005
Quantity Level 4	0.0115	0.0007	16.94	<.0001	0.0005	0.0002	2.75	0.003
Crack								
Quantity Level 1	0.0849	0.0057	14.95	<.0001	0.0052	0.0024	2.21	0.014
Quantity Level 2	0.0317	0.0022	14.64	<.0001	0.0041	0.0012	3.3	0.001
Quantity Level 3	0.0082	0.0006	13.11	<.0001	0.0007	0.0002	2.85	0.002
Heroin								
Quantity Level 1	0.1159	0.0054	21.43	<.0001	0.0281	0.0074	3.8	<.0001
Quantity Level 2	0.1288	0.0073	17.76	<.0001	0.0228	0.0060	3.8	<.0001
Quantity Level 3	0.1480	0.0076	19.58	<.0001	0.0026	0.0010	2.73	0.003
D-Methamphetamine								
Quantity Level 1	0.0050	0.0029	1.74	0.041	0.1018	0.0086	11.79	<.0001
Quantity Level 2	0.0074	0.0029	2.56	0.005	0.0520	0.0044	11.81	<.0001
Quantity Level 3	0.0115	0.0040	2.91	0.002	0.0294	0.0051	5.78	<.0001
Marijuana								
Quantity Level 1					0.1162	0.0401	2.9	0.002
Quantity Level 2	0.2983	0.0368	8.12	<.0001				
Quantity Level 3	0.1184	0.0140	8.44	<.0001	0.0012	0.0006	1.9	0.029

Table C.1 and C.2 suggest that the primary model fits the majority of quantity levels within each substance except marijuana better than the other models. Because specific quantity levels were determined on criteria independent of model specification, we imposed the primary model on all quantity levels so that they could all be estimated in a consistent fashion (for comparability). The results in Tables C.1 and C.2 suggest that some efficiency may be lost in doing this, but certainly not much. In the case of marijuana, there is no clearly superior model that can be applied to all quantity levels. However, the loss in goodness of fit from adapting a slightly more parameterized specification is not great, and the predicted trends generated from differing models are not very different. Thus, again for consistency, we imposed the same basic functional form as that of the other models.

Appendix D

Annual City-Specific Price and Purity Predictions Based on City-Specific Price/Purity Models

Table D.1. Predicted Price Per Gram of Powder Cocaine, Lowest Quantity Level (Evaluated at 0.75 Grams) in Various Cities

	Atlanta			Chicago			New York			San Diego			Washington DC		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1981	380.31	632.48	1115.27	269.25	500.54	925.13	223.59	431.84	659.23	279.10	393.85	517.36	271.44	497.29	664.71
1982	356.07	620.81	1387.83	315.54	609.54	1419.59	314.53	475.52	713.65	256.41	368.41	421.36	290.53	564.85	826.18
1983	291.31	538.50	1001.97	260.42	434.54	647.57	207.48	332.04	473.68	217.25	305.06	399.27	296.54	424.32	584.30
1984	227.99	487.70	787.64	273.16	464.49	846.95	148.64	266.35	494.32	231.78	302.12	431.56	242.79	323.57	441.92
1985	234.71	429.51	793.37	231.35	368.05	531.23	128.58	235.19	359.97	256.53	314.54	392.88	206.25	272.36	294.69
1986	181.13	344.07	605.37	181.18	365.32	732.65	95.37	195.49	342.16	187.31	245.74	286.95	203.10	241.79	270.22
1987	147.37	308.22	630.68	92.71	212.07	474.84	79.30	195.55	273.73	135.71	186.37	242.88	171.05	208.64	246.20
1988	132.51	248.01	453.23	87.94	195.99	375.96	84.70	179.43	324.26	67.58	154.26	222.31	125.38	163.83	201.80
1989	115.59	226.70	464.32	108.26	177.94	272.12	81.20	124.71	183.25	94.60	148.97	230.59	103.77	156.64	240.84
1990	143.32	305.20	611.65	106.64	201.72	444.05	79.77	117.11	201.27	92.05	167.32	276.97	138.34	193.84	267.59
1991	119.71	240.50	490.01	112.75	160.20	303.27	62.43	99.17	177.18	78.16	130.41	221.22	120.93	183.43	290.57
1992	93.00	170.42	340.55	78.15	144.83	280.37	48.73	75.14	106.54	61.23	102.03	209.36	91.66	160.96	247.17
1993	83.87	173.52	391.32	80.21	132.75	210.72	60.68	94.94	221.62	52.55	101.52	228.44	92.85	173.93	289.32
1994	86.85	174.42	366.59	69.01	135.69	286.64	55.16	77.78	120.69	51.64	115.40	203.12	76.58	130.78	278.77
1995	100.41	210.98	428.11	79.58	166.08	334.77	58.82	115.74	242.54	58.81	90.95	127.96	82.73	209.69	418.43
1996	80.71	174.36	376.21	51.47	136.05	294.15	61.25	118.30	238.37	34.64	94.22	176.65	66.33	132.85	230.06
1997	91.20	171.36	314.31	72.29	142.95	245.60	49.84	105.29	169.26	51.35	112.12	236.83	66.86	136.18	238.94
1998	77.73	153.94	336.67	61.61	128.38	264.71	41.89	89.14	190.69	43.08	80.50	130.54	68.30	141.04	302.40
1999	75.84	182.52	320.24	51.77	115.65	237.71	43.77	94.95	172.21	45.10	84.42	154.75	70.46	112.05	155.61
2000	83.32	182.21	436.13	66.04	152.29	341.01	53.83	145.08	372.87	49.54	108.78	254.64	64.51	142.86	331.63
2001	94.19	199.38	400.19	76.02	141.11	312.91	62.71	117.10	226.68	63.25	128.11	241.00	72.90	144.27	304.31
2002	65.63	147.94	335.21	35.55	106.24	262.12	39.07	85.86	169.68	29.89	76.50	195.74	50.81	123.91	265.87
2003*	54.19	121.71	211.71	42.93	93.18	165.61	41.71	81.12	125.49	32.39	59.28	85.52	41.93	93.38	161.09

***** 2003 prices are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.
Note: These predicted prices are based on econometric models estimated for quarters, not years. We transformed the prices from quarterly to annual by taking the average for the “predicted” and taking the minimum “lower bound” and maximum “upper bound” for the 95 percent confidence intervals. All prices are adjusted for inflation and reported in 2002 dollars. Estimates to the penny are provided to facilitate replication/confirmation and are not intended to be meaningfully interpreted, given how broad the uncertainty bands are.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.2. Predicted Purity of Powder Cocaine, Lowest Quantity Level (Evaluated at 0.75 Grams) in Various Cities

	Atlanta			Chicago			New York			San Diego			Washington DC		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1981	0.06	0.30	0.56	0.15	0.41	0.72	0.15	0.36	0.62	0.37	0.53	0.69	0.09	0.28	0.57
1982	0.12	0.34	0.55	0.18	0.41	0.61	0.21	0.46	0.77	0.47	0.58	0.70	0.11	0.29	0.44
1983	0.17	0.38	0.65	0.28	0.45	0.65	0.31	0.53	0.77	0.47	0.62	0.76	0.26	0.40	0.52
1984	0.28	0.44	0.64	0.34	0.47	0.62	0.41	0.59	0.79	0.44	0.63	0.80	0.26	0.37	0.45
1985	0.21	0.42	0.61	0.32	0.44	0.60	0.33	0.56	0.80	0.48	0.60	0.71	0.27	0.34	0.42
1986	0.29	0.55	0.83	0.26	0.55	0.86	0.50	0.69	0.91	0.63	0.78	0.97	0.30	0.41	0.52
1987	0.43	0.68	0.89	0.49	0.71	0.91	0.53	0.74	0.97	0.66	0.82	0.94	0.41	0.51	0.64
1988	0.32	0.64	0.89	0.54	0.73	0.94	0.47	0.73	0.94	0.72	0.87	1.05	0.50	0.62	0.76
1989	0.32	0.60	0.80	0.52	0.72	0.92	0.45	0.69	0.92	0.65	0.82	0.96	0.44	0.61	0.78
1990	0.24	0.51	0.72	0.30	0.54	0.73	0.43	0.62	0.84	0.50	0.72	0.90	0.29	0.46	0.61
1991	0.31	0.55	0.79	0.43	0.59	0.85	0.40	0.60	0.89	0.54	0.76	0.98	0.42	0.63	0.74
1992	0.34	0.62	0.85	0.40	0.66	0.88	0.46	0.69	0.88	0.59	0.80	1.02	0.32	0.53	0.75
1993	0.40	0.62	0.83	0.48	0.67	0.88	0.50	0.67	0.91	0.59	0.81	1.01	0.37	0.62	0.84
1994	0.34	0.59	0.82	0.40	0.65	0.87	0.52	0.67	0.88	0.59	0.79	1.01	0.28	0.57	0.66
1995	0.23	0.53	0.74	0.29	0.59	0.79	0.33	0.64	0.86	0.44	0.73	0.85	0.19	0.48	0.68
1996	0.39	0.66	0.97	0.45	0.71	1.03	0.47	0.72	1.05	0.61	0.85	1.13	0.30	0.58	0.84
1997	0.31	0.59	0.84	0.37	0.65	0.90	0.44	0.69	0.90	0.47	0.76	1.00	0.36	0.56	0.77
1998	0.34	0.62	0.83	0.40	0.68	0.88	0.44	0.70	0.92	0.62	0.80	0.98	0.33	0.59	0.80
1999	0.32	0.59	0.86	0.34	0.61	0.88	0.41	0.64	0.92	0.52	0.77	0.97	0.43	0.61	0.77
2000	0.29	0.53	0.76	0.34	0.59	0.81	0.43	0.67	0.87	0.48	0.74	0.95	0.25	0.51	0.73
2001	0.23	0.50	0.79	0.28	0.57	0.84	0.32	0.58	0.88	0.46	0.68	0.92	0.19	0.47	0.75
2002	0.34	0.62	0.89	0.39	0.68	0.94	0.48	0.74	0.94	0.54	0.78	1.08	0.34	0.59	0.86
2003	0.37	0.62	0.85	0.42	0.66	0.90	0.52	0.75	0.94	0.60	0.82	1.01	0.33	0.59	0.80

*2003 expected purities are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.
Note: These predicted purities are based on econometric models estimated for quarters, not years. We transformed the prices from quarterly to annual by taking the average for the "predicted" and taking the minimum "lower bound" and maximum "upper bound" for the 95 percent confidence intervals.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.3. Predicted Price Per Gram of Crack Cocaine, Lowest Quantity Level (Evaluated at 0.3 Grams) in Various Cities

	Atlanta			Chicago			New York			San Diego			Washington DC		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1986	158.18	342.05	644.34	118.87	264.56	471.33	119.61	179.19	240.41	129.56	278.54	358.79	260.56	397.54	478.65
1987	107.35	323.68	605.45	80.66	237.71	443.02	119.01	300.11	585.29	127.31	214.45	235.03	193.21	270.47	349.55
1988	116.01	233.79	477.32	87.16	162.48	314.31	69.27	162.94	352.67	135.80	180.60	249.75	159.17	180.32	195.97
1989	86.59	207.94	346.67	73.11	148.72	299.22	76.43	138.13	247.06	124.66	174.08	229.60	152.25	172.90	194.10
1990	111.45	242.71	521.33	83.76	198.84	381.22	111.79	163.61	242.05	116.34	194.14	326.69	190.34	217.95	245.65
1991	76.99	192.51	379.50	61.72	152.70	209.88	70.50	119.46	274.49	93.16	169.28	315.26	166.55	174.91	181.38
1992	97.92	218.02	419.40	90.21	175.94	430.30	82.11	127.50	192.08	80.65	148.90	343.87	154.55	178.51	199.06
1993	87.07	186.93	353.92	85.98	131.92	183.36	73.66	119.70	174.40	67.40	142.86	265.14	155.23	175.77	192.16
1994	90.76	168.75	336.26	74.69	139.64	271.38	76.85	107.85	140.64	75.20	113.88	142.42	123.75	171.06	202.57
1995	89.94	180.92	383.22	61.81	126.60	309.53	79.52	154.34	231.96	84.01	125.23	168.31	163.25	189.81	231.87
1996	80.08	160.67	319.20	61.85	145.02	336.07	59.10	118.34	180.13	63.70	144.10	252.27	104.19	155.77	216.43
1997	87.61	200.46	438.14	65.87	142.80	320.37	104.95	175.21	237.76	67.84	145.20	328.17	134.57	185.27	229.21
1998	91.31	154.98	313.39	60.93	118.81	229.12	63.66	122.40	226.57	62.75	129.76	208.72	137.63	153.86	166.35
1999	72.99	181.63	368.81	76.34	132.49	254.14	81.32	152.93	276.32	77.71	167.13	276.21	135.23	153.41	160.95
2000	110.20	244.16	564.85	109.31	172.92	219.79	85.42	215.20	349.12	76.31	128.12	214.48	135.29	170.08	209.04
2001	98.69	252.39	370.80	74.21	124.93	246.94	109.19	196.06	279.19	76.43	143.27	252.95	143.20	168.34	202.07
2002	87.69	171.96	308.79	66.45	134.95	231.41	70.68	121.62	219.74	67.89	150.00	227.05	119.72	147.71	163.39
2003*	85.35	163.89	295.58	72.78	123.18	111.67	113.72	221.66	228.59	66.08	140.70	221.39	126.60	183.51	256.63

* 2003 prices are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.

Note: These predicted prices are based on econometric models estimated for quarters, not years. We transformed the prices from quarterly to annual by taking the average for the "predicted" and taking the minimum "lower bound" and maximum "upper bound" for the 95 percent confidence intervals. All prices are adjusted for inflation and reported in 2002 dollars. Estimates to the penny are provided to facilitate replication/confirmation and are not intended to be meaningfully interpreted, given how broad the uncertainty bands are.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.4. Predicted Purity of Crack Cocaine, Lowest Quantity Level (Evaluated at 0.3 Grams) in Various Cities

	Atlanta			Chicago			New York			San Diego			Washington DC		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1986	0.67	0.86	1.00	0.67	0.84	0.99	0.73	0.88	1.00	0.74	0.88	0.99	0.46	0.72	0.88
1987	0.62	0.85	1.00	0.60	0.82	1.00	0.69	0.87	1.00	0.47	0.74	1.00	0.62	0.79	0.90
1988	0.69	0.88	1.00	0.66	0.86	1.00	0.62	0.85	1.00	0.77	0.90	1.00	0.75	0.82	0.90
1989	0.75	0.88	1.00	0.72	0.86	0.98	0.78	0.88	0.99	0.76	0.89	0.99	0.79	0.84	0.86
1990	0.66	0.82	0.98	0.64	0.81	0.97	0.75	0.86	1.00	0.71	0.84	0.96	0.71	0.76	0.81
1991	0.73	0.87	1.00	0.73	0.86	0.98	0.76	0.89	1.00	0.72	0.88	1.00	0.80	0.83	0.86
1992	0.68	0.83	0.99	0.68	0.81	0.94	0.67	0.84	0.98	0.67	0.85	0.99	0.74	0.81	0.85
1993	0.67	0.82	0.97	0.62	0.78	0.90	0.71	0.86	0.98	0.68	0.84	0.97	0.70	0.75	0.78
1994	0.69	0.84	0.98	0.68	0.81	0.93	0.74	0.84	0.98	0.71	0.85	0.94	0.70	0.76	0.84
1995	0.59	0.77	0.91	0.59	0.74	0.87	0.65	0.80	0.93	0.58	0.73	0.83	0.60	0.68	0.75
1996	0.60	0.77	0.92	0.59	0.74	0.88	0.65	0.80	0.94	0.63	0.78	0.90	0.59	0.72	0.82
1997	0.52	0.72	0.85	0.50	0.71	0.86	0.63	0.79	0.89	0.53	0.75	0.88	0.54	0.64	0.70
1998	0.62	0.77	0.92	0.56	0.74	0.89	0.61	0.77	0.92	0.64	0.77	0.91	0.63	0.68	0.73
1999	0.57	0.72	0.87	0.55	0.70	0.84	0.62	0.76	0.88	0.60	0.74	0.87	0.50	0.64	0.71
2000	0.54	0.69	0.82	0.50	0.60	0.70	0.59	0.70	0.84	0.59	0.72	0.82	0.55	0.61	0.68
2001	0.53	0.68	0.82	0.51	0.66	0.79	0.59	0.72	0.86	0.54	0.68	0.81	0.56	0.60	0.64
2002	0.57	0.71	0.85	0.56	0.68	0.82	0.61	0.76	0.86	0.58	0.71	0.83	0.59	0.64	0.73
2003	0.62	0.76	0.89	0.58	0.73	0.87	0.66	0.76	0.87	0.62	0.75	0.89	0.60	0.67	0.70

* 2003 expected purities are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.

Note: These predicted purities are based on econometric models estimated for quarters, not years. We transformed the prices from quarterly to annual by taking the average for the “predicted” and taking the minimum “lower bound” and maximum “upper bound” for the 95 percent confidence intervals.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.5. Predicted Price Per Gram of Heroin, Lowest Quantity Level (Evaluated at 0.4 Grams) in Various Cities

	Atlanta			Chicago			New York			San Diego			Washington DC		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1981	640.01	1284.54	2782.55	638.75	2010.17	3027.40	295.53	651.24	979.50	879.59	1494.09	2180.85	981.82	1429.86	1826.68
1982	604.32	1180.29	2199.18	543.00	1396.51	3330.40	209.51	732.18	1874.98	783.22	1075.07	1376.76	759.26	1291.04	1796.90
1983	571.31	1092.65	2932.05	525.78	1314.00	2680.63	399.06	692.32	941.68	704.65	1186.54	1708.74	840.28	1870.81	3085.53
1984	432.68	908.92	1804.25	391.63	1042.69	2479.53	584.37	1093.55	1853.18	587.22	821.09	1166.99	461.29	1047.55	2059.78
1985	338.82	961.34	3271.90	418.88	961.66	1900.58	495.72	927.89	1578.15	637.18	1017.65	1444.89	446.38	892.13	1394.93
1986	248.01	763.42	2282.28	460.85	1038.53	2369.28	439.40	977.36	1517.80	501.33	849.06	1157.41	839.88	1226.20	1636.08
1987	373.22	865.66	1882.38	455.33	1193.79	2390.38	465.97	972.50	947.08	416.39	769.95	1428.79	577.82	930.12	1098.10
1988	407.46	609.20	831.65	380.50	743.92	1361.73	405.80	549.26	670.23	294.52	560.76	804.73	441.00	732.95	1155.30
1989	297.90	576.37	979.45	434.23	797.69	1192.00	450.63	573.45	725.15	380.47	613.67	1041.53	316.02	613.06	958.28
1990	274.78	689.21	1155.13	389.55	834.18	1396.08	472.10	596.61	692.20	258.92	601.34	1456.10	565.19	1048.66	1757.68
1991	391.82	556.00	889.43	314.83	566.55	1280.80	410.63	482.46	557.03	596.44	801.32	1070.86	569.58	877.17	1315.48
1992	237.78	489.50	961.10	235.28	462.93	613.43	335.75	444.47	553.45	423.38	609.66	803.90	540.25	762.84	1010.85
1993	322.53	481.24	611.38	230.25	469.23	931.95	300.76	378.36	450.80	260.97	424.99	547.98	548.71	690.58	869.68
1994	354.13	481.93	615.85	339.78	493.39	663.00	275.57	346.96	411.30	304.15	397.00	498.03	567.04	734.72	1015.53
1995	321.83	424.11	535.55	315.00	501.59	650.70	268.60	316.86	366.75	190.00	300.83	371.75	443.25	626.37	814.28
1996	329.19	427.05	568.45	197.60	403.34	601.65	272.64	369.98	463.83	229.23	316.31	519.84	446.66	662.00	859.00
1997	239.61	410.34	730.38	173.03	382.46	713.58	248.11	309.40	406.20	149.63	252.31	328.91	457.93	679.15	1021.90
1998	293.33	369.64	442.18	208.70	330.91	515.40	260.16	329.18	381.68	172.44	230.36	264.52	304.82	429.41	468.38
1999	282.58	452.24	1104.10	225.00	334.84	525.28	255.56	344.20	467.25	146.33	222.04	341.82	294.46	374.41	400.98
2000	293.43	370.22	459.80	195.53	327.61	471.38	225.99	292.72	415.70	156.88	239.34	349.52	317.08	425.92	589.38
2001	293.78	385.75	566.60	186.95	267.30	373.15	231.51	284.30	324.85	125.80	192.38	249.47	245.03	334.09	478.80
2002	287.44	373.87	478.58	155.80	293.61	440.25	228.70	280.39	325.80	158.98	209.02	253.90	230.02	329.79	428.73
2003	251.83	370.48	532.55	154.00	216.74	215.38	120.84	231.92	297.90	118.46	209.93	480.21	230.43	298.96	398.73

* 2003 Prices are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.
Note: These predicted prices are based on econometric models estimated based on quarters, not years. We transformed the prices from quarterly to annual by taking the average for the “predicted” and taking the minimum “lower bound” and maximum “upper bound” for the 95 percent confidence intervals. All prices are adjusted for inflation and reported in 2002 dollars. Estimates to the penny are provided to facilitate replication/confirmation and are not intended to be meaningfully interpreted, given how broad the uncertainty bands are.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.6. Predicted Purity of Heroin, Lowest Quantity Level (Evaluated at 0.4 Grams) in Various Cities

	Atlanta			Chicago			New York			San Diego			Washington DC		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1981	0.00	0.06	0.27	0.00	0.00	0.19	0.12	0.25	0.38	0.00	0.15	0.37	0.00	0.05	0.11
1982	0.00	0.15	0.40	0.00	0.04	0.30	0.16	0.29	0.43	0.11	0.23	0.36	0.01	0.08	0.25
1983	0.00	0.08	0.36	0.00	0.02	0.26	0.10	0.25	0.45	0.06	0.19	0.40	0.00	0.10	0.31
1984	0.00	0.11	0.43	0.00	0.07	0.36	0.10	0.33	0.57	0.07	0.26	0.45	0.00	0.19	0.40
1985	0.00	0.18	0.51	0.00	0.07	0.32	0.15	0.36	0.57	0.10	0.31	0.50	0.00	0.11	0.28
1986	0.00	0.20	0.55	0.00	0.11	0.42	0.22	0.40	0.58	0.03	0.36	0.61	0.00	0.11	0.19
1987	0.00	0.19	0.51	0.00	0.08	0.35	0.15	0.38	0.61	0.08	0.19	0.32	0.01	0.15	0.31
1988	0.02	0.16	0.28	0.00	0.11	0.28	0.35	0.45	0.56	0.15	0.33	0.54	0.05	0.17	0.29
1989	0.04	0.20	0.43	0.00	0.18	0.42	0.36	0.47	0.56	0.24	0.44	0.64	0.00	0.28	0.55
1990	0.00	0.13	0.24	0.00	0.08	0.22	0.30	0.43	0.50	0.04	0.32	0.53	0.02	0.15	0.22
1991	0.00	0.24	0.38	0.00	0.14	0.26	0.43	0.51	0.61	0.05	0.27	0.43	0.01	0.17	0.32
1992	0.09	0.32	0.51	0.00	0.19	0.40	0.41	0.61	0.81	0.12	0.35	0.56	0.04	0.18	0.33
1993	0.09	0.30	0.43	0.09	0.31	0.46	0.58	0.69	0.80	0.27	0.43	0.60	0.11	0.23	0.35
1994	0.25	0.44	0.61	0.04	0.22	0.38	0.53	0.63	0.71	0.34	0.51	0.67	0.10	0.21	0.33
1995	0.35	0.52	0.60	0.12	0.26	0.42	0.60	0.72	0.85	0.25	0.48	0.67	0.14	0.28	0.41
1996	0.21	0.40	0.55	0.00	0.29	0.45	0.41	0.56	0.68	0.22	0.47	0.58	0.13	0.23	0.33
1997	0.19	0.42	0.69	0.13	0.30	0.50	0.45	0.64	0.77	0.15	0.48	0.68	0.12	0.23	0.34
1998	0.37	0.53	0.65	0.12	0.32	0.41	0.50	0.63	0.75	0.38	0.56	0.66	0.14	0.28	0.38
1999	0.42	0.55	0.65	0.13	0.29	0.46	0.47	0.62	0.77	0.37	0.55	0.71	0.16	0.23	0.29
2000	0.25	0.44	0.61	0.09	0.24	0.38	0.52	0.63	0.76	0.36	0.52	0.69	0.12	0.23	0.33
2001	0.26	0.46	0.60	0.05	0.22	0.36	0.45	0.60	0.69	0.29	0.45	0.59	0.12	0.21	0.30
2002	0.23	0.43	0.64	0.04	0.20	0.41	0.49	0.61	0.70	0.34	0.47	0.58	0.13	0.23	0.34
2003	0.26	0.43	0.59	0.08	0.20	0.31	0.36	0.58	0.74	0.20	0.43	0.67	0.03	0.19	0.30

* 2003 expected purities are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.
Note: These predicted purities are based on econometric models estimated based on quarters, not years. We transformed the prices from quarterly to annual by taking the average for the “predicted” and taking the minimum “lower bound” and maximum “upper bound” for the 95 percent confidence intervals.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.7. Predicted Price Per Gram of Methamphetamine, Middle Quantity Level (Evaluated at 27.5 Grams) in Various Cities

	Los Angeles			Phoenix			San Diego			San Francisco		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1981	29.70	58.88	116.29	35.75	71.09	140.83	29.37	57.84	113.51	41.43	82.09	162.04
1982	40.56	77.47	138.35	48.81	93.53	167.63	40.11	86.08	164.35	69.36	119.96	194.92
1983	32.65	78.69	189.16	39.30	95.01	229.11	32.28	82.34	184.52	58.08	118.04	263.59
1984	43.08	96.41	147.14	51.83	123.99	209.24	71.08	125.33	170.46	60.09	131.31	205.08
1985	26.83	81.02	173.95	32.28	89.54	207.32	26.54	87.28	179.66	49.86	105.03	191.65
1986	34.85	91.25	183.55	41.93	122.82	275.87	30.81	92.92	177.74	48.58	127.20	255.85
1987	41.55	73.61	125.74	49.99	96.23	152.31	60.94	80.37	105.27	57.93	105.12	163.31
1988	33.00	61.49	118.78	41.65	72.81	143.93	46.63	59.99	72.87	46.00	88.99	151.86
1989	33.08	63.23	120.01	39.80	81.20	138.17	43.93	60.00	85.59	49.10	97.03	178.04
1990	49.62	108.45	241.31	66.41	143.82	292.46	62.80	108.28	189.36	66.39	153.02	336.44
1991	44.42	107.03	236.91	67.73	133.17	243.14	58.76	88.01	136.84	61.92	152.27	330.30
1992	32.87	63.74	100.95	39.54	78.34	145.41	42.27	56.92	73.03	45.82	95.41	138.16
1993	25.62	52.14	84.64	30.73	60.43	121.27	33.08	49.86	76.71	35.72	70.41	139.50
1994	21.40	35.35	62.60	24.96	43.74	72.20	26.97	35.58	42.53	28.14	48.52	76.34
1995	25.82	60.03	174.89	25.97	68.66	200.40	23.24	53.25	145.87	31.24	77.82	243.84
1996	25.64	58.37	96.53	38.90	66.61	106.27	34.71	55.21	75.39	34.96	76.86	122.24
1997	23.83	41.71	75.26	38.32	53.09	73.15	29.17	36.31	43.20	33.69	52.14	81.37
1998	36.25	68.83	121.06	45.16	89.79	146.75	36.01	69.47	99.62	44.07	91.87	163.09
1999	41.24	66.82	107.76	37.72	72.15	97.81	46.27	57.86	71.07	49.81	86.97	126.45
2000	32.04	53.70	80.68	33.38	59.21	105.29	38.44	48.69	61.85	44.66	78.63	129.92
2001	24.15	45.18	73.16	33.74	55.74	96.83	36.67	44.82	50.46	38.53	60.23	103.21
2002	25.56	41.29	61.25	27.69	45.91	81.44	34.21	42.75	50.64	38.26	56.05	91.45
2003*	27.06	34.37	44.07	26.74	42.28	70.40	31.73	38.47	48.18	43.36	57.21	85.30

* 2003 Prices are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.
Note: These predicted prices are based on econometric models estimated based on quarters, not years. We transformed the prices from quarterly to annual by taking the average for the “predicted” and taking the minimum “lower bound” and maximum “upper bound” for the 95 percent confidence intervals. All prices are adjusted for inflation and reported in 2002 dollars. Estimates to the penny are provided to facilitate replication/confirmation and are not intended to be meaningfully interpreted, given how broad the uncertainty bands are.
Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Table D.8. Predicted Purity of Methamphetamine, Middle Quantity Level (Evaluated at 27.5 Grams) in Various Cities

	Los Angeles			Phoenix			San Diego			San Francisco		
Year	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound	Lower Bound	Average	Upper Bound
1981	0.38	0.72	1.00	0.18	0.52	0.86	0.28	0.62	0.96	0.26	0.59	0.94
1982	0.38	0.70	1.00	0.18	0.50	0.80	0.28	0.62	0.89	0.31	0.59	0.87
1983	0.28	0.68	1.00	0.08	0.47	0.86	0.18	0.57	0.96	0.15	0.56	0.91
1984	0.23	0.60	0.94	0.03	0.42	0.74	0.14	0.45	0.73	0.11	0.49	0.81
1985	0.31	0.68	1.00	0.15	0.50	0.95	0.24	0.58	1.00	0.28	0.57	0.98
1986	0.31	0.62	0.91	0.10	0.40	0.65	0.26	0.50	0.67	0.18	0.49	0.78
1987	0.39	0.67	0.92	0.19	0.45	0.72	0.37	0.59	0.77	0.25	0.50	0.76
1988	0.44	0.71	0.98	0.24	0.52	0.82	0.45	0.63	0.80	0.36	0.60	0.85
1989	0.44	0.73	1.00	0.27	0.52	0.80	0.35	0.57	0.74	0.24	0.58	0.88
1990	0.21	0.50	0.78	0.01	0.28	0.58	0.08	0.31	0.53	0.09	0.37	0.65
1991	0.25	0.54	0.78	0.02	0.29	0.56	0.19	0.40	0.62	0.04	0.36	0.65
1992	0.35	0.67	0.96	0.15	0.45	0.75	0.41	0.68	0.94	0.13	0.45	0.82
1993	0.36	0.73	1.00	0.20	0.55	0.88	0.48	0.71	0.93	0.28	0.62	0.93
1994	0.65	0.91	1.00	0.44	0.70	0.96	0.67	0.85	1.00	0.51	0.77	1.00
1995	0.23	0.75	1.00	0.04	0.58	0.95	0.08	0.69	1.00	0.10	0.64	1.00
1996	0.38	0.64	0.92	0.17	0.45	0.73	0.25	0.55	0.84	0.23	0.50	0.77
1997	0.44	0.69	0.87	0.24	0.43	0.60	0.56	0.72	0.88	0.31	0.55	0.78
1998	0.14	0.43	0.70	0.02	0.24	0.49	0.14	0.32	0.37	0.06	0.31	0.51
1999	0.20	0.56	0.89	0.07	0.32	0.60	0.19	0.33	0.46	0.10	0.44	0.68
2000	0.29	0.55	0.82	0.08	0.32	0.59	0.26	0.38	0.53	0.18	0.42	0.69
2001	0.35	0.68	0.93	0.17	0.47	0.77	0.23	0.43	0.63	0.25	0.59	0.82
2002	0.49	0.70	0.91	0.27	0.48	0.70	0.50	0.61	0.76	0.29	0.56	0.76
2003	0.56	0.73	0.93	0.34	0.56	0.77	0.61	0.74	0.81	0.48	0.69	0.82

*2003 expected purities are based on information from only the first two quarters of the year and thus are likely to be updated in future reports.
Note: These predicted purities are based on econometric models estimated based on quarters, not years. We transformed the prices from quarterly to annual by taking the average for the “predicted” and taking the minimum “lower bound” and maximum “upper bound” for the 95 percent confidence intervals.

Source: System to Retrieve Information on Drug Evidence (STRIDE).
Prepared by the RAND Corporation, April 2004.

Appendix E

Weights Used in Construction of the National Price and Purity Indices

Table E.1. Weights Used in Aggregation of City and Area Prices into National Price Indices

	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Atlanta	0.028	0.029	0.029	0.030	0.030	0.031	0.031	0.032	0.032	0.033
Baltimore	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027	0.027
Boston	0.038	0.038	0.038	0.038	0.037	0.037	0.037	0.037	0.036	0.036
Buffalo	0.015	0.015	0.015	0.014	0.014	0.014	0.014	0.014	0.013	0.013
Chicago	0.089	0.088	0.087	0.087	0.086	0.085	0.085	0.084	0.083	0.083
Cleveland	0.028	0.027	0.027	0.027	0.026	0.026	0.026	0.025	0.025	0.025
Dallas	0.026	0.026	0.027	0.027	0.027	0.028	0.028	0.029	0.029	0.030
Denver	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
Detroit	0.053	0.053	0.052	0.051	0.051	0.050	0.049	0.049	0.048	0.048
Houston	0.034	0.035	0.035	0.035	0.035	0.036	0.036	0.036	0.037	0.037
Kansas City	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
Los Angeles	0.093	0.093	0.094	0.095	0.095	0.096	0.097	0.097	0.098	0.099
Miami	0.020	0.020	0.020	0.021	0.021	0.021	0.021	0.021	0.021	0.022
Milwaukee	0.017	0.017	0.017	0.017	0.017	0.016	0.016	0.016	0.016	0.016
Minneapolis-St. Paul	0.027	0.027	0.027	0.028	0.028	0.028	0.028	0.028	0.028	0.028
New Orleans	0.016	0.016	0.016	0.015	0.015	0.015	0.015	0.015	0.014	0.014
New York	0.101	0.101	0.100	0.099	0.099	0.098	0.097	0.097	0.096	0.095
Newark	0.024	0.024	0.023	0.023	0.023	0.022	0.022	0.022	0.022	0.021
Philadelphia	0.058	0.058	0.058	0.057	0.057	0.056	0.056	0.056	0.055	0.055
Phoenix	0.020	0.021	0.021	0.022	0.022	0.023	0.023	0.024	0.024	0.025
Pittsburgh	0.031	0.031	0.030	0.030	0.029	0.029	0.028	0.028	0.027	0.027
Portland	0.016	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017
San Antonio	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.015	0.015
San Diego	0.023	0.024	0.024	0.025	0.025	0.026	0.026	0.027	0.027	0.028
San Francisco	0.018	0.018	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.018
Seattle	0.021	0.021	0.021	0.021	0.021	0.022	0.022	0.022	0.022	0.023
St. Louis	0.030	0.029	0.029	0.029	0.029	0.029	0.028	0.028	0.028	0.028
Tampa	0.020	0.020	0.021	0.021	0.021	0.022	0.022	0.022	0.023	0.023
Washington DC	0.043	0.044	0.044	0.045	0.045	0.045	0.046	0.046	0.047	0.047
Census Divisions:										
EN Central	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
ES Central	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Mid Atlantic	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
Mountain	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
New England	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Pacific	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
South Atlantic	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
WN Central	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
WS Central	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004

Note: Annual weights are used for each quarter in a given year.

Table E.1 (Continued)

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Atlanta	0.034	0.034	0.035	0.036	0.037	0.037	0.038	0.039	0.040	0.041
Baltimore	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.025	0.025	0.025
Boston	0.036	0.036	0.035	0.035	0.035	0.035	0.034	0.034	0.034	0.034
Buffalo	0.013	0.013	0.013	0.013	0.012	0.012	0.012	0.012	0.012	0.012
Chicago	0.083	0.083	0.082	0.082	0.082	0.082	0.082	0.082	0.082	0.082
Cleveland	0.024	0.024	0.024	0.024	0.023	0.023	0.023	0.023	0.022	0.022
Dallas	0.030	0.031	0.031	0.032	0.032	0.033	0.033	0.034	0.034	0.035
Denver	0.018	0.019	0.019	0.019	0.019	0.020	0.020	0.020	0.021	0.021
Detroit	0.047	0.047	0.046	0.046	0.046	0.045	0.045	0.045	0.044	0.044
Houston	0.037	0.038	0.038	0.039	0.039	0.040	0.040	0.040	0.041	0.041
Kansas City	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
Los Angeles	0.098	0.098	0.097	0.097	0.097	0.096	0.096	0.095	0.095	0.094
Miami	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022
Milwaukee	0.016	0.016	0.016	0.016	0.015	0.015	0.015	0.015	0.015	0.015
Minneapolis-St. Paul	0.028	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.029
New Orleans	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.013	0.013	0.013
New York	0.095	0.095	0.094	0.094	0.094	0.093	0.093	0.093	0.092	0.092
Newark	0.021	0.021	0.021	0.021	0.021	0.021	0.020	0.020	0.020	0.020
Philadelphia	0.054	0.054	0.054	0.053	0.053	0.052	0.052	0.051	0.051	0.050
Phoenix	0.026	0.026	0.027	0.028	0.028	0.029	0.030	0.031	0.031	0.032
Pittsburgh	0.026	0.026	0.026	0.025	0.025	0.025	0.024	0.024	0.024	0.023
Portland	0.017	0.017	0.018	0.018	0.018	0.018	0.018	0.019	0.019	0.019
San Antonio	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.016	0.016	0.016
San Diego	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.028
San Francisco	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
Seattle	0.023	0.023	0.023	0.023	0.023	0.023	0.024	0.024	0.024	0.024
St. Louis	0.028	0.027	0.027	0.027	0.027	0.027	0.026	0.026	0.026	0.026
Tampa	0.023	0.023	0.023	0.023	0.023	0.023	0.024	0.024	0.024	0.024
Washington D.C.	0.047	0.047	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.049
Census Divisions:										
EN Central	0.006	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
ES Central	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Mid Atlantic	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
Mountain	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
New England	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Pacific	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
South Atlantic	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.007	0.007
WN Central	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
WS Central	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004

Note: Annual weights are used for each quarter in a given year.

Table E.1 (Continued)

	2001	2002	2003
Atlanta	0.041	0.042	0.043
Baltimore	0.025	0.025	0.025
Boston	0.033	0.033	0.033
Buffalo	0.011	0.011	0.011
Chicago	0.082	0.081	0.081
Cleveland	0.022	0.022	0.022
Dallas	0.035	0.036	0.036
Denver	0.021	0.021	0.022
Detroit	0.043	0.043	0.043
Houston	0.042	0.042	0.043
Kansas City	0.018	0.018	0.017
Los Angeles	0.094	0.093	0.092
Miami	0.022	0.022	0.022
Milwaukee	0.015	0.015	0.014
Minneapolis-St. Paul	0.029	0.029	0.030
New Orleans	0.013	0.013	0.013
New York	0.092	0.091	0.091
Newark	0.020	0.020	0.020
Philadelphia	0.050	0.049	0.049
Phoenix	0.033	0.034	0.035
Pittsburgh	0.023	0.023	0.022
Portland	0.019	0.019	0.020
San Antonio	0.016	0.016	0.016
San Diego	0.028	0.028	0.028
San Francisco	0.018	0.018	0.017
Seattle	0.024	0.024	0.024
St. Louis	0.026	0.025	0.025
Tampa	0.024	0.024	0.024
Washington D.C.	0.049	0.049	0.049
Census Divisions:			
EN Central	0.005	0.005	0.005
ES Central	0.003	0.003	0.003
Mid Atlantic	0.004	0.004	0.004
Mountain	0.002	0.002	0.002
New England	0.002	0.002	0.002
Pacific	0.004	0.004	0.005
South Atlantic	0.007	0.007	0.007
WN Central	0.002	0.002	0.002
WS Central	0.004	0.004	0.004

Note: Annual weights are used for each quarter in a given year.

Appendix F

Tables and Charts Related to the Assessment of Uncertainty in the Model

Table F.1. Comparison of Key Regression Coefficients Using Bayesian and REML Methods to Estimate the Hierarchical Model Parameters in the Purity Model

Drug	Quantity Level	Number of Observations	Posterior Mean Estimate of Coefficient on Amount	REML-Based Estimate of Coefficient on Amount
Heroin	1	12,865	-1.967 (1.779)	-2.1277 (1.7540)
Powder cocaine	1	6,056	-0.4793 (0.6345)	-0.7726 (0.6807)

Table F.2. Comparison of Key Regression Coefficients Using Bayesian and REML Methods to Estimate the Hierarchical Model Parameters in the Price Model

Drug	Quantity Level	Number of Observations	Posterior Mean Estimate of Coefficient on log(pure amount)	REML-Based Estimate of Coefficient on log(pure amount)
Powder cocaine	1	12,711	0.5281 (0.03461)	0.5268 (0.03013)
Heroin	1	6,122	0.7054 (0.02731)	0.7091 (0.01870)

Figure F.1
50 and 95 Percent Posterior Probability Intervals of the National Price Index for Powder Cocaine, Quantity Level 1, Employing Bayesian Model

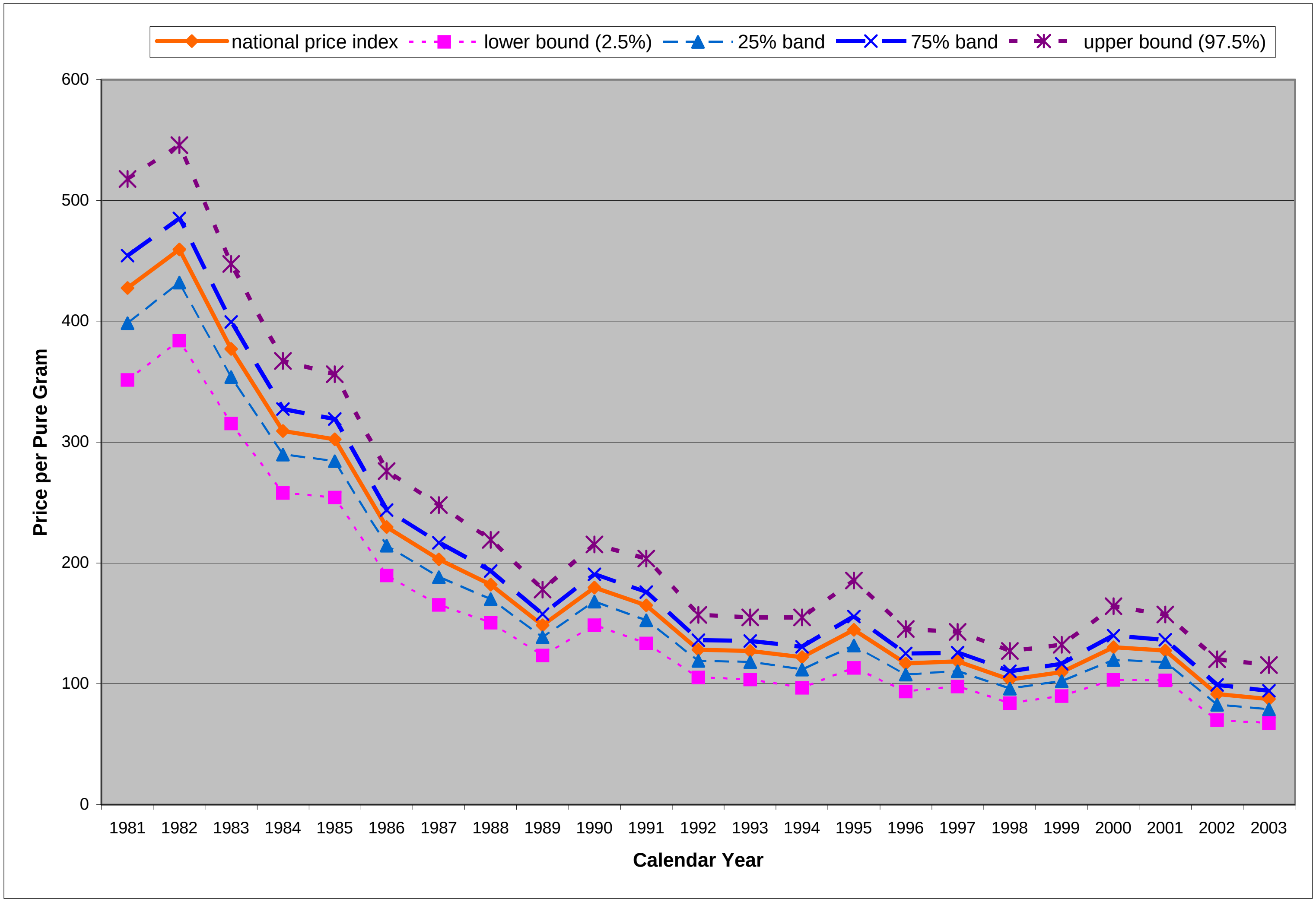
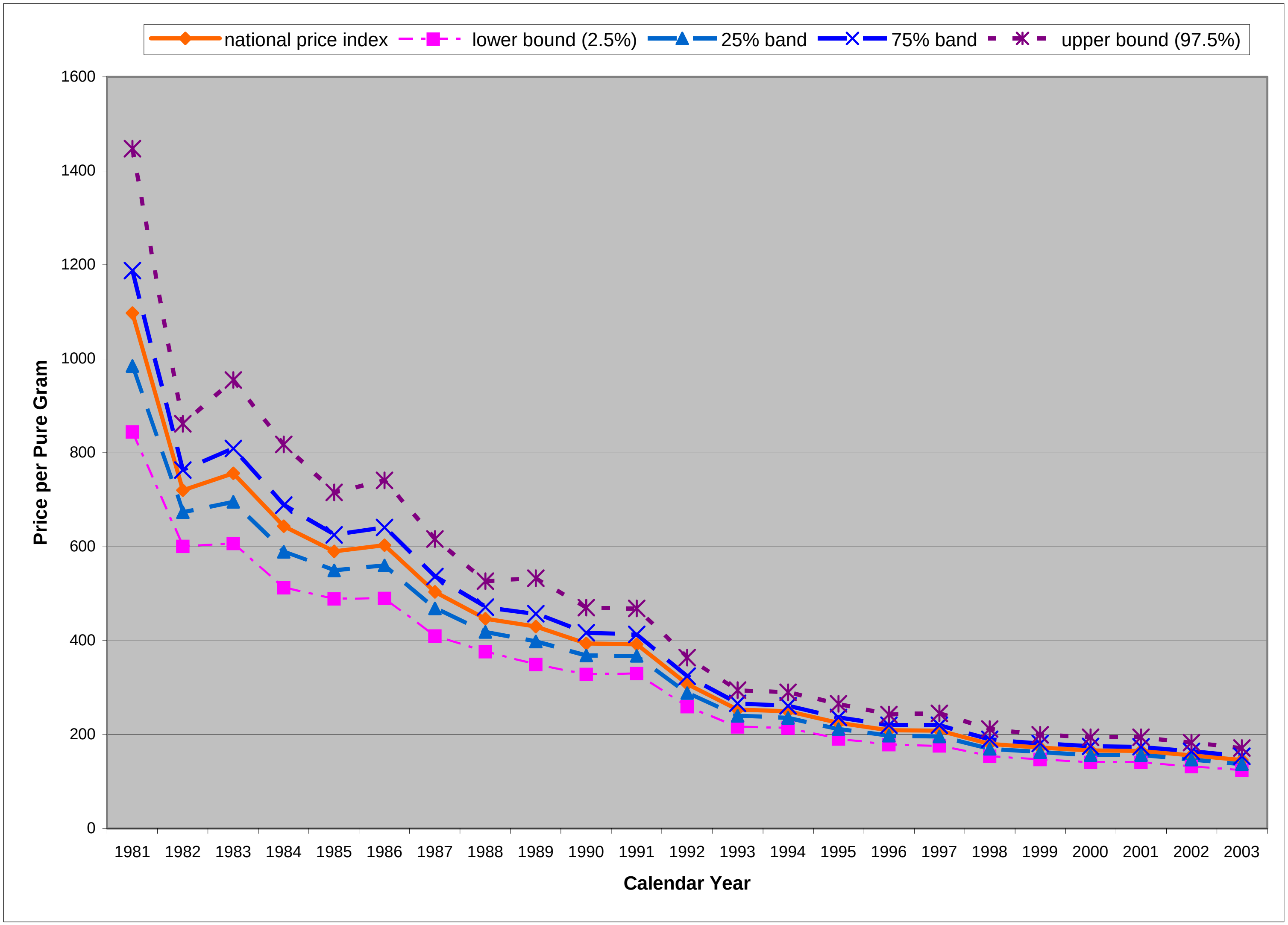


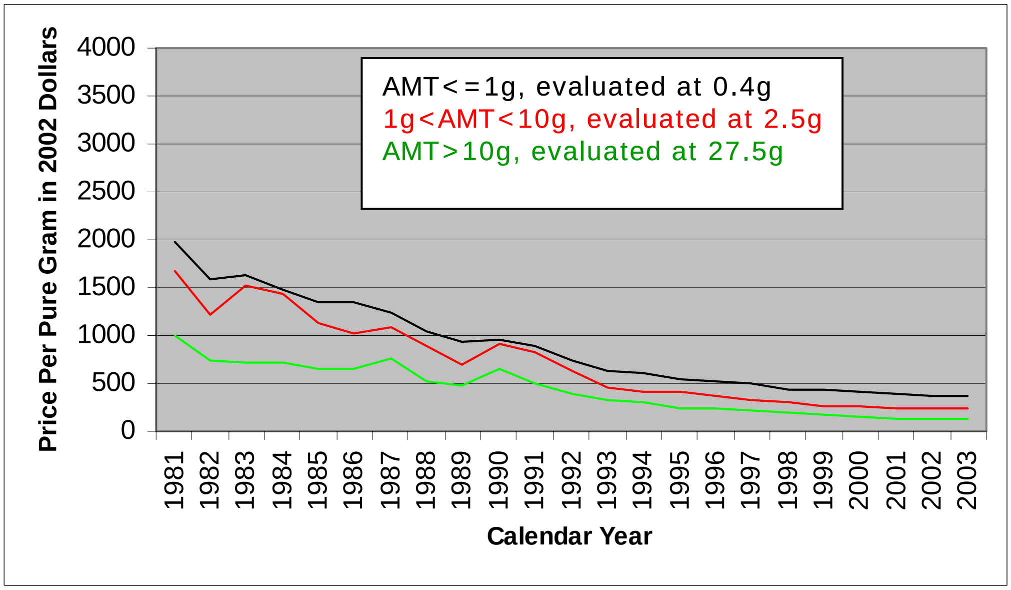
Figure F.2
50 and 95 Percent Posterior Probability Intervals of the National Price Index for Heroin, Quantity Level 1, Employing Bayesian Model



Appendix G

Comparisons of Main Heroin Models and Heroin, Salt-Undetermined Model

Figure G.1. Heroin Price per Pure Gram in 2002 Dollars.



Note: This figure comes from the main report. We changed the scale to permit direct comparison with the salt-undetermined series.

Figure G.2. Heroin, -Salt-Undetermined Price per Pure Gram in 2002 Dollars.

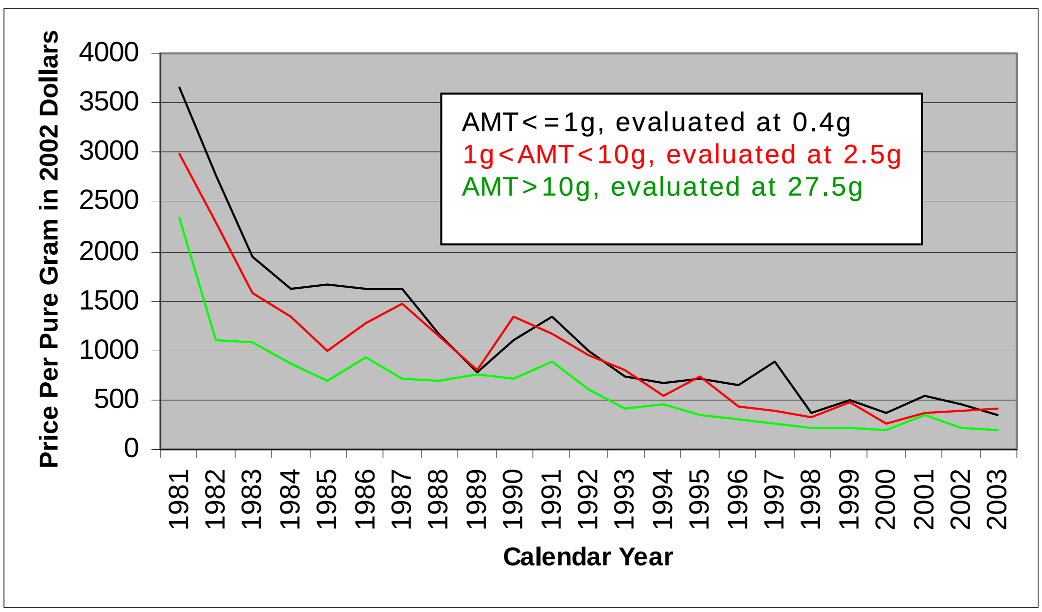


Figure G.3. Heroin Purity

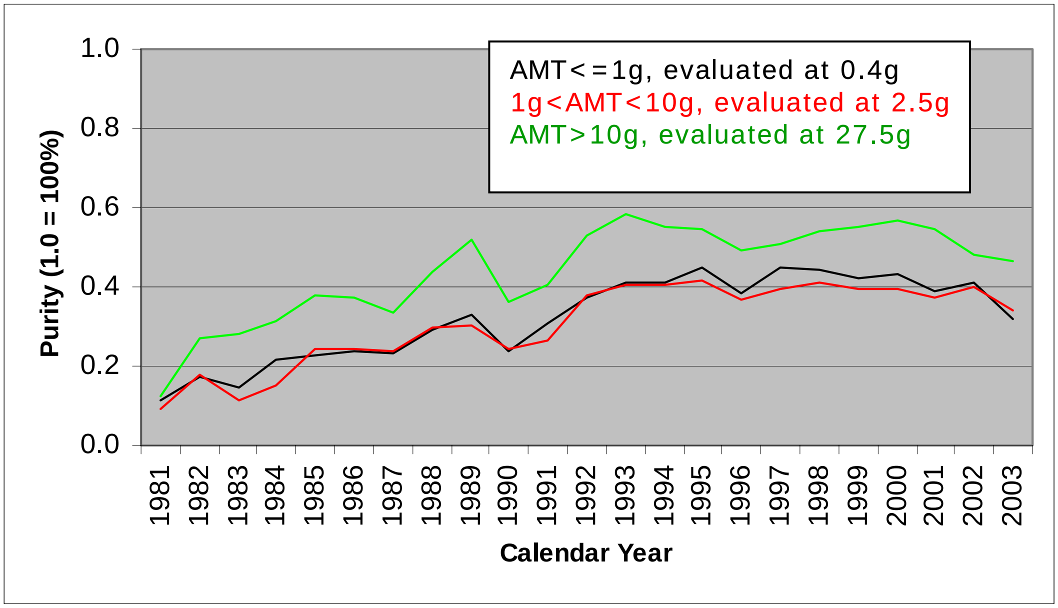
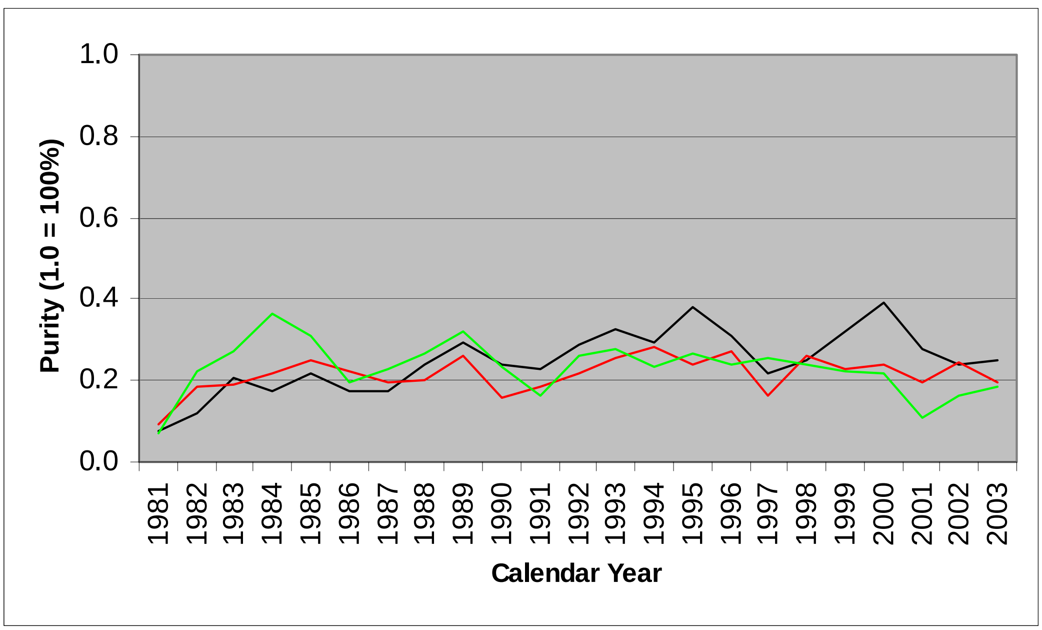


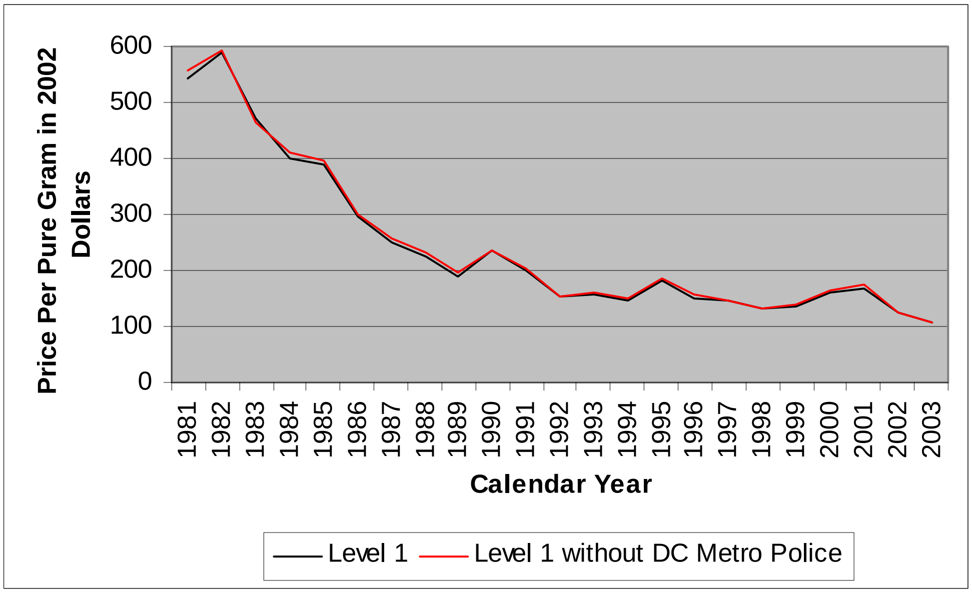
Figure G.4. Heroin, Salt-Undetermined Purity



Appendix H

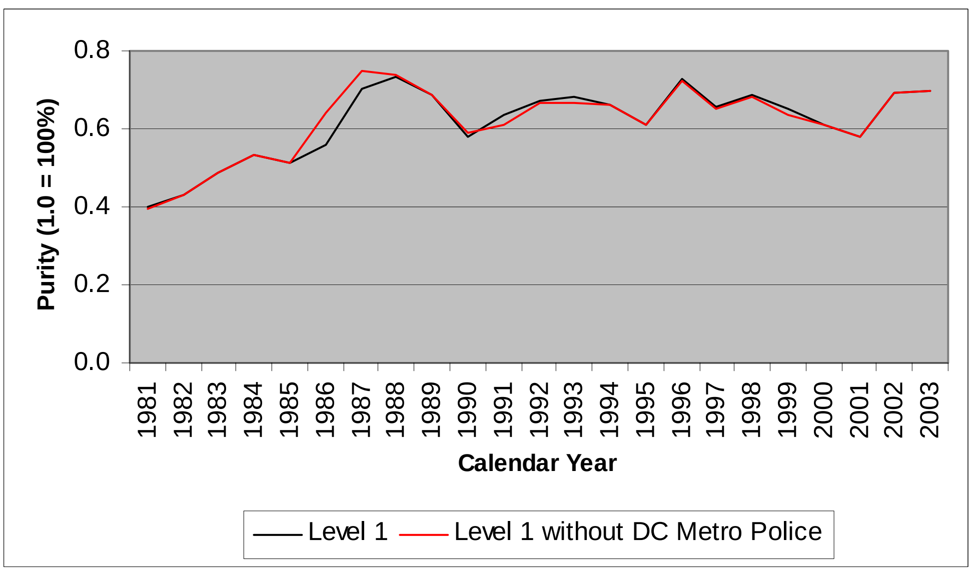
Impact of Excluding Washington D.C. Metropolitan Police Observations

Figure H.1. Powder Cocaine Price per Pure Gram, First Quantity Level (Amount ≤ 2 Grams), With and Without DC Metropolitan Police Transactions



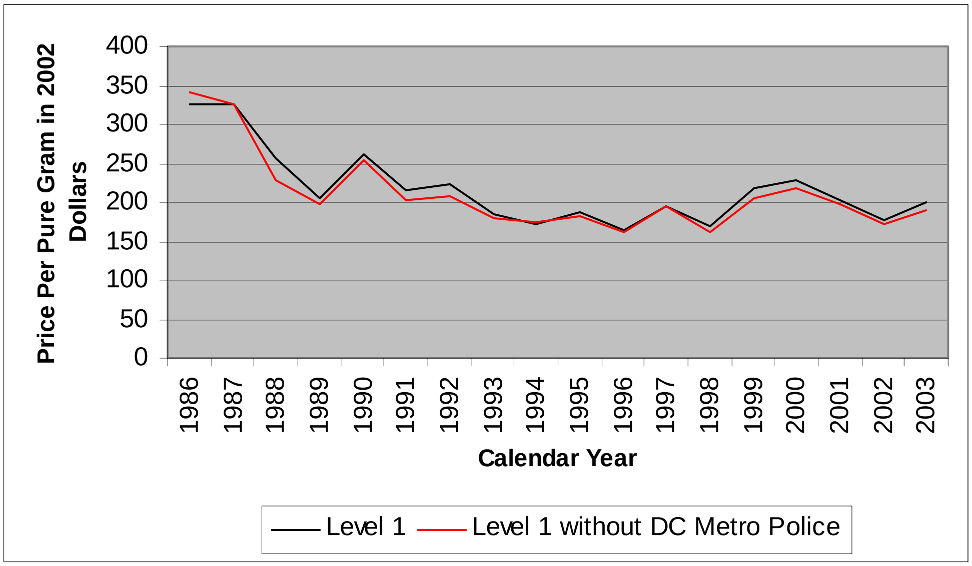
Note: Evaluated at 0.75 grams.

Figure H.2. Powder Cocaine Purity, First Quantity Level (Amount ≤ 2 Grams), With and Without DC Metropolitan Police Transactions



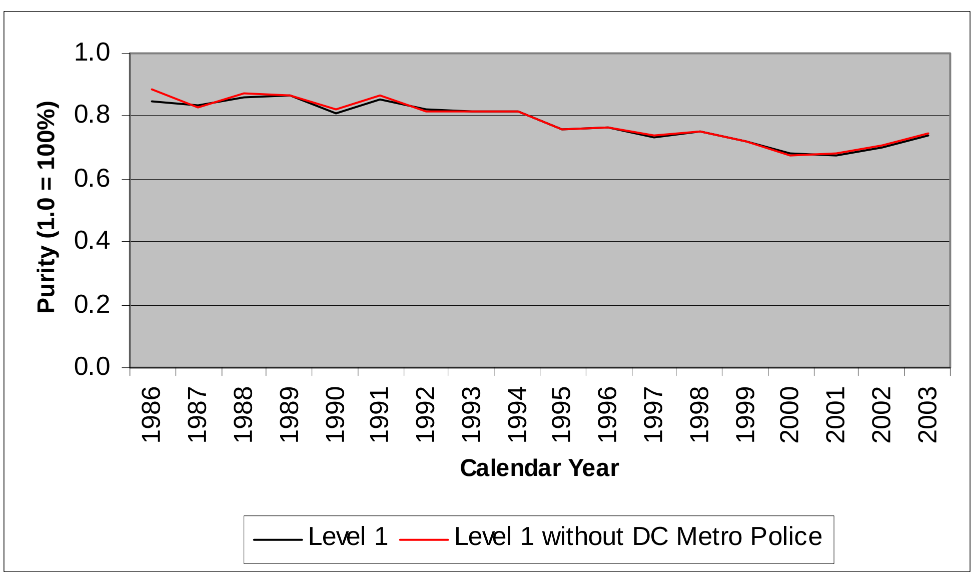
Note: Evaluated at 0.75 grams.

Figure H.3. Crack Cocaine Price per Pure Gram, First Quantity Level (Amount ≤ 1 Gram), With and Without DC Metropolitan Police Transactions



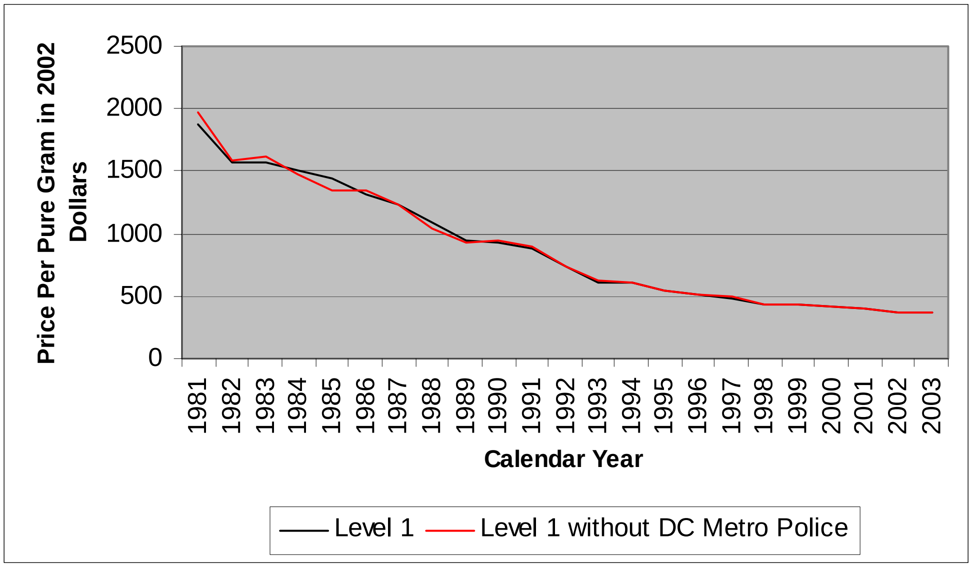
Note: Evaluated at 0.3 grams.

Figure H.4. Crack Cocaine Purity, First Quantity Level (Amount ≤ 1 Gram), With and Without DC Metropolitan Police Transactions



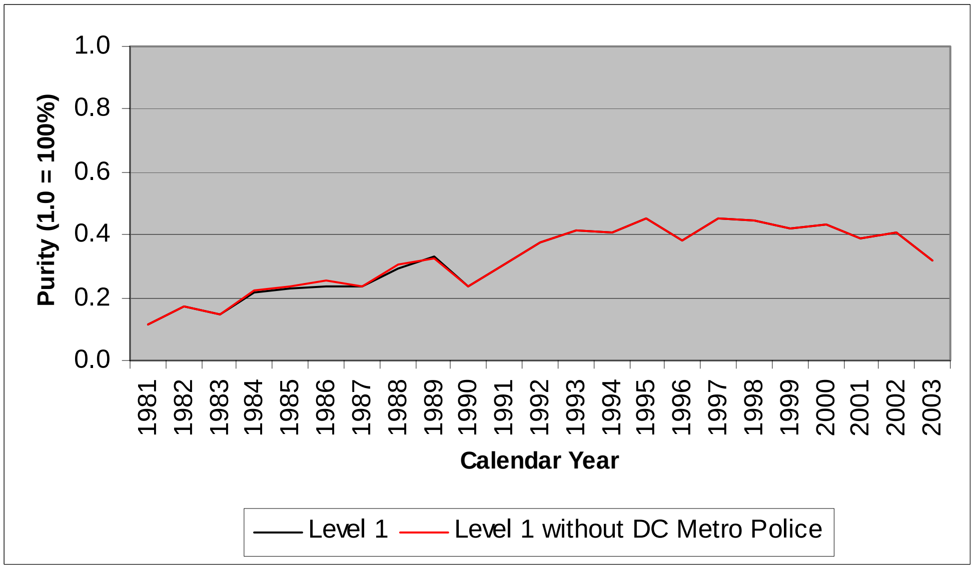
Note: Evaluated at 0.3 grams.

Figure H.5. Heroin Price per Pure Gram, First Quantity Level (Amount ≤ 1 Gram), With and Without DC Metropolitan Police Transactions



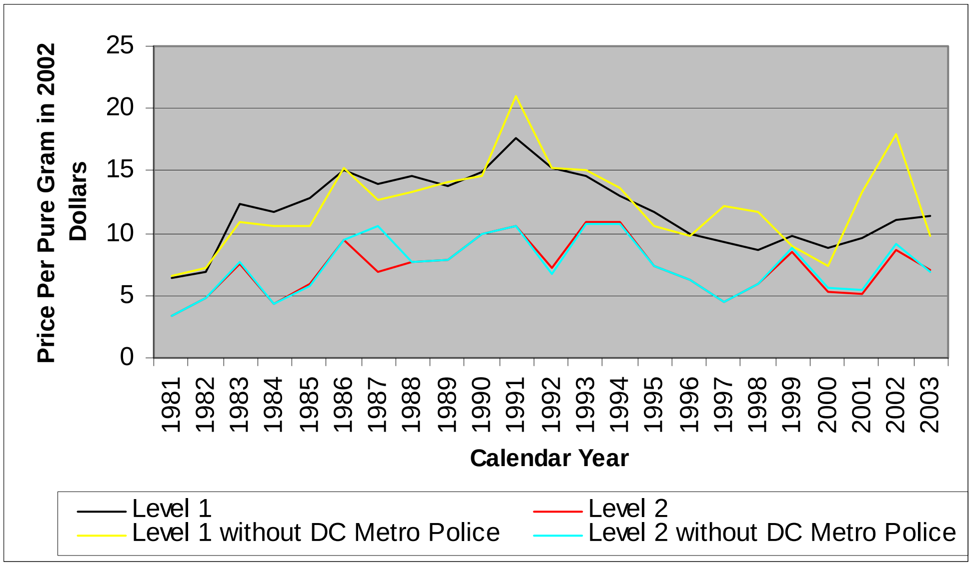
Note: Evaluated at 0.4 grams.

Figure H.6. Heroin Purity, First Quantity Level (Amount ≤ 1 Gram), With and Without DC Metropolitan Police Transactions



Note: Evaluated at 0.4 grams.

Figure H.7. Marijuana Price per Gram, First (Amount ≤ 10 Grams) and Second Quantity Levels (10 ≤ Amount ≤ 100 Grams), With and Without DC Metropolitan Police Transactions



Note: Evaluated at 2.5 grams for quantity level 1 and 26 grams for quantity level 2.

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF NATIONAL DRUG CONTROL POLICY
WASHINGTON, DC 20503

NCJ 207769