IDA

INSTITUTE FOR DEFENSE ANALYSES

Technical Report for the Price and Purity of Illicit Drugs: 1981– 2007

Arthur Fries, Project Leader
Robert W. Anthony
Andrew Cseko, Jr.
Carl C. Gaither
Eric Schulman

October 2008

IDA Paper P-4370
Log: H 08-001216

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About This Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract DASW01-04-C-0003, Task ES-55-2448, "Drug Use Indicators," for the Office of National Drug Control Policy. The views, opinions, and findings should not be construed as representing the official position of the sponsoring organization.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-4370

Technical Report for the Price and Purity of Illicit Drugs: 1981– 2007

Arthur Fries, Project Leader
Robert W. Anthony
Andrew Cseko, Jr.
Carl C. Gaither
Eric Schulman

PREFACE

This document was prepared by the Institute for Defense Analyses (IDA) for the Office of National Drug Control Policy (ONDCP), in partial fulfillment of the task Drug Use Indicators. This *Technical Report* is a companion to the associated *Results Report* entitled *The Price and Purity of Illicit Drugs: 1981 – 2007*. The present pair of reports update previous estimates published in the 2004 ONDCP report entitled *The Price and Purity of Illicit Drugs: 1981 Through the Second Quarter of 2003*.

All price and purity estimates were derived from records in the System To Retrieve Information from Drug Evidence (STRIDE) database maintained by the Drug Enforcement Administration (DEA) and provided to ONDCP and IDA. As noted in the following DEA disclaimer, these records should be considered to be "unvalidated DEA data":

> *Official Disclaimer: DEA responses to external data requests include all releasable records requested, without regard to analytic value. DEA analyses, by contrast, may exclude selected records, as closer inspection of such records may reveal errors, inaccuracies, or otherwise unverifiable data. External analyses of DEA data, accordingly, may not always yield conclusions consistent with DEA's own findings. Your acceptance and/or use of the information accompanying this disclaimer indicates your agreement (1) to refer to same information as "unvalidated DEA data," (2) to apply the guidance provided with same information competently, (3) to claim authorship/responsibility for any inferences/conclusions you may draw from same information, and (4) not to transmit same information to any other party without including this Official Disclaimer in your transmission.*

The IDA Technical Review Committee was chaired by Rear Admiral Richard B. Porterfield, USN (Ret.), and consisted of Mr. Saul A. Grandinetti, Mr. William B. Simpkins, and Dr. Richard H. White.

The authors are indebted to Dr. Michael A. Cala, our ONDCP sponsor, and Dr. Rosalie Liccardo Pacula and Dr. Jeremy Arkes from the RAND Drug Policy Research Center – for furnishing the software modules that constituted the EPH modeling, and for numerous related amplifying discussions. Quest Diagnostics, Inc., and Dr. Barry Sample,

Director of Science and Technology at the Employer Solutions business unit, are acknowledged for providing workforce drug testing data. Finally, the authors thank the DEA for review comments on an earlier draft of this document.

The viewpoints, results, and conclusions expressed in this document are solely those of the authors. No official endorsement by or attribution to ONDCP, the Rand Drug Policy Research Center, Quest Diagnostics, Inc., or the DEA is intended or should be inferred.

TECHNICAL REPORT FOR

THE PRICE AND PURITY OF ILLICIT DRUGS: 1981 – 2007

List of Figures

List of Tables

CHAPTER I

INTRODUCTION

I. INTRODUCTION

This *Technical Report* is a companion document to the associated IDA Paper P-4332, *The Price and Purity of Illicit Drugs: 1981 – 2007,* hereafter referred to as the *Results Report.* The primary objectives of these reports are to update previous estimates published by ONDCP in 2004, using the same Expected Purity Hypothesis (EPH) modeling methodology that produced the estimates given in the 2004 report, and to discuss the new results.

All reported price and purity estimates were derived from records in the System To Retrieve Information from Drug Evidence (STRIDE) database maintained by the Drug Enforcement Administration (DEA) and furnished to ONDCP. As noted in the following DEA disclaimer, the STRIDE records that we based our analyses on should be considered to be "unvalidated DEA data:"

> *Official Disclaimer: DEA responses to external data requests include all releasable records requested, without regard to analytic value. DEA analyses, by contrast, may exclude selected records, as closer inspection of such records may reveal errors, inaccuracies, or otherwise unverifiable data. External analyses of DEA data, accordingly, may not always yield conclusions consistent with DEA's own findings. Your acceptance and/or use of the information accompanying this disclaimer indicates your agreement (1) to refer to same information as "unvalidated DEA data," (2) to apply the guidance provided with same information competently, (3) to claim authorship/responsibility for any inferences/conclusions you may draw from same information, and (4) not to transmit same information to any other party without including this Official Disclaimer in your transmission.*

Our *Results Report* presents an overview portrayal of that EPH construct and related data processing steps. Neither our *Results Report* nor this *Technical Report* attempts to give an exhaustive description of these methods, e.g., explaining the structure

and implementation procedure of sophisticated regression models. Interested readers may turn to the 2004 ONDCP reports for those details.[1]

This *Technical Report* provides material that complements and supplements our *Results Report*. Some results are preliminary and might serve to motivate follow-on research.

Five chapters and an appendix of acronyms follow this introductory chapter. The subsequent chapters address in turn the following topics:

- IDA replication of the implementation steps undertaken in support of the 2004 ONDCP report and subsequent minor software modifications
- Expanded discussions of the value, limitations, and methods to compensate for limitations associated with STRIDE-centric studies
- Scatter plots of pure quantity-price data
- A median-based methodology for conveniently portraying the STRIDE data and facilitating direct simple comparisons to the EPH model results – to provide a context for assessing the EPH formulation and to motivate potential future enhancements
- Correlation analyses comparing STRIDE estimates to independent databases – price compilations from law enforcement sources (including local police and DEA) reported by the National Drug Intelligence Center (NDIC), and general workforce drug testing results.

Chapters IV and VI summarize analyses undertaken with earlier versions of the STRIDE database (i.e., not extending through the end of 2007). Time limitations precluded updating these analyses.

---

1 Office of National Dug Control Policy (2004). *The Price and Purity of Illicit Drugs: 1981 Through the Second Quarter of 2003*, Washington, D.C.: Executive Office of the President (Publication Number NCJ 207768), electronically accessible through the following World Wide Web address http://www.whitehousedrugpolicy.gov/publications/price_purity/. The accompanying *Technical Report* is available at http://www.whitehousedrugpolicy.gov/publications/price_purity_tech_rpt/. Sponsored by ONDCP, both reports were produced at RAND's Drug Policy Research Center and Public Safety and Justice Division.

CHAPTER II

EPH MODELING DETAILS – REPLICATION OF 2004 RESULTS AND FOLLOW-ON SOFTWARE MODIFICATIONS

II. EPH MODELING DETAILS – REPLICATION OF 2004 RESULTS AND FOLLOW-ON SOFTWARE MODIFICATIONS

Section A documents our replication (essentially identically in nearly all cases) of the data count totals, estimates, figures and tables presented in the 2004 ONDCP *Results Report* and *Technical Report*.

Section B details the few minor modifications that IDA introduced to the 2004 EPH software modules that were provided by ONDCP and the RAND Drug Policy Research Center. These revisions were incorporated to update inputs to encompass the 2007 timeframe, to reconcile the descriptions of the code provided in the 2004 ONDCP *Technical Report* with the actual content of the code, and to incorporate new information about STRIDE from the DEA.

A. REPLICATION OF 2004 RESULTS

The copies of RAND's code initially supplied by ONDCP turned out to be incomplete. RAND researchers kindly provided the missing modules and offered insightful explanations. ONDCP furnished us the identical STRIDE data set that RAND analyzed, encompassing the 22-year period 1 January 1981 to 31 May 2003. We verified that the total numbers of observations per drug type matched exactly, and proceeded to attempt to match all of the data counts for each of the subsequent data processing steps as well as the final model estimates.

The data processing counts matched exactly except for 13 suspect observations – 12 heroin and one marijuana. Contrary to their written intent, the RAND code included 12 heroin salt undetermined observations from the DMP. Similarly, RAND included a marijuana observation with drugcode 8222.000 that should have been deleted because the only valid marijuana drug codes for this study were stated to be 7600.000 (no plant material detected), 7360.4 (all plant material), and 7371.000 (Tetrahydrocannabinol-Organic).

Our constructed time series generally matched exactly those reported by RAND, with the exception of some very modest discrepancies for some of the quarters for heroin

price (< 0.40%) and d-methamphetamine price (< 0.01%).[1] The former likely results from our deletion of 12 heroin observations incorporated into the RAND analyses. Additionally, as discussed in Section B.2 below, there were some minor differences in how our population weights were calculated, which may have contributed to the minor discrepancies in d-methamphetamine price for a few isolated quarters.

B. FOLLOW-ON SOFTWARE MODIFICATIONS

1. Updates

Three sets of inputs to the EPH modeling computations must be updated whenever additional data from new calendar quarters are integrated into the existing STRIDE database.

The first required update is that ranges for variables in the software that run through the total number of quarters must be adjusted to match the new count for the total number of quarters. This value is hard-wired in several locations within the software. In our case, we changed "90" quarters to "108" quarters.

Second, for each of the new years encompassed by data updates, population weights must be determined for the 29 major cities and nine Census divisions identified in Chapter I of the *Results Report*, i.e., the 38 analysis entities, central to the data sorting and statistical estimation procedures. With the lone exception discussed in Section B below, we adhered to the detailed data processing code that translates individual city locations given in the STRIDE database (i.e., site of the subject purchase or seizure) uniquely to one of the specific 29 major cities or nine divisions. For each of these 38 analysis entities, decennial population estimates were obtained for the years 1980, 1990, and 2000. A constant linear population growth trend was assumed to hold for all intermediate years, as well as for the post-2000 years. We continued the extrapolation beyond 2003, as executed in support of the 2004 ONDCP reports, to the new years represented in our data, i.e., to 2004, 2005, 2006, and 2007.

---

[1] For heroin, $0.99602 \leq$ (RAND Quarter Estimate) / (IDA Quarter Estimate) ≤ 1.00310; while for d-methamphetamine, $0.99994 \leq$ (RAND Quarter Estimate) / (IDA Quarter Estimate) ≤ 1.00008. For the other three drug categories, the ratios of estimates were all identically 1 (to 5 decimal places).

When sufficient new quarters are introduced to the STRIDE database, one can contemplate adjusting the estimated price values for inflation. The 2004 ONDCP *Results Report* and *Technical Report* presented results in constant 2002 dollars, and we express our price estimates in terms of constant 2007 dollars. Inflation adjustments are based on the annual Consumer Price Index (CPI) for All Urban Consumers (non-seasonally adjusted) figures provided by the Bureau of Labor Statistics.[2]

2. Reconciliations

In our execution of the EPH modeling, we deleted the 13 suspect data points described above in Section A – to conform exactly to the written documentation.

The 2004 *Technical Report* noted that CPI adjustments were based on seasonally adjusted figures, but inspection of the code revealed that the actual values implemented in the code were non-seasonally adjusted CPIs. In our implementation, we retained the non-seasonally adjusted CPIs so that we could replicate the previous study's results.

The 2004 documentation reported that Oakland drug transactions were incorporated into "San Francisco" city, but the code in fact places them into the "Pacific" division. This specific assignment allowed us to replicate the 1980 population weights used in the 2004 study, but failed when we attempted to reproduce the 1990 population weights associated with San Francisco and the Pacific region. To replicate exactly those 1990 figures, we were forced to exclude the Oakland population from both San Francisco city and the Pacific division.

To reconcile all of this, and to match the given documentation, we subsequently absorbed all of the Oakland records into the San Francisco city designation (and thereby also excluded them from the Pacific division designation). Further, we accordingly modified all of the population weights over the entire time span of 1981 through 2007. This change, consistent with the intent articulated in the 2004 *Technical Report*, gives relatively more weight to San Francisco and visibly alters (shifts in level, but no changes in trends) early year portions of some of the time series for d-methamphetamine and marijuana.

---

[2] CPI figures are available at http://www.bls.gov/cpi/home.htm. As noted below in Section B.2, we use non-seasonally adjusted values to coincide with what was implemented in the code that generated the results published in the 2004 ONDCP report.

3. New Information

Per DEA Office of Forensic Science, zero purity is entered in STRIDE in lieu of a purity value when a quantitation is not performed. Thus, "zero purity" records in STRIDE are not necessarily indicative of samples that necessarily had zero or even low purity. With the concurrence of ONDCP, we therefore removed observations with "zero purity" for all drugs but marijuana from the STRIDE database. In particular, these observations did not contribute to any of the price and purity indices depicted in our *Results Report*, but they were used in analyses discussed in Chapters IV and VI of this *Technical Report*.

CHAPTER III

OBJECTIVES AND LIMITATIONS OF STRIDE-BASED ANALYSES

III. OBJECTIVES AND LIMITATIONS OF STRIDE-BASED ANALYSES

This chapter first identifies STRIDE analysis topics of potential value to the counter-drug policy and law enforcement communities. It then describes many of the inherent and technical limitations of STRIDE data in supporting such analyses. Finally, within the context of each limitation, this chapter discusses the prospects of, and methods for, reducing the uncertainty associated with the limitation.

A. SOME BASIC TOPICS OF VALUE TO THOSE IN DRUG CONTROL ACTIVITIES

To explore for topics of interest, consider the following progression of topics from the simple to the complex:

- Absolute measures of price, purity, or transactional quantities
- Long-term time variation of price, purity, or transactional quantities
- Short-term time variation of price, purity, or transactional quantities
- Regional or local variation of price, purity, or transactional quantities
- Correlations in time or location among drug types or forms, or correlations between STRIDE indicators and other indicators of drug abuse.

1. Absolute Measures

Several topics require absolute measures:

- Consistent year-to-year reports to the public on the price and purity of illicit narcotics
- Price per pure amount as one of the inputs to estimating the total amount of money spent on each illicit drug
- Law enforcement awareness of illicit market conditions to protect and inform undercover officers
- Estimates of price accessibility of illicit drugs to drug abusers and youth
- Anticipation of the attractiveness of illicit drugs and their physical effects based upon purity and likely form of delivery.

Note that since the logging and chemical analysis of seized or purchased drugs takes months to complete and enter into STRIDE, the law enforcement use of current drug prices on the street must be obtained locally as the samples are collected because prices can change significantly over the course of a month.

2. Short-Term Time Variations

Short-term time variation can be detected even without the rigor required for long-term calibration of trends. Knowing short-term variation provides useful information on several important topics:

- Measures of effectiveness of recent counter-drug operations or activities
- Indicators of emerging threats of increasing drug abuse (e.g., with sudden price declines or the creation of more attractive types or forms of drugs)
- Indicators of the risk of hospital emergencies as purity increases.

The first two topics could benefit from immediate observation of price increases or decreases, as well as from delayed intelligence reporting of distinct changes. For the third topic to be useful, either purity increases would have to be inferred from price drops, or the pacing of laboratory processing of drug samples would have to be accelerated greatly.

3. Long-Term Time Variations

Long-term time variation might be accurately tracked even without an absolute measure of prices. Knowing the long-term variation provides useful information on several important topics of value:

- Overall increase or decrease in affordability and attractiveness of the drug to users
- An indicator of the cumulative effects of counter-drug activities (interdiction, treatment, education) as well as changing societal attitudes toward drug abuse
- Tracking the shifts in drug forms (e.g., powder to crack cocaine) and types (e.g., cocaine to amphetamines) within drug markets.

Clearly, these topics are of more strategic than tactical value.

4. Regional or Local Variations

Persistent regional patterns of drug type, form, or price or purity differences can be informative:

- Evidence for underlying drug distribution networks and trading relationships (regionally or across the central-city, suburban, satellite-city, rural rings making up urban-rural zones)
- Evidence for the existence of middlemen of additional transactional steps in distribution
- Evidence for the effects of local laws or attitudes in shaping long-standing drug abuse patterns.

Short-term local variations in patterns of drug type, form, or price differences can sometimes be informative:

- Evidence for the impact of abrupt local changes (whether transient or persistent) in drug access, law enforcement, or social attitudes.

Here, comparative analysis can be of value if the variations from locale to locale can be distinguished from biases or distortions arising from the limitations of STRIDE as a scientific data set.

5. Correlations in Time or Among Locations of Various Indicators

Relationships among various indicators can illuminate several topics of interest:

- Evidence for competition among drug types (or forms)
- Validation of the significance of STRIDE indicators (time series or spatial patterns) by co-variation with non-STRIDE indicators of drug abuse
- Hypothesis testing for causal relationships in drug markets and drug abuse behavior
- Calibration of warning indicators for monitoring changes or operational impacts.

A consistent pattern of relationships across a wide range of indicators provides evidence for the validity of each of the contributing indicators. It also becomes the basis for calibrating indicators and predicting changes in some based on changes in others. Thus, the value of STRIDE indicators is amplified by relating those indicators to the family of other known indicators.

B. LIMITATIONS AND METHODS OF ANALYSIS

Employing STRIDE data to address the above topics encounters several types of limitations, some inherent in drug "markets" themselves, some specific to STRIDE data acquisition, and others arising from STRIDE data management. Certain limitations create significant unavoidable uncertainty, but for others, there are methods of analysis or

interpretation that can largely compensate for the limitation. Organizing the following discussion around the types of limitations motivates the discussion of methods, interpretations, and relevant topics of value in context.

One overarching fact influences both the limitations and methods: the unit price of illicit drugs is deeply discounted for larger transactional quantities. Also, transactional quantities cluster tightly around agreed upon amounts, which form a spectrum of values corresponding to levels in the distribution hierarchy.[1]

1. Inherent Limitations

Several limitations arise from the characteristics of drug markets themselves, and are inherent problems for any analysis of those markets.

a. Highly Variable Prices

Problem: Even within a given local area or for a single drug dealer, price differences for the same product can persistently vary by factors of two or more. This is probably a consequence of the risks of shopping around, both for customers and dealers. Also, the urgent needs of heavy users might limit their willingness to wait for the best or even the expected deal. Such behavior slows the process by which the market might converge to a common price. Even as the market slowly converges, disruptions in drug supply and turbulence in the associated criminal activities continually disperse the expected prices of individual transactions. The net effect is a market that never fully "clears" by converging to a commonly accepted price for a given standard product.

Methodological Comment: Given such variability, it is not possible to refer with any precision to the "price" for an illicit drug. For law enforcement purposes, the range of prices and practices is important to know, but for policy summaries, a single *index* might adequately represent the *distribution* of prices rather precisely. Selection of an adequate index, however, must deal with additional limitations.

b. Highly Variable Purity

Problem: Drug traffickers dilute the products they sell to make up for shortages; they also frequently cheat customers by selling a product with very low, even zero purity.

---

1 For graphical depictions and related discussions, see Figures III-1 and IV-1 to IV-3 of this *Technical Report*, as well as Office of National Dug Control Policy (2004), *The Price and Purity of Illicit Drugs: 1981 Through the Second Quarter of 2003*, Washington, D.C.: Executive Office of the President.

The distribution of prices per pure gram can, therefore, include many sales with very high unit prices, all the way up to infinite price. Mathematically, this dispersion of prices makes it possible for the price variance to become a divergent expression, and without a finite variance, the "average" of the distribution becomes meaningless. Even if the variance expression slowly converges, the "average" of such a distribution can be quite volatile for sampling reasons and not be representative of the drug market behavior.

Method 1: Define market levels for transactions by gross transaction quantity, and compute average purity as a basis for then calculating average price per pure gram. One can justify such an approach based on the customer's *expected* purity for the transaction. The 2004 ONDCP report refined this approach by further adjusting for the average price according to the discount on larger transaction quantities.

Another version of this method could take advantage of the known structure of the drug markets – nearly all transactions cluster about either one of a set of standard quantities (e.g., kilogram, ounce) or for small quantities a set of standard prices. Figure III-1 is an illustrative example, with evidence of fixed price levels on the left and fixed quantity levels (e.g., at 28 gm = 1 ounce) on the right. For these sorts of representations, average purities and prices could be computed for each standard quantity (or price) level in the market, and the prices per pure gram could then be computed for each standard quantity. Chapter IV presents related figures and discussions.

Method 2: The median of the distribution of unit prices per pure quantity has a finite variance even if the distribution itself does not. Generally, the median characterizes the bulk of the transactions rather than the exceptions in the divergent tails of a price-per-pure-gram distribution. Unfortunately, the median also can be unstable or misleading if it falls in a low-density region between peaks in the underlying distribution – this must be checked. To some extent, the median price for all transaction quantities of a drug is an artificial value representing some arbitrary quantity dictated by the sampling pattern; on the other hand, the medians of standard transaction quantities are well defined but are also difficult to interpret when reassembled into an overall characterization of the drug.

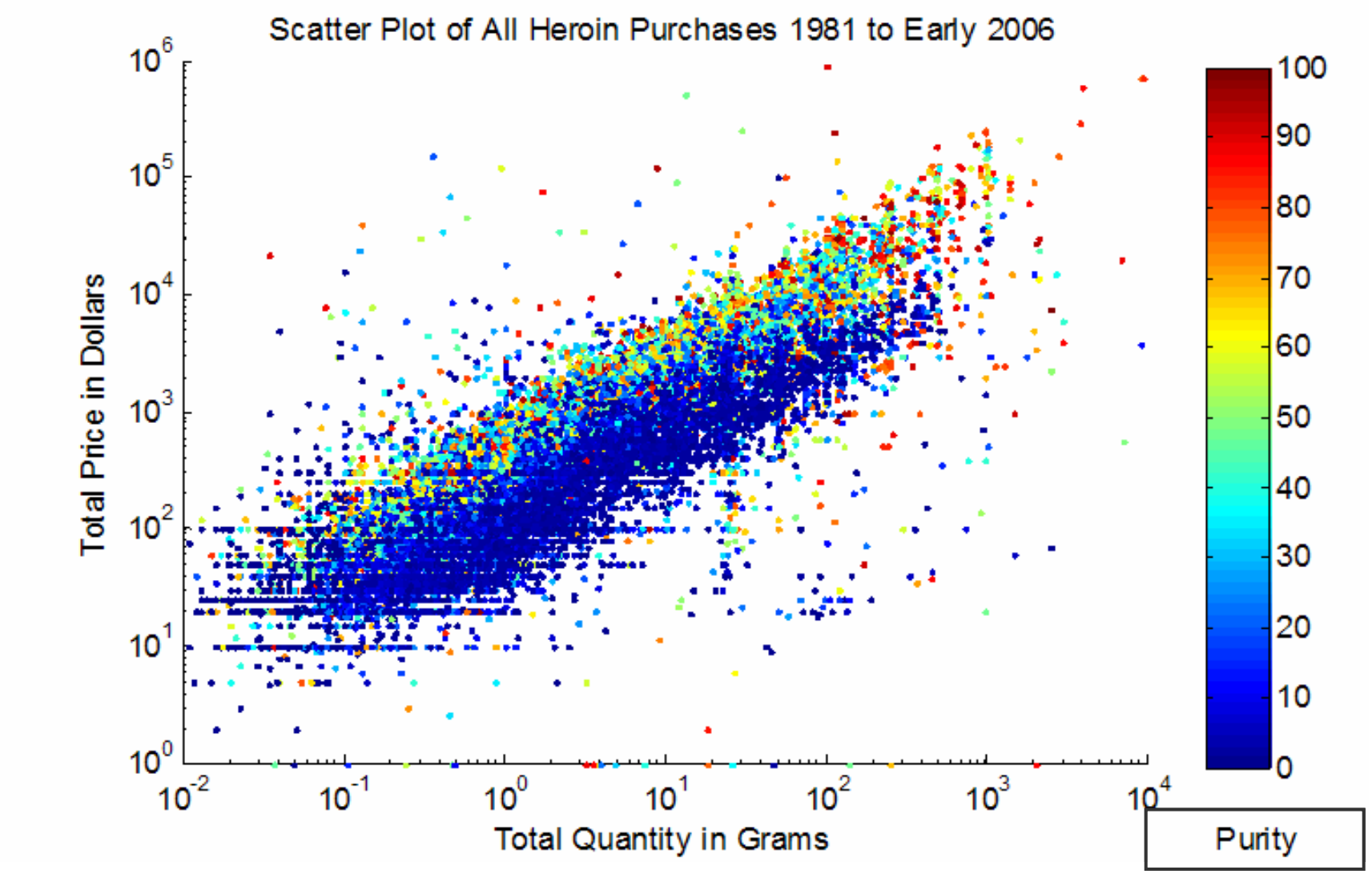

Figure III-1. Scatter Plot of STRIDE Heroin Purchases

Method 3: A model of drug market behavior might be constructed based on calculating the result of competing processes of dispersion versus convergence. The mathematics has been worked out by physicists for thermodynamic processes that have been frustrated by non-equilibrium diffusion from achieving a thermal equilibrium. The resulting distribution of individual particle energies has a characteristic form with long tails and poorly behaved variance. If the shape of the distribution of prices for standard transaction levels could be shown to result from an analogous economic process, this modeling approach might yield insight into the drug market transaction processes and provide a possible basis for estimating the underlying target price.

c. Unknown Heavy User Buying Patterns

Problem: Heavier users consume most of the supply for each type of drug, but the quantities of their purchases, hence their price discounts, are not known. Therefore, even if we knew the actual "prices" of drugs by quantity level, we still could not compute the total dollar transactions represented by the drug market as a whole. For example, if heavy cocaine users typically binge for several days, the bulk of the sales would, at minimum, be in the few gram range necessary to support a binge or, at maximum, be skimmed off at a dealer's discount for ounce purchases.

Although one of the principal topics of value is an accurate estimate of the revenues for illicit drug markets, all current estimates rely on assumptions concerning buying patterns of traffickers at higher transactional levels and of heavy users at lower transactional levels.

Method: Without an understanding of heavy user purchasing patterns, possibly obtained from anthropological studies of drug users, no method can accurately estimate illicit revenues. In general, without independent knowledge of the pattern of sales at all levels of transaction, revenues from those levels will also be dependent upon expert assumptions about those patterns. Law enforcement experience and intelligence insights might provide a sufficiently accurate understanding of the frequencies of transactions at different levels and for heavy user buying patterns.

d. Shifts in Types and Forms of Drugs

Problem: Drug preferences are constantly changing as a consequence of drug availability and periodic local fads. There may even be a continuum of brands, dosage forms, drug types, and mixtures representing an unregulated marketplace. Each form, brand, or supplier establishes some price and purity expectations among the user

population. Therefore, price and purity estimates involve assumptions regarding what forms or types of drugs to aggregate or disaggregate.

Method: For estimating long-term trends, analysts need to check whether price and purity differ significantly for the different forms of each drug. (Form might also include average packet size when transaction quantities consist of many packets.) If there are significant differences, separate series would probably have both law enforcement and policy value. Correlations among the time series for different types of drugs or forms of a single drug could provide insight into whether and how they compete with one another.

2. STRIDE-Specific Sampling Limitations

Some analytical limitations arise from potential biases of the law enforcement priorities that determine the seizures or purchases of illicit drugs reported in STRIDE. Because STRIDE is a convenience sample of data, it presents a variety of analytical problems.

a. Non-Random Sampling Rate

Problem: Law enforcement objectives determine the rate of sampling so that it is not uniform over time. Grouping these samples into equal time intervals produces different numbers of counts in each interval. Thus, the statistical uncertainty varies from interval to interval in the resulting time series.

Method 1: Estimates of statistical uncertainty can be computed and reported for each time interval. There are methods for proper statistical analysis of time series with varying counts per time interval; however, these methods are much more cumbersome and can encounter technical limitations.

Method 2: Variations in the STRIDE sampling rate might reveal shifts in law enforcement focus or surges of law enforcement attention. Statistical tests for time correlations for surges or troughs in sampling rates might be shown to correspond to law enforcement operations or shifts in counter-drug focus. If measured prices and purities of drugs remained constant through such surges or troughs, this would argue against there being a significant bias in those measures from one time interval to another. However, if surges in sampling did correspond to changes in price and purity, it would be difficult to separate the influence of the law enforcement activity from the possibility of bias associated with greater sampling effort.

Method 3: The fraction of drug samples from each transaction quantity level can be checked to see if it remains steady over time. If not, this suggests a systematic shift in sampling emphasis and a risk that sampling might no longer be as representative for those transaction quantity levels receiving less attention. Also, such long-term systematic shifts would affect indicators that combine all quantity levels, such as the global median.

b. Under-Sampling Hidden Flows of Drugs

Problem: By the very nature of illicit markets, drug flows are not readily visible to law enforcement or those who determine the sample frequency for purchases. For example, law enforcement may focus on violent and socially disruptive segments of the illicit drug markets more than on the less disruptive exchanges and consumption patterns. Such a focus could skew the choice of informants and their resulting purchases; it would also skew the probability of seizure in that component of STRIDE data.

Methodological Comment 1: This bias might be inherent to illicit markets – the less visible flows will be under-represented in samples. This bias further undermines the use of STRIDE as a measure of the "true" price or purity of illicit drugs. For similar reasons, there has been continual controversy surrounding efforts to balance the total flows of drugs as they move down the distribution chain to final consumers.

Methodological Comment 2: If there are ample indicators of consistent sampling of the more visible transactional flows, then time series constructed from STRIDE should faithfully reveal price or purity shifts in those flows. Assuming that the visible flows compete with and therefore reflect the behavior of the undetected flows, price and purity changes from STRIDE data should represent similar changes for that drug. This hypothesis could be partially tested by comparing changes in price and purity of different forms of a drug, e.g., powder versus crack cocaine.

Methodological Comment 3: Random effects methods cannot compensate for this bias because there is no means of determining that the sampling process represented by the random effect model is itself unbiased. This inherent dilemma cannot be resolved by analysis of STRIDE data alone; it could be resolved only by exogenous sources of understanding of the distribution of less visible flows and the likelihood of undersampling those flows. One possible means of resolution would be to present a convincing argument for the *a priori* mathematical form of the distribution to be sampled by the random effects. Distortions in the observed STRIDE sampling distribution relative to the known distribution might compensate for the undersampled data. Unfortunately, for illegal drug markets, one cannot assume *a priori* that the sampling

follows a normal distribution. Moreover, this technique requires very large data sets to implement, and STRIDE might not be sufficiently robust.

c. Different Sampling Practices by Region or City

Problem: Law enforcement in different regions or cities might focus on different drugs or forms of a drug, and, therefore, informant access might differ from place to place. Therefore, the drug purchases might not represent local drug prevalence or sample many of the distribution channels for a given drug. Note that drug availability and user prevalence most likely also vary by region or city; therefore, one cannot assume uniform behavior, flows, or rates as a basis for compensating for a law enforcement bias, and local consumption does not necessarily scale with local population.

Method 1: Time correlations of price or purity fluctuations throughout a region or among neighboring cities would be evidence for a real effect representing a distribution chain relationship or consumption pattern of preferences among the correlated locations. However, the observed price level in each region or city might still be biased according to the peculiarities of local sampling focus. Analysts should check that surges in local sampling are not correlated with abrupt changes in price or purity values.

Methodological Comment: Because sampling rates might reflect law enforcement interest rather than the prevalence of the drug in that the region or city, a true estimate of a national price for a drug would require weighing the samples according to true prevalence of abuse. However, true prevalence is seldom known, especially for each locality. Using population data for weighting assumes a uniform prevalence of each drug across all regions – an unlikely pattern. Conversely, using independent drug testing data on prevalence could establish regional and local patterns that might then be employed to corroborate STRIDE analyses of price, purity, drug type, and drug form.

3. STRIDE Data Management Limitations

Some problems with interpreting and analyzing STRIDE data arise from errors, misunderstandings, or lack of detail.

a. Data Entry Errors and Outliers

Problem: Some data describe highly implausible transactions, such as 16 kilograms of cocaine purchased for less than $25,000. However, the drug markets are known to be quite variable, so being unusual by itself should not exclude a transaction. Moreover, the statistical distribution of prices deviates from the generally expected

Gaussian (normal distribution). Therefore, "exceptions" in long tails of the distribution can be representative transactions, and should be retained.

Method 1: Accumulate suspicious "outliers" in a separate term to determine whether their inclusion or exclusion would affect the final result. If not, the problem is minor and can be ignored.

Method 2: Averages can be very sensitive to a few outlier values, while medians are not. Therefore, if the topics of interest can be addressed by methods based on medians, averages could be avoided.

Various statistical methods could be employed to validate the use of medians or other similar indicators as the basis for addressing topics of interest, especially comparisons such as evaluating abrupt time changes or regional differences. These methods provide statistical measures of the probability of the change or difference occurring at random. By validating the identification of real changes for known events, the methods can establish trustworthiness for indicating current change or difference.

For smooth distributions that arise in drug markets, variance about a median is always a well-behaved statistical measure (as long as the median corresponds to a quantity within a preferred level and not pathologically between levels). Also, percentiles are well-behaved measures because they are always defined and easy to interpret.

Method 3: If there are too many values identified as outliers to ignore, consider that the outliers are legitimate components of a highly variable illicit market. One way to test this hypothesis is to examine the form of the statistical distribution including the outliers. If those outliers form a smooth continuum, especially an inverse power law tail characteristic of a Pareto-Levy distribution, this is evidence that those outliers are in fact part of a coherent but highly variable process (e.g., a "fractal" marketplace).

b. Misunderstandings and Lack of Detail

Problem: STRIDE has evolved over many years, and some modifications had to be implemented within the confines of the existing data structure. This has led to confusing or apparently illogical data values being employed to designate a special situation.

Method: As STRIDE analysts expand the scope of their studies, consultation with appropriate DEA staff should be encouraged. At a minimum, the deletion of data values that cannot be properly interpreted should not affect the aggregate result.

CHAPTER IV

SCATTER PLOTS OF PURE QUANTITY-PRICE DATA

IV. SCATTER PLOTS OF PURE QUANTITY-PRICE DATA

The preceding chapter displayed a scatter plot of prices and purities for STRIDE heroin purchases (Figure III-1). Comparable graphical depictions have appeared in other sources, e.g., to contrast the slopes of quantity discount curves for different illicit drugs,[1] and to analyze subsets of STRIDE data via regression methods.[2] Section A of this chapter presents similar scatter plots, incorporating format changes to more clearly reveal fixed price levels (vertical clusters of data points) and fixed quantity levels (horizontal clusters of data points), as well as their respective dispersions by purity. Section B relates these scatter plots to methodological approaches for analyzing STRIDE data.

A. SCATTER PLOTS

Figures IV-1 to IV-3 present three pairs of scatter plots – respectively for cocaine (all variants), heroin (all variants), and methamphetamine (all variants) – all extracted from the ONDCP version of the STRIDE database provided to IDA. Transactions that involved zero purity sales were omitted from the analysis. For a given drug type, each pair of side-by-side scatter plots portrays STRIDE price and purity data from two distinct years. The years have been selected to highlight major feature identified in the *Results Report*.

On a log-log scale, the vertical axis tracks the total pure quantity (in grams) of a transaction and the horizontal axis gives the sales price (in dollars, unadjusted for inflation) for that same transaction. The individual data points are color coded to reflect the purity of the purchased drug, running from "blue" for low purity to "red" for high purity. Through each collection of data, a best-fit line is displayed. The data always exhibit a high degree of variability, but general adherence to a linear slope remains evident. The bottom of the figure gives the equation of that line, the number of data points, and the time period encompassed.

---

1 R. Anthony and A. Fries, "Empirical Modelling of Narcotics Trafficking from Farm Gate to Street," *Bulletin on Narcotics*, United Nations Office on Drugs and Crime, Volume LVI, Nos. 1 and 2, 2004, pp. 1-48.

2 J. Horowitz, "Should the DEA's STRIDE Data Be Used for Economic Analyses of Markets for Illegal Drugs," *Journal of the American Statistical Association*, 96, No. 456, December 2001, 1254-1262.

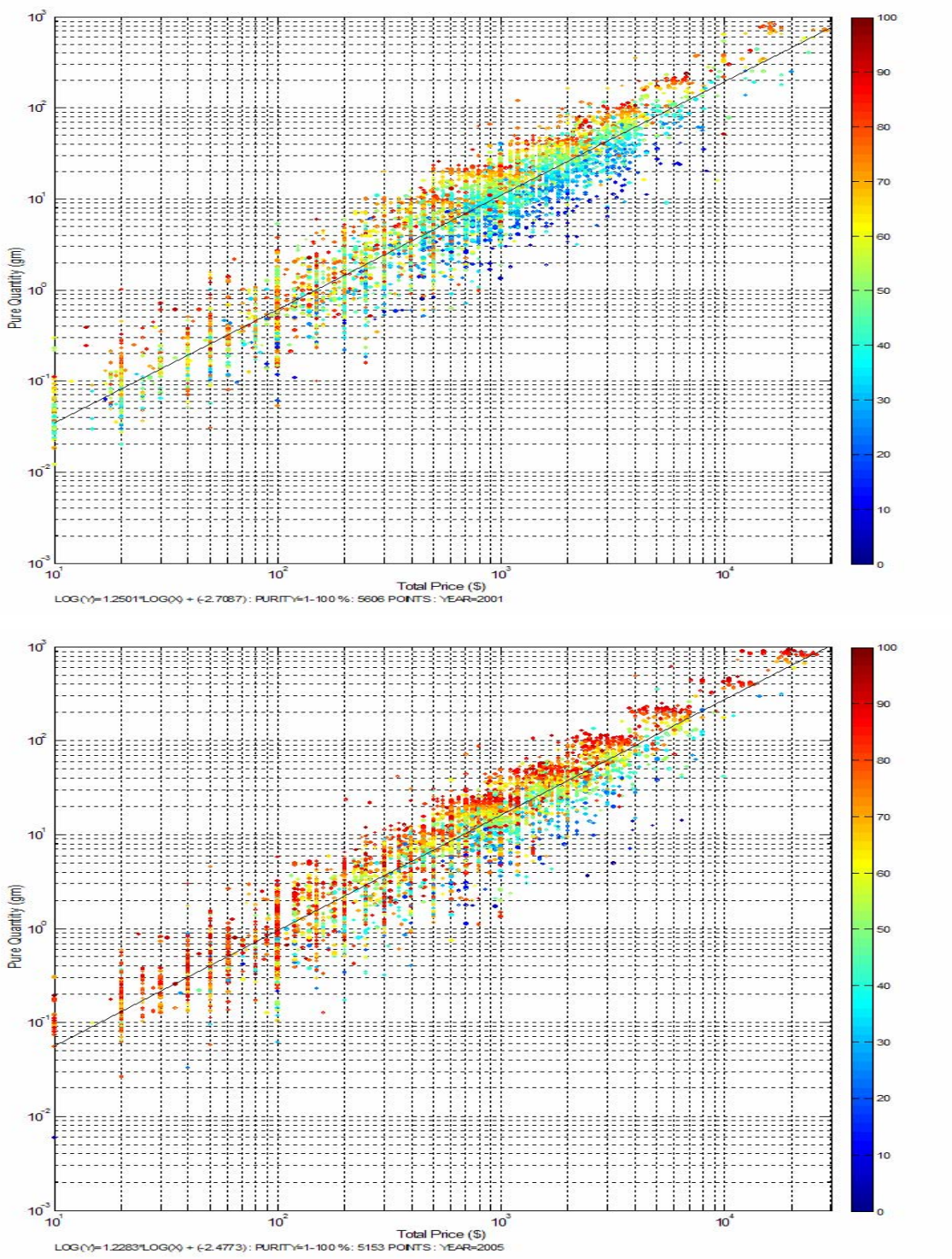

Figure IV-1. Scatter Plots of Pure Quantity-Price Data (Color Coded by Purity) for STRIDE Cocaine Purchases: Years 2001 (top) and 2005 (bottom)

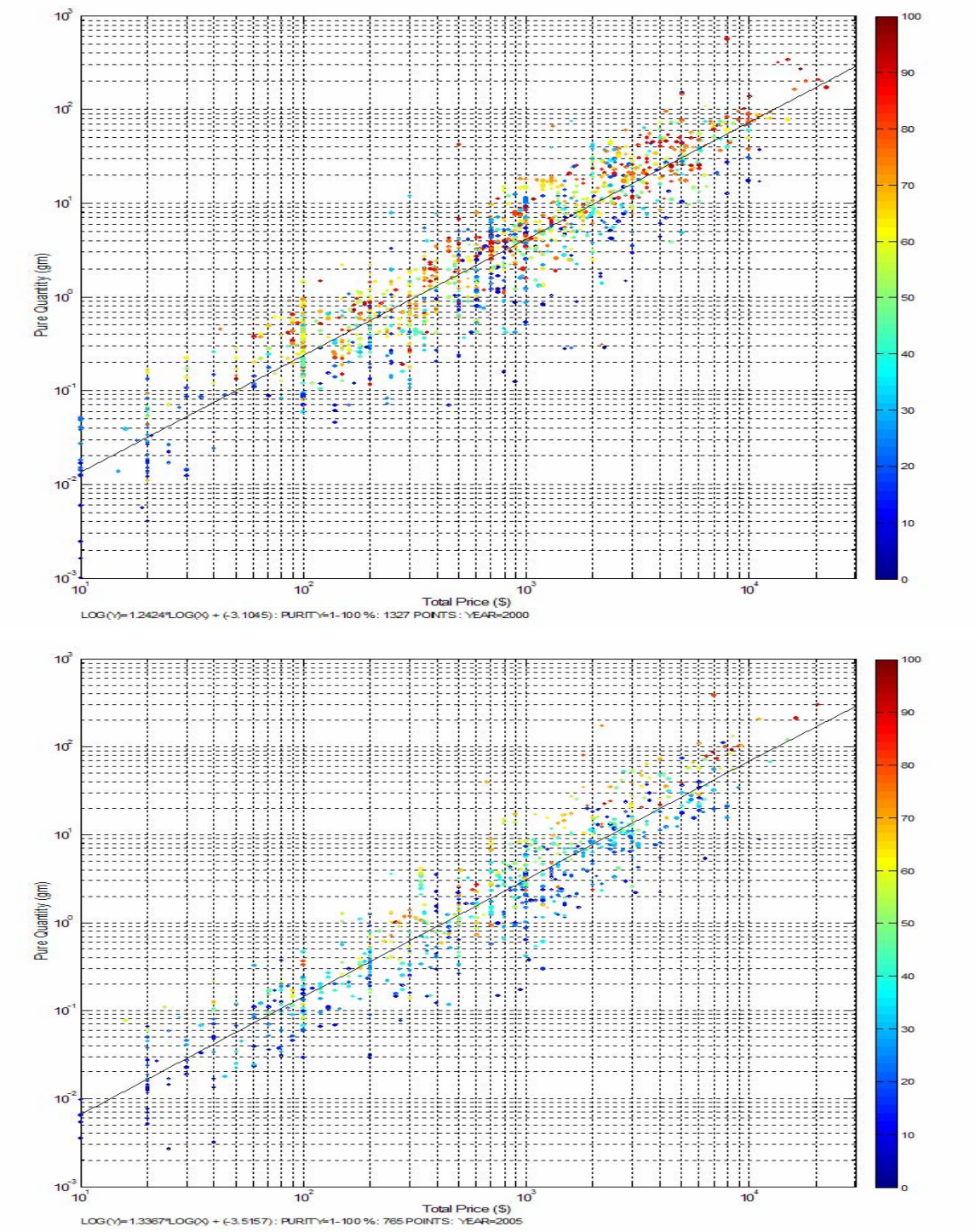

Figure IV-2. Scatter Plots of Pure Quantity-Price Data (Color Coded by Purity) for STRIDE Heroin Purchases: Years 2000 (top) and 2005 (bottom)

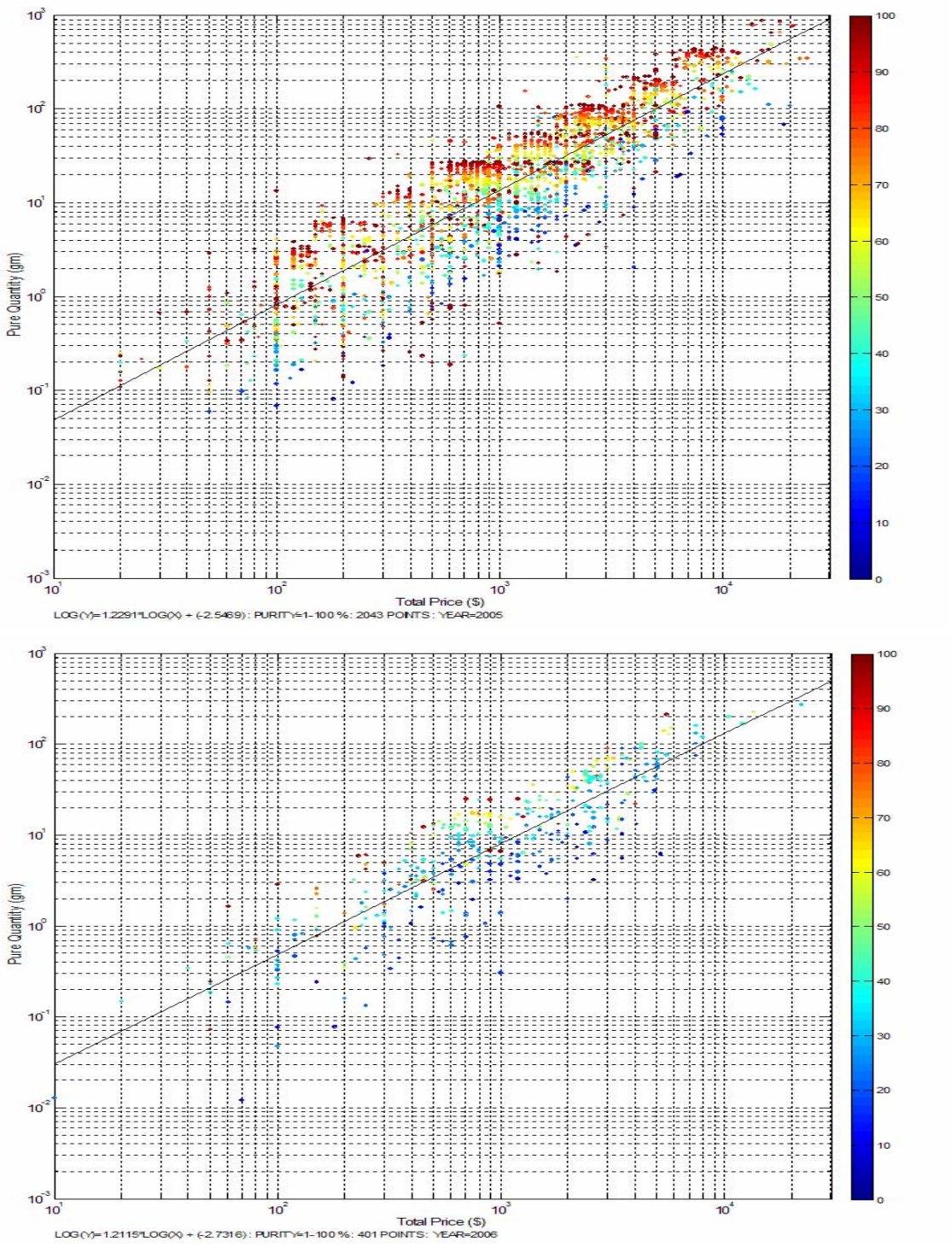

Figure IV-3. Scatter Plots of Pure Quantity-Price Data (Color Coded by Purity) for STRIDE Methamphetamine Purchases: Years 2005 (top) and 2006 (bottom)

If drug traffickers throughout the distribution chain made no profit, the theoretical slope of the fitted equation would be exactly 1. The degree to which the slope exceeds 1 is indicative of the price mark-up that traffickers impose on their transactions. Fitting a straight line through the entire span of the data is tantamount to assuming, in an average sense, a constant mark-up at each distribution level, e.g., wholesale and retail. Near constancy of the estimated slope over a time period is consistent with an illicit drug market in which the profit margin for traffickers has remained steady.

The annual plots for cocaine show a remarkable consistency for the estimated slope. The values all lie between 1.20 and 1.28 from 1981 to 2006. The two selected plots, depicted in Figure IV-1, clearly show the transformation between 2001 and 2005. The density of red data points is more prevalent in the latter figure (indicating an increase in purity), and the entire curve has been shifted upward (indicating an increase in average pure quantity per unit price, i.e., a decrease in the average price per pure gram).

Figure IV-2 displays two similar scatter plots for heroin, now for the years 2000 and 2005. The estimated annual slopes of STRIDE data between 1981 and 2006 vary between 1.17 and 1.34. Transitioning from 2000 to 2005, both the average purity and the overall data variability decreased. Also, the fitted straight line has moved downward, i.e., starts (to the left) at a smaller pure quantity value and ends (to the right) at the same pure quantity value.

The Figure IV-3 curves for methamphetamine contrast the years 2005 and 2006. The relative change is the opposite of that displayed in Figure IV-1; i.e., when advancing in time, the purity is seen to decrease while the average price per pure gram increases. The count of STRIDE methamphetamine samples is low, and the estimated slopes exhibit more variability. For the time period 1981 through 2006, the observed values span from 1.07 to 1.37.

B. METHODOLOGIES

The STRIDE analysis methodology suggested by Figures IV-1 to IV-3, i.e., imposing a regression structure to link pure amount and price, in some ways can be viewed as being intermediate to the EPH modeling construct (which was the focus of the *Results Report*) and simple median-based methods (discussed in Chapter V of this *Technical Report*). Similar to the regression approach taken in this chapter, the EPH analysis approach relates the log of the price to the log of the amount, but it incorporates additional complexity by permitting different coefficients for distinct geographical areas.

The EPH formulation actually performs a separate regression for each quantity level, but no attempt is made to check to see whether the estimates are compatible at the end points of the quantity levels. This is notionally equivalent to implementing the general quantity discount portrayal while partitioning the quantity axis into distinct regions, fitting a separate straight-line fit within each region, but not imposing any constraints to require the end points of the fitted line segments to coincide.

The classical median method, on the other hand, aggregates all of the subject data without any consideration of the specific value of the quantity for a given STRIDE sample (other than it falls within some prescribed quantity level). This suggests that a plausible first-order adjustment to median-based methodologies can be introduced, by incorporating a straightforward regression structure that accounts for the contribution of varying amounts within a given quantity level. Likewise, a similar adjustment procedure could be applied to methods based on means vice medians. This latter approach is utilized in the analyses summarized in Chapter VI, dealing with the construction of separate time series of estimated price for distinct variants of heroin.

CHAPTER V

COMPARISONS OF EPH- AND MEDIAN-BASED ESTIMATES

V. COMPARISONS OF EPH- AND MEDIAN-BASED ESTIMATES

A. OVERVIEW

This chapter compares the EPH-based national indices to time series constructed from simple median estimates. Each point of a median-based index is obtained as the sample median, or 50^{th} percentile, of the *entire* population of relevant STRIDE data points, i.e., aggregated across all geographic locations irrespective of the perceived importance or weight ordinarily attributable to each locale. In other words, the sampling and frequencies of the DEA, and other agencies contributing to STRIDE, induce an implicit weighting function. Median metrics can be computed separately for purity and for normalized price (each recorded transaction price is divided by its associated purity). The medians are not purported to represent national measures, but merely serve as simple representations of the observable data against which the EPH-derived results can be compared. The medians thus provide a context for potential detailed assessments that could motivate possible enhancements to the EPH formulation and/or suggest suitable alternative methodologies.

Median-based approaches are in contrast to regression-based methods, such as the EPH formulation, that *disaggregate* STRIDE data into distinct geographical units, estimate a summary statistic for each, and then reassemble all these into a nationally representative metric via a weighted linear combination. The individual combinations of geographical unit, calendar quarter, drug type, and quantity level can parse the STRIDE database so fine that few data are available to support statistical calculations. Moreover, what data are present can be extremely variable (considering intrinsic data randomness as well as volatility in sampling processes).[1]

IDA first published median-based portrayals of STRIDE data trends in a 1997 chronicle of methods that had been used within the interdiction community to support

---

[1] Within this context, the 2004 *Technical Report* (p. 28) notes: "We recognize that developing a national average from these relatively sparse and unrepresentative data is not advisable for a number of reasons."

operational assessments of counter-cocaine operations in the source and transit zones.[2] The median-based metric aggregated over all geographic locales as well as over all volume levels, the latter data consolidation justified by separate studies that established homogenous DEA sampling patterns across the distribution of transaction volumes. The metrics explicitly were acknowledged to be measures of the content of STRIDE, and not direct measures of retail price and purity. Additional corroboration for the viability of the derived median-based indices as indicators or price and purity was provided by strong correlations with other databases indicative of illicit drug use, and by persistent correspondences with major counter-cocaine operations.

These IDA medians were computed on an equal sample size basis, vice for predetermined time intervals (e.g., weekly) that would entail different, and possibly very small, sample sizes. STRIDE observations were chronologically ordered and separated into samples of size 100, and a median STRIDE value and a median time value were calculated for each such batch – determining the coordinates of a point on the resultant index. A follow-on formal time series analysis of median normalized cocaine prices segregated STRIDE data into monthly vice equal-sample size bins, because standard time series methodologies and software packages are defined only for equally spaced time points.[3] That study further corroborated the viability of median-based metrics, showing that the impact on domestic markets of counter-cocaine interdiction operations varied logically with the intensity of the activities and their relative proximity to domestic shores.

B. SIDE-BY-SIDE COMPARISONS

The side-by-side comparisons of the quarterly EPH-based estimates (solid curves) and their median counterparts (dotted curves) are portrayed in Figures V-1 to V-5. Separate curves are depicted for each quantity level, with the EPH estimates evaluated at the indicated intermediate point within the quantity level. The medians are determined for the exact same data that the EPH methodology retained in its regression modeling. Median prices rely on normalization by actual purity vice expected purity. All are

---

2 B. Crane, A. Rivolo, and G. Comfort, *An Empirical Examination of Counterdrug Interdiction Program Effectiveness*, Paper P-3219, Institute for Defense Analyses, Alexandria, Virginia, January 1997.

3 S. Soneji, R. Anthony, A. Fries, and B. Crane, "Time Series Intervention Analyses on US Cocaine Prices," in *Proceedings of the Fifth Annual U.S. Army Conference on Applied Statistics*, United States Military Academy, West Point, New York, October 19-21 1999, B.B. Bodt (ed.), Army Research Laboratory: ARL-SR-110, Aberdeen Proving Ground, Maryland, July 2001, pp. 67-76.

presented on a quarterly basis, since the existing software modules yield quarterly EPH estimates. Line segments are used to connect the depicted quarterly values, but they have no physical interpretation.

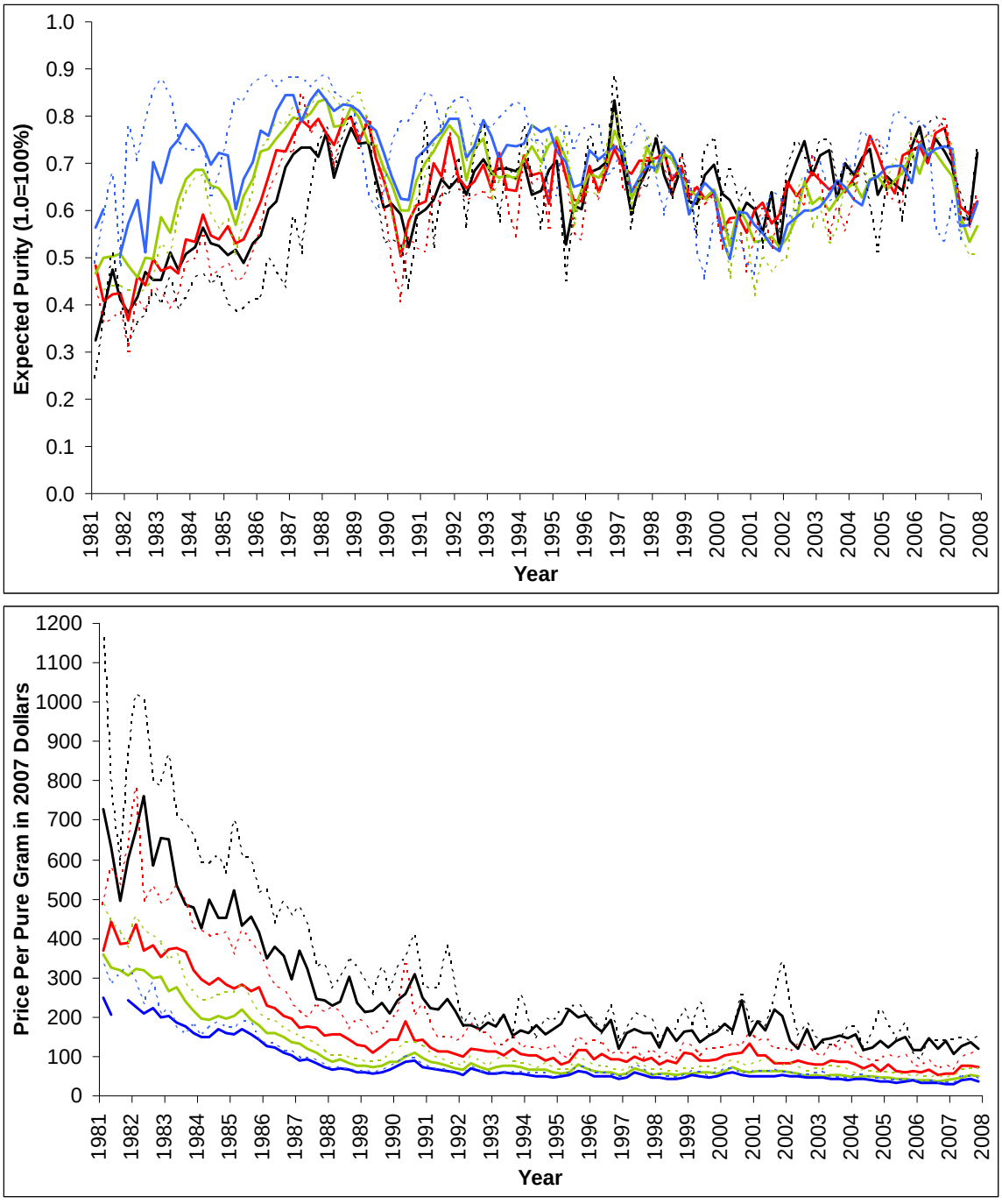

Figure V-1. EPH Predicted Price of One Expected Pure Gram of Powder Cocaine (bottom) and EPH Expected Purity of Powder Cocaine (top) Compared to Counterpart Medians (≤ 2 g @ 0.75 g, 2 to 10 g @ 5 g, 10 to 50 g @ 27 g, > 50 g @ 108 g)

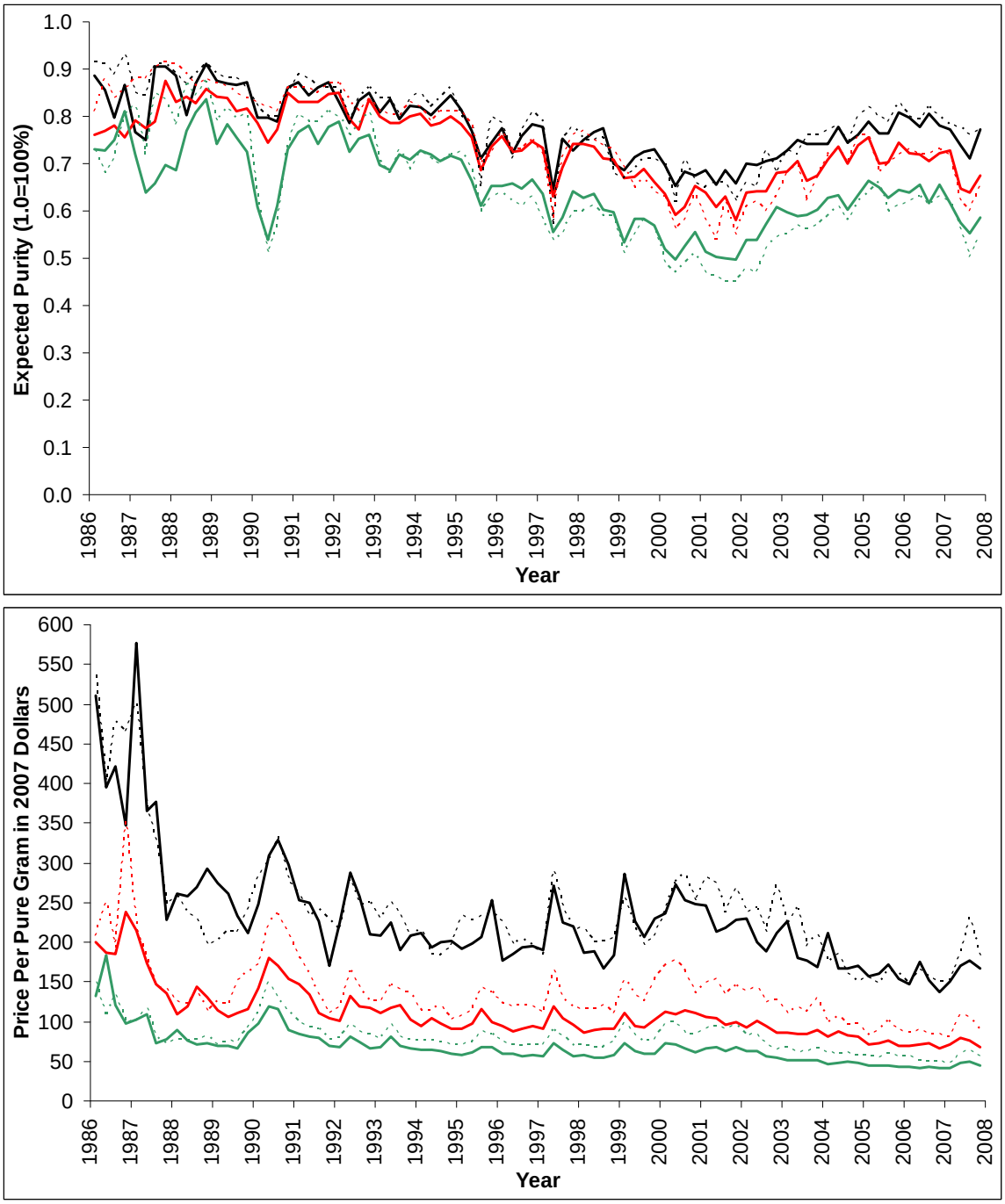

Figure V-2. EPH Predicted Price of One Expected Pure Gram of Crack Cocaine (bottom) and EPH Expected Purity of Crack Cocaine (top) Compared to Counterpart Medians (≤ 1 g @ 0.3 g, 1 to 15 g @ 5 g, > 15 g @ 38 g)

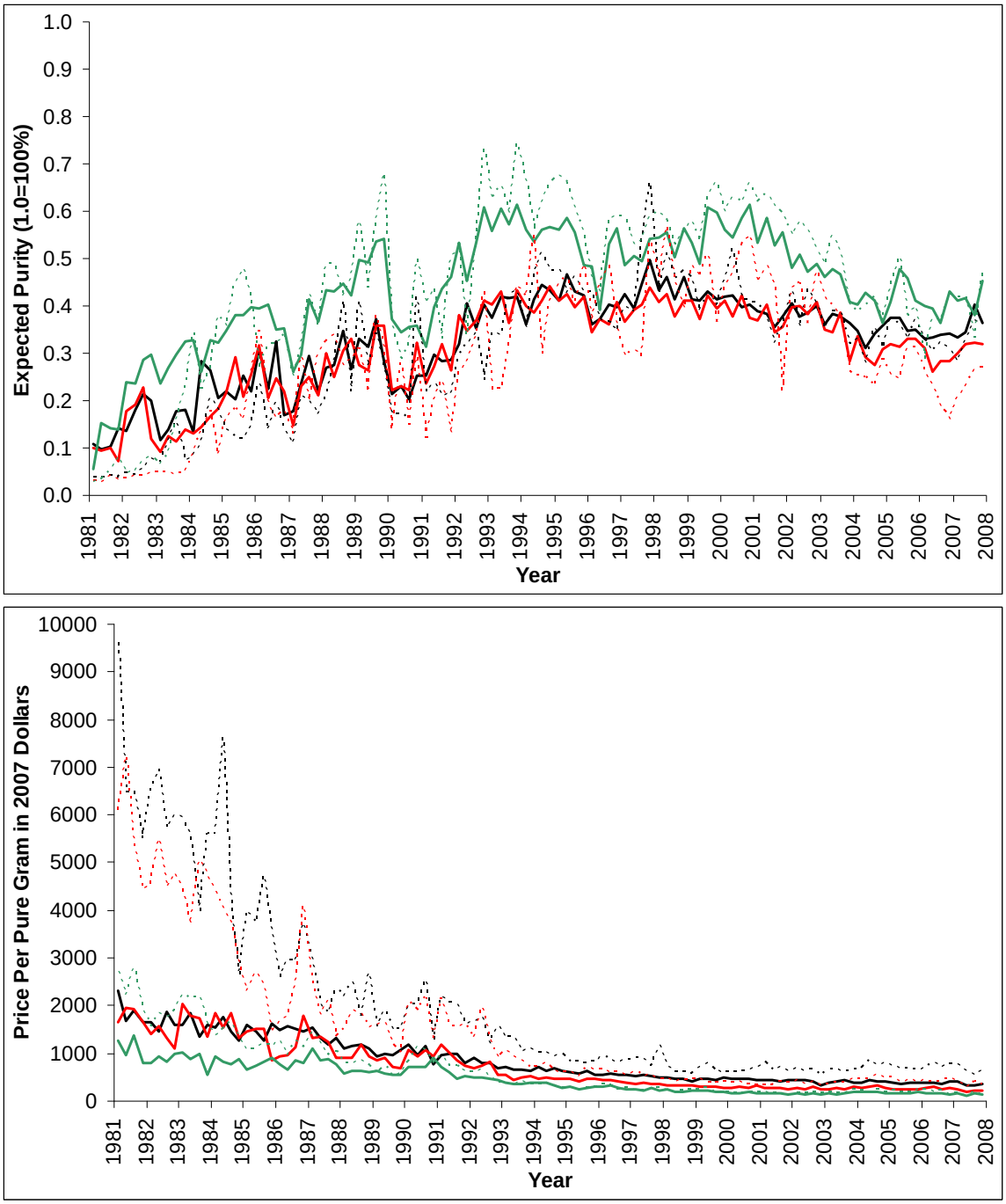

Figure V-3. EPH Predicted Price of One Expected Pure Gram of Heroin (bottom) and EPH Expected Purity of Heroin (top) Compared to Counterpart Medians (≤ 1 g @ 0.4 g, 1 to 10 g @ 2.5 g, > 10 g @ 27.5 g)

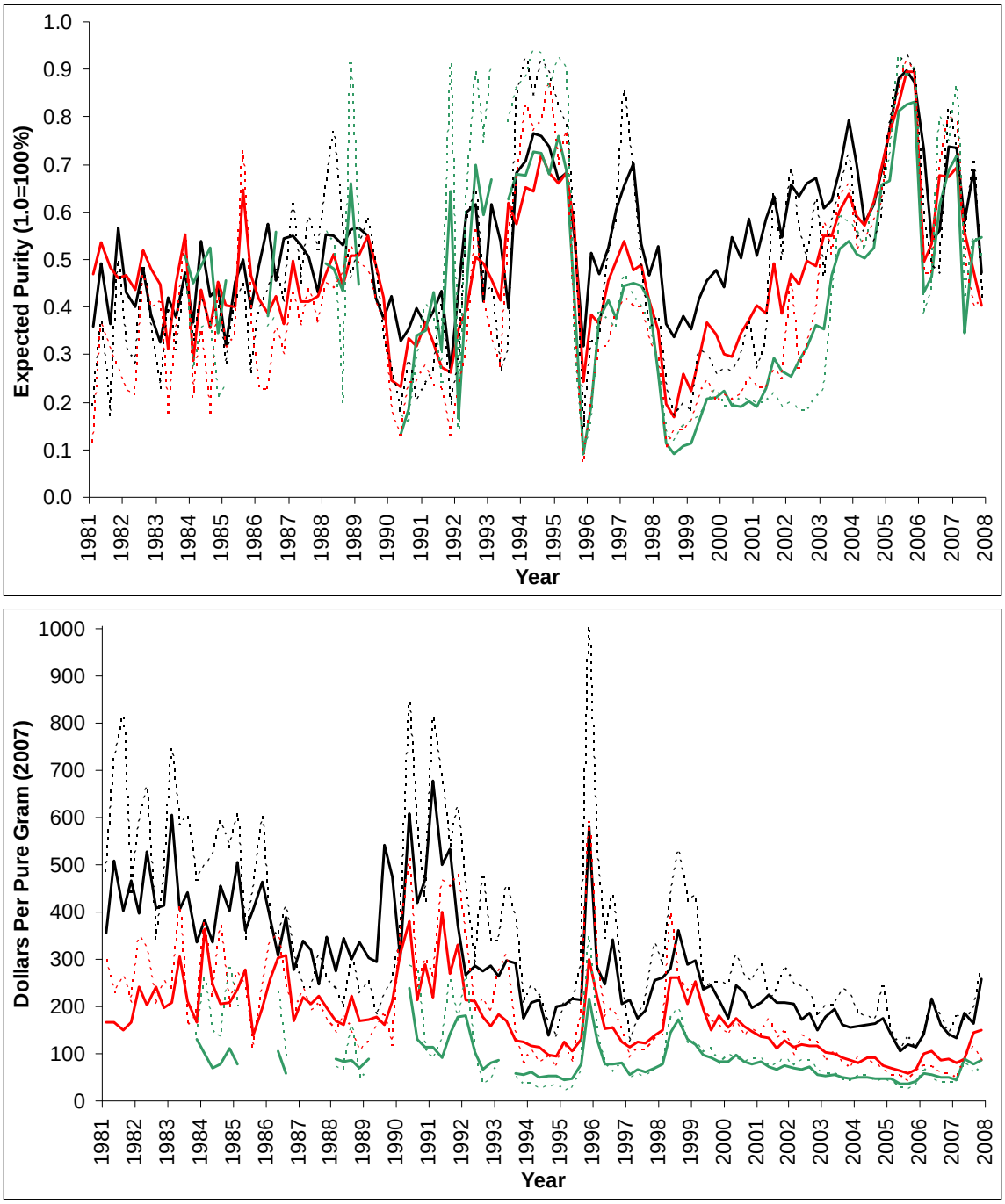

Figure V-4. EPH Predicted Price of One Expected Pure Gram of d-Methamphetamine (bottom) and EPH Expected Purity of d-Methamphetamine (top) Compared to Counterpart Medians (≤ 10 g @ 2.5 g, 10 to 100 g @ 27.5 g, > 100 g @ 225 g)

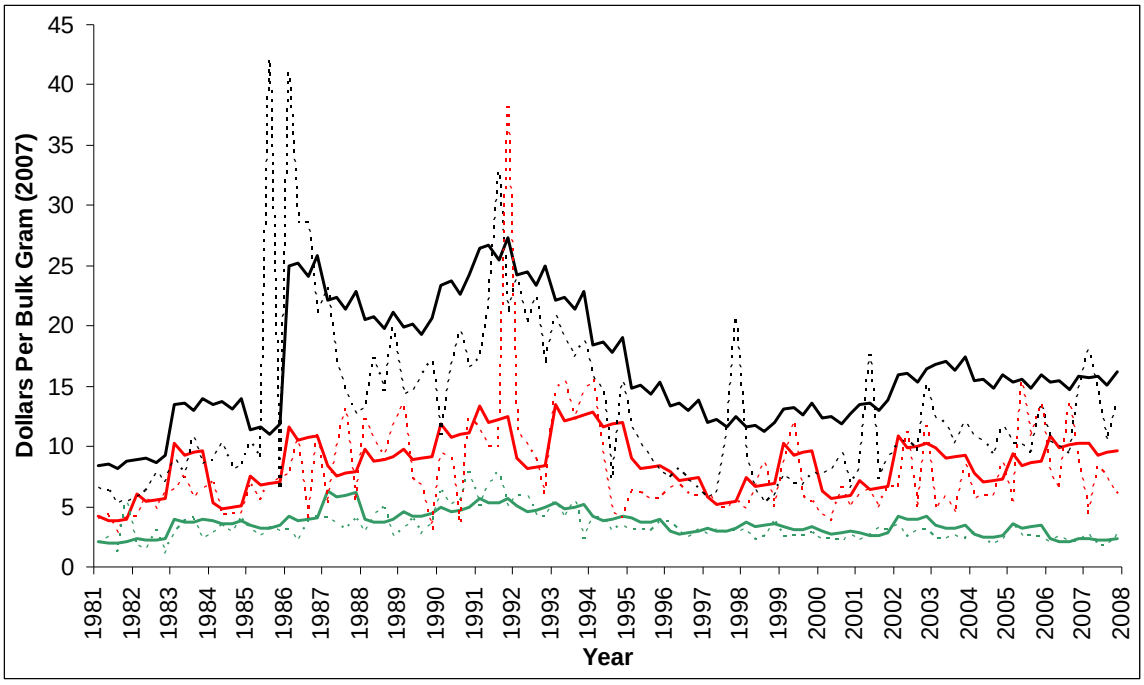

Figure V-5. EPH Predicted Price of One Bulk Gram of Marijuana Compared to Counterpart Medians (≤ 10 g @ 2.5 g, 10 to 100 g @ 26 g, > 100 g @ 443 g)

Several overall observations are evident. First, the EPH quarterly estimates are much more variable than their yearly averaged values. The exhibited vacillation suggests that not too much significance should be attached to any result reported for a single or even two consecutive quarterly estimates, especially in light of the inherent variability in the STRIDE data and underlying sampling processes. The EPH yearly estimates can be viewed as a smoothing procedure that reduces the apparent variation, but it also diminishes the apparent magnitude of features and departures from long-term trends. Similar comments also apply to the median values, which show even more variation than their EPH counterparts.

A second observation is that the figures reveal nearly universal agreement in major trends and features, although there certainly are differences in precise levels and timing. This general agreement can be viewed as a confirmatory direct comparison of the EPH results to the raw data.[4] There are some instances when the absolute magnitudes of the two sets of estimates vary substantially, but the trends nonetheless remain common. Detailed study of some of these examples suggests that small sample circumstances are

[4] The 2004 *Technical Report* presents Akaike Information Criteria test statistics for assessing model goodness of fit, but these merely compare alternative model representations against each other.

the cause. Here, the EPH model weights very heavily the few available observations from major cities and discounts the many but disparate data from the remainder of the country. In contrast, the median method weights all of the STRIDE samples equally.

Table V-1 illustrates the situation for the case of d-methamphetamine at the lowest quantity level in 2001. For these circumstances, the EPH estimates of purity exceed 50 percent while the median value counterparts are near 20 percent. The higher EPH estimates are determined essentially by a total of only 28 major city observations, scattered across eight major cities and the four quarters of the calendar year.[5] No city-quarter combination has more than four observations. It is the remainder of the country, 250 data counts in total but not captured within Table V-1, that drive the overall national aggregation medians down under 20 percent. A fundamental question here is to what degree should 10 percent of the available national data dominate the national metric?

Table V-1. STRIDE Data Counts for d-Methamphetamine Quantity Level 1 in 2001 – By Major City and Quarter

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	TOTAL
Chicago	0	0	3	2	5
Dallas	1	0	0	0	1
Los Angeles	0	2	1	0	3
New Orleans	0	0	0	1	1
New York	0	0	2	0	2
San Diego	3	2	4	1	10
Tampa	1	0	0	0	1
Washington DC	1	1	2	1	5
TOTAL	6	5	12	5	28

When there are substantial differences between the two sets of estimates, the EPH model result generally is less extreme than the median-based estimate, i.e., high and low median values are associated with EPH counterparts that are not as close to the possible extreme values (e.g., 100 percent purity and 0 percent, respectively). At times, coherent

[5] Recall that with the EPH modeling construct the prescribed 29 major cities contribute a share of 29/30, or 97 percent, to the national index.

trends observable in the quarterly median estimates are not readily discernible in the EPH results. Reasons for these differences remain under study. An initial conjecture is that they might be attributable to, depending on the specific circumstances, one or more possible factors: low sample sizes, inherent differences in the data collected by law enforcement agencies, normalization of price by the expected purity vice the actual purity, and an intrinsic property of the specific EPH regression modeling constructs that tend to produce estimates that are "shrunk" toward a common mean value.[6]

Neither median- nor regression-based estimates by themselves address the widely acknowledged limitations of the STRIDE database – including the non-scientific sampling processes that generate the data recorded within STRIDE. Likewise, additional layers of sophisticated statistical modeling cannot overcome entirely this fundamental shortcoming (although analysis of clusters of STRIDE samples could lead to practical adjustments). Within this context, however, constructing different sets of indices based on disparate methodological approaches can prove insightful. Comparable results can be interpreted as plausible portrayals of the STRIDE content, laying the groundwork for additional confirmatory checks with relevant external databases and/or sources of information. Substantially different results, and an understanding of the underlying causes, can motivate the development of appropriate methodological enhancements.

Given the many analytical challenges that confront a comprehensively rigorous examination of STRIDE data, the merits of alternative methodologies for developing price and purity indices cannot be argued persuasively solely on theoretical grounds. For example, for a given location-time combination regression-based methods that mimic established sample survey techniques must contend with small sample sizes that can be dominated be statistical noise. Median-based methods may have a better chance to extract a real signal from the midst of noise, but the interpretation of the results can be more problematic when the underlying sampling processes that populate the STRIDE database are not consistent over time. In addition, using medians to capture changes in

---

6 The EPH methods are based on *random coefficient regression models*, which assume that the parameters being estimated adhere to sets of underlying normal distributions. This construct imposes some relationships among the parameters, and leads to estimates that can be thought of as weighted averages of standard estimates (obtained via traditional fixed regression model techniques) and the overall average of any set of estimates that are under consideration. Thus, the resultant estimate is "shifted" away from the traditional estimate in the direction of the relevant overall average, i.e., "shrunk" towards the mean. The degree of the translation depends on the relative variation exhibited by the traditional estimators and the spread across the family of related estimators. Additional detailed discussions can be found in C.E. McCulloch and S.R. Searle, *Generalized, Linear, and Mixed Models*, John Wiley & Sons, New York, New York, 2001.

STRIDE data may not be prudent under specific small sample size circumstances. For instance, sampling processes could concentrate purchases at particular quantity and/or price amounts in such a way that the 50th percentile of the observed price essentially is "trapped" into a narrow range of possible variation. Under these circumstances, sample means may prove to be more insightful. Both median- and mean-based methodologies can be enhanced by incorporating simple adjustments, i.e., regression models, which permit estimated prices and purities for individual transactions within some data aggregation to vary with the precise quantity amounts associated with the transactions. Simulation studies, which can interject various representations of nominal STRIDE data, can be pursued to explore the relative performance of these and other specific analytical approaches for detecting and characterizing short- and long-term trends.

For the present, use of complementary methodologies is a reasonable analysis strategy for intermittently monitoring data trends. For non-standard investigations (e.g., focused on smaller geographical regions, especially if areas do not correspond one-to-one with the formal definition of divisions prescribed within the EPH construct), median-based methods will be much easier to implement. In any case, median estimates will be generated much more expeditiously. Depending on the computer system used, it can take multiple days to finish a single EPH run for all of the drug types.

CHAPTER VI

CORRELATION ANALYSES

VI. CORRELATION ANALYSES

Section A below compares price data from select cities with ample STRIDE sample sizes to independent compilations of illicit drug prices obtained from local law enforcement sources (including police and DEA) as reported by NDIC.[1] Similar results serve to corroborate the interpretability of STRIDE results. Gross departures, on the other hand, may signal the need for follow-on research to understand the differences and ultimately motivate specific enhancements to STRIDE data collection processes and/or analysis approaches.

Section B follows with comparisons of STRIDE-based estimated price and purity time series for specific illicit drugs with associated time series of positivity rates (i.e., percent of tests with a "positive" outcome) from general workforce drug tests. In a classical supply-driven drug market, all three drug use indicators would move in concert – price up/down, purity down/up, and positivity rate down/up. Departures from this classic model may signal a waning user demand and/or highlight situations that warrant further study. Concordance, on the other hand, is a partial affirmation of the credibility of the major features exhibited by the subject data sets.

A. STRIDE-NDIC COMPARISONS

NDIC's Field Program Specialists collect illicit drug price data from various law enforcement officials (including local police and DEA) and semiannually tabulate the findings in series of NDIC Intelligence Bulletins entitled *National Illicit Drug Prices*. These publications compile prices recorded in 126 cities for the five illicit drugs central to our STRIDE studies (as well as 3,4-methylenedioxymethamphetamine, i.e., ecstasy).

---

1 Established in 1993, NDIC is a component of the U.S. Department of Justice and a member of the Intelligence Community. The General Counterdrug Intelligence Plan, signed by the President in February 2000, designated NDIC as the nation's principal center for strategic domestic counterdrug intelligence.

Our analyses in this section are based on results appearing in the December 2005 and June 2006 editions of *National Illicit Drug Prices*, encompassing data recorded in 2005.[2] The NDIC bulletins do not provide any sense of sample sizes, nor do they indicate whether reported prices are based on data tabulations, general impressions, or some combination of the two.

There are a number of reasons why the NDIC and STRIDE data should not necessarily be expected to match closely. They are obtained from different sources and via different data selection processes. The fundamental characteristics of the underlying sets of purchase transactions may vary substantially. Illicit drug prices are inherently volatile, and our sample sizes are limited. Finally, purity information is not an explicit factor in the NDIC depictions and yet the STRIDE transactions can span a wide range of purity values. Thus, the degree of comparability of our NDIC and STRIDE data sets should be judged accordingly, i.e., rough correspondence is likely the best that could be attained.

Our synopses of NDIC data are presented in Tables VI-1 to VI-5, contrasted side-by-side with 2005 STRIDE data representations (in 2005 dollars) for select U.S. cities. The cities generally were chosen so that the total number of STRIDE price records, across all quantity levels, is 30 or more. One exception was the inclusion of San Diego, with 26 marijuana price observations, in Table VII-5 – inserted to provide a second city for the table, and to illustrate the variation in prices for different varieties of marijuana. The columns in each table correspond to quantity values, or ranges of values, specified in the NDIC compilations. Since the associated STRIDE quantity values are not necessarily concentrated at these particular quantity levels, we imposed a regression equation structure to represent the STRIDE data. Consistent with the general construct of the EPH model formulation, we set $\log(\text{Price}) = \alpha + \beta \cdot \log(\text{Amount})$ and use all of the available STRIDE data to estimate the parameters. For each column in the table, we then substitute the quantity corresponding to that column in for the amount term and calculate the associated price from the fitted regression equation. When the equation does not fit the data very well, we additionally report a rough estimate based upon visual inspection of the local data.

---

[2] These publications were provided by ONDCP near the end of our study's term. Follow-on research should include more detailed examinations of these Intelligence Bulletins as well as of NDIC Intelligence Information Reports (summaries of regular interviews conducted with federal, state, and local law enforcement officials, often detailing trends related to availability and short-term disruptions in drug prices for a region).

Given the acknowledged differences between the two data sources, we consider the NDIC and STRIDE data to match reasonably well in Tables VI-1, VI-2, and VI-4. The agreement is somewhat less so for Table VI-3, particularly for the Orlando lower quantity level and for all of the Washington, D.C. entries. The marijuana data in Table VI-5 match reasonably well for Washington, D.C., except at the one pound quantity for which no STRIDE data are available. In that same table, the two sets of numbers for San Diego differ substantially.

Table VI-1. Reported Powder Cocaine Prices – NDIC and STRIDE

		0.1 g	.25 g	1 g	1/8 oz	1 oz	31 g	62 g	125 g	1 lb	1 kg
Boston	NDIC Jun 2006			$50 to $60		$650 to $1200					$18K to $30K
	NDIC Dec 2005			$50 to $90		$650 to $1200					$18K to $28K
(N=37)	STRIDE 2005			$79		$1048					$16K[a]
Philadelphia	NDIC Jun 2006										
	NDIC Dec 2005				$70 to $125	$800 to $1600					
(N=38)	STRIDE 2005				$125[a]	$1600					
San Diego	NDIC Jun 2006	$10 to $25	$30 to $100	$60 to $160	$120 to $160	$500 to $800				$6.5K to $10K	$12K to $18K
	NDIC Dec 2005	$10	$20 to $40	$60 to $80	$100 to $120	$500 to $800				$8K to $10K	$12K to $14K
(N=49)	STRIDE 2005	$9[a]	$19	$57	$150	$790				$7K	$13K
Washington DC	NDIC Jun 2006			$80 to $100	$125 to $150	$900 to $1200	$1000	$1850	$3700 to $4000		
	NDIC Dec 2005			$100	$125 to $150	$800 to $1200	$1000	$1700 to $2000	$3500 to $3800		
(N=32)	STRIDE 2005			$77[a]	$210	$1050	$1130	$1940	$3350		

a Data are sparse in this range of the STRIDE records.

Table VI-2. Reported Crack Cocaine Prices – NDIC and STRIDE[a]

		1 g	1/8 oz	1 oz	1/8 kg
Camden	NDIC Jun 2006	$35 to $100		$750 to $1200	
	NDIC Dec 2005	$35 to $55		$800 to $1000	
(N=45)	STRIDE 2005	$62		$546[b]	
Chicago	NDIC Jun 2006			$850	
	NDIC Dec 2005	$75 to $100		$600 to $1000	
(N=49)	STRIDE 2005	$64		$922	
New York	NDIC Jun 2006	$23 to $34	$175	$1000 to $1500	
	NDIC Dec 2005	$20 to $60	$100 to $180	$600 to $1200	
(N=150)	STRIDE 2005	$78	$202	$970	
Philadelphia	NDIC Jun 2006	$70		$800 to $1600	
	NDIC Dec 2005	$70		$1600	
(N=45)	STRIDE 2005	$88		$997	
Washington DC	NDIC Jun 2006	$100		$900 to $1200	$3700 to $4000
	NDIC Dec 2005	$100		$800 to $1200	$3500 to $3800
(N=315)	STRIDE 2005	$84[c]		$1153	$3693[d]

a Following the convention adopted in the 2004 ONDCP report, we apply the label "crack cocaine" to results derived from the analysis of cocaine base observations in STRIDE, the majority but not necessarily all of which are literally crack.

b Local data ≈ $800; fitted line does not fit well here.

c Local data ≈ $110; fitted line does not fit well here.

d Data are sparse in this range of the STRIDE records.

Table VI-3. Reported Heroin Prices – NDIC and STRIDE

		.05-.1 g	.5-1.3 g	1 g	¼ oz	½ oz	1 oz	1 kg
Baltimore	NDIC Jun 2006			$65 to $165			$2500 to $3250	
	NDIC Dec 2005	$6 to $10	$60 to $100	$70 to $100			$2800 to $3000	
(N=54)	STRIDE 2005	$10 to $17[a]	$62 to $132	$107			$1502[b]	
Chicago	NDIC Jun 2006			$70 to $200			$2000 to $3000	
	NDIC Dec 2005			$70 to $200			$2500 to $3500	
(N=56)	STRIDE 2005			$74			$1505[c]	
New York	NDIC Jun 2006							
	NDIC Dec 2005			$45 to $100			$1500 to $2000	$40000 to $80000
(N=57)	STRIDE 2005			$153[d]			$2087	$33770[e]
Orlando	NDIC Jun 2006			$70 to $100			$1960 to $2800	
	NDIC Dec 2005			$80 to $115			$2240 to $3220	
(N=108)	STRIDE 2005			$270			$2666	
Washington DC	NDIC Jun 2006					$1800 to $2200	$3700 to $4000	
	NDIC Dec 2005				$1000	$1700 to $2000	$3700 to $4000	
(N=315)	STRIDE 2005				$615[a]	$1112[a]	$2011	

a Data are sparse in this range of the STRIDE records.

b Local data ≈ $2400; fitted line does not fit well here.

c Local data ≈ $2500; fitted line does not fit well here.

d Local data ≈ $100; fitted line does not fit well here.

e Local data ≈ $60000 from one data point; fitted line does not fit well here.

Table VI-4. Reported d-Methamphetamine Prices – NDIC and STRIDE

		1 g	1/16 oz	1/8 oz	¼ oz	1 oz	1 lb
Las Vegas	NDIC Jun 2006					$900	
	NDIC Dec 2005					$600 to $700	
(N=30)	STRIDE 2005					$754	
Phoenix	NDIC Jun 2006					$500 to $650	$9500 to $11000
	NDIC Dec 2005					$800 to $900	$9500 to $10000
(N=39)	STRIDE 2005					$613	$7784[a]
Philadelphia	NDIC Jun 2006	$100 to $250		$125 to $500		$700 to $3300	
	NDIC Dec 2005	$250		$500		$3,300	
(N=49)	STRIDE 2005	$275[a]		$518		$1462[b]	
San Diego	NDIC Jun 2006	$50 to $100	$75 to $130	$120 to $150	$150 to $300	$800 to $1100	$9000 to $11000
	NDIC Dec 2005	$40 to $50	$50 to $80	$100 to $150	$140 to $250	$550 to $1100	$3500 to $11000
(N=125)	STRIDE 2005	$56	$85	$146	$248	$711	$5871[c]

a Data are sparse in this range of the STRIDE records.

b Local data ≈ $3300; fitted line does not fit well here.

c Local data ≈ $8000; fitted line does not fit well here.

Table VI-5. Reported Marijuana Prices – NDIC and STRIDE

		.5-1 g	1 g	1–3 g	2 g	¼ oz	½ oz	1 oz	1 lb
San Diego	NDIC Jun 2006	$5		$10		$30 to $50			$250 to $5.5K[a]
	NDIC Dec 2005	$5		$10		$20 to $40		$75 to $400[a]	$250 to $6K[a]
(N=26)	STRIDE 2005[b]	$107 to $126		$126 to $162		$198		$273	$517
Washington, DC	NDIC Jun 2006		$10		$20			$120 to $150	$1.2K to $1.5K
	NDIC Dec 2005		$10		$20		$75 to $160	$125 to $300	$1.2K to $3.5K[a]
(N=119)	STRIDE 2005		$14		$20		$60	$88[c]	$409[d]

a Wide ranges reflect different types of marijuana recorded by NDIC, i.e., domestic, Mexican, Canadian, and hydroponic. (STRIDE does not categorize marijuana variants.)

b The model fit is poor, but the STRIDE data, especially at the lower volumes, are clearly different than their NDIC counterparts.

c Data are sparse in this range of the STRIDE records.

d There are no data in this range of the STRIDE records; estimate is a pure extrapolation.

Some extremely wide marijuana price ranges are evident in Table VI-5, reflecting NDIC's reporting of substantially different prices for distinct types of marijuana (i.e., domestic, Mexican, Canadian, and hydroponic).[3] This may be a possible explanation for the large discrepancies for San Diego at the three lower quantity levels. For instance, some of the purchases recorded in STRIDE may correspond to the more expensive marijuana variants while the NDIC numbers might be associated with the less expensive varieties. Additional research is required to resolve the apparent differences.

Relating these observations back to the general analytical issue of constructing time series of marijuana prices, we conclude that, if feasible, follow-on studies should explore the degree to which these price spreads influence the current construction of national and local level price estimates.

B. STRIDE-QUEST COMPARISONS

Previous studies demonstrated the strong correspondence between U.S. cocaine price and purity estimates with cocaine metabolite positivity results recorded from

3 The STRIDE database does not categorize variants of marijuana.

random drug testing on general workforce sectors.[4] The data period covered in that study spanned up to 2001. Here we update these comparisons and extend their scope to additionally encompass the other illicit drugs addressed in our *Results Report*. Our source of positivity results is the Workforce Drug Testing Database provided to IDA by Quest Diagnostics, Inc., a leading commercial drug testing firm that annually performs more than 8.5 million drug tests across the U.S. The timeframe for these data begins in July 2001 and ends in June 2006.

Tables VI-6 to VI-10 present compilations of correlation coefficients computed separately for each drug type,[5] spanning pairings of variables from among the Quest Diagnostics positivity rate metric and the EPH-price and purity measures for each quantity level. Correlations are calculated based on comparisons of nationwide-quarterly estimates, running from the third quarter of 2001 through the second quarter of 2006. In the tables, the different STRIDE quantity levels (defined in Table I-1 of the *Results Report*) are denoted by Q1, Q2, etc.

One complicating factor here is that positives for the Quest Diagnostics opiates drug group can be triggered from having consumed any of a variety of substances (e.g., heroin, morphine, codeine, oxycodone, oxymorphone, hydrocodone, and hydromorphone). Thus the STRIDE to Quest Diagnostics comparisons do not directly contrast heroin-specific metrics. Another similar complication is that the Quest Diagnostics "amphetamines" drug group encompasses d-methemphatamine (the target drug group in our STRIDE analyses), other variants of methamphetamine, and other classes of amphetamines. Again, these STRIDE to Quest Diagnostics comparisons do not directly contrast metrics specific to d-methamphetamine.

Tables VI-6 and VI-7 share the same Quest Diagnostics positivity rate, based on random drug testing within the U.S. general workforce for a metabolite of cocaine. The entries in the tables strongly conform to a classical supply driven market – all price-price and purity-purity correlations are positive, all price-purity correlations are negative, and all Quest-price/purity correlations are negative/positive. Moreover, the magnitudes of the coefficients are all substantial. These associated three time series are portrayed in Figure

---

4 "Empirical Modelling of Narcotics Trafficking from Farm Gate to Street," R. Anthony and A. Fries, *Bulletin on Narcotics*, United Nations Office on Drugs and Crime, Volume LVI, Nos. 1 and 2, 2004, pp. 1-48.

5 Following the convention adopted in the 2004 ONDCP report, we apply the label "crack cocaine" to results derived from the analysis of cocaine base observations in STRIDE, the majority but not necessarily all of which are literally crack.

VI-1, in which the general increase/decrease through 2005 is evident for the estimated prices/positivity rate. Also, the general correspondence between upward and downward excursions between pairs of time series can be seen.

Table VI-6. Correlation Coefficients (EPH STRIDE Powder Cocaine - Quest Diagnostics Cocaine): Quarterly Data from July 2001 to June 2006

	Q1Price	Q1Purity	Q2Price	Q2Purity	Q3Price	Q3Purity	Q4Price	Q4Purity	Quest
Q1Price	1.0000	-0.6333	0.3443	-0.7121	0.5566	-0.5213	0.4976	-0.5800	-0.5614
Q1Purity	-0.6333	1.0000	-0.0867	0.4971	-0.4760	0.5240	-0.3774	0.4267	0.4827
Q2Price	0.3443	-0.0867	1.0000	-0.5974	0.7640	-0.6347	0.7140	-0.6147	-0.2651
Q2Purity	-0.7121	0.4971	-0.5974	1.0000	-0.6938	0.7344	-0.6431	0.6718	0.4378
Q3Price	0.5566	-0.4760	0.7640	-0.6938	1.0000	-0.9544	0.9034	-0.8621	-0.4506
Q3Purity	-0.5213	0.5240	-0.6347	0.7344	-0.9544	1.0000	-0.8581	0.8252	0.4487
Q4Price	0.4976	-0.3774	0.7140	-0.6431	0.9034	-0.8581	1.0000	-0.9709	-0.5424
Q4Purity	-0.5800	0.4267	-0.6147	0.6718	-0.8621	0.8252	-0.9709	1.0000	0.5731
Quest	-0.5614	0.4827	-0.2651	0.4378	-0.4506	0.4487	-0.5424	0.5731	1.0000

Table VI-7. Correlation Coefficients (EPH STRIDE Crack Cocaine - Quest Diagnostics Cocaine): Quarterly Data from July 2001 to June 2006

	Q1Price	Q1Purity	Q2Price	Q2Purity	Q3Price	Q3Purity	Quest
Q1Price	1.0000	-0.8654	0.7312	-0.7487	0.7712	-0.7550	-0.5884
Q1Purity	-0.8654	1.0000	-0.9103	0.9317	-0.9471	0.9141	0.6135
Q2Price	0.7312	-0.9103	1.0000	-0.8435	0.9274	-0.8902	-0.4531
Q2Purity	-0.7487	0.9317	-0.8435	1.0000	-0.9129	0.9230	0.6539
Q3Price	0.7712	-0.9471	0.9274	-0.9129	1.0000	-0.9671	-0.5618
Q3Purity	-0.7550	0.9141	-0.8902	0.9230	-0.9671	1.0000	0.6093
Quest	-0.5884	0.6135	-0.4531	0.6539	-0.5618	0.6093	1.0000

Table VI-8. Correlation Coefficients (EPH STRIDE Heroin - Quest Diagnostics Opiates): Quarterly Data from July 2001 to June 2006

	Q1Price	Q1Purity	Q2Price	Q2Purity	Q3Price	Q3Purity	Quest
Q1Price	1.0000	-0.1773	0.3356	0.0546	0.0268	0.1172	0.2784
Q1Purity	-0.1773	1.0000	-0.3943	0.7614	-0.5087	0.5496	-0.4622
Q2Price	0.3356	-0.3943	1.0000	-0.5056	0.6771	-0.3549	0.3222
Q2Purity	0.0546	0.7614	-0.5056	1.0000	-0.5415	0.6751	-0.5783
Q3Price	0.0268	-0.5087	0.6771	-0.5415	1.0000	-0.7243	0.4365
Q3Purity	0.1172	0.5496	-0.3549	0.6751	-0.7243	1.0000	-0.3428
Quest	0.2784	-0.4622	0.3222	-0.5783	0.4365	-0.3428	1.0000

Table VI-9. Correlation Coefficients (EPH STRIDE d-Methamphetamine - Quest Diagnostics Amphetamines): Quarterly Data from July 2001 to June 2006

	Q1Price	Q1Purity	Q2Price	Q2Purity	Q3Price	Q3Purity	Quest
Q1Price	1.0000	-0.8269	0.7599	-0.8041	0.8168	-0.8351	-0.6986
Q1Purity	-0.8269	1.0000	-0.7526	0.8346	-0.6902	0.7958	0.5287
Q2Price	0.7599	-0.7526	1.0000	-0.9559	0.9186	-0.9649	-0.8094
Q2Purity	-0.8041	0.8346	-0.9559	1.0000	-0.8887	0.9699	0.7210
Q3Price	0.8168	-0.6902	0.9186	-0.8887	1.0000	-0.9167	-0.7601
Q3Purity	-0.8351	0.7958	-0.9649	0.9699	-0.9167	1.0000	0.7808
Quest	-0.6986	0.5287	-0.8094	0.7210	-0.7601	0.7808	1.0000

Table VI-10. Correlation Coefficients (EPH STRIDE Marijuana - Quest Diagnostics Marijuana): Quarterly Data from July 2001 to June 2006

	Q1Price	Q2Price	Q3Price	Quest
Q1Price	1.0000	0.5042	0.2831	-0.2048
Q2Price	0.5042	1.0000	0.4715	-0.0830
Q3Price	0.2831	0.4715	1.0000	0.0310
Quest	-0.2048	-0.0830	0.0310	1.0000

The signs of the correlation coefficients in Table VI-8 that involve the Quest Diagnostics term are all inconsistent with the expected structure of a supply driven market. What is driving this pattern is unknown. As noted earlier, a positive opiates test result reported by Quest Diagnostics could have been triggered by the use of heroin or a number of other different drugs.

The entries in Table VI-9 agree impressively with the expectation of a supply driven market, despite the potential complication introduced by comparing STRIDE d-methamphetamine to the broader class of Quest Diagnostics amphetamines. The large magnitude values in the table are driven by consistent and dramatic trends. For instance, Figure VI-2 shows both the EPH expected purity and the associated positivity rate increasing steadily through 2005, with sharp decreases afterwards. The starting points for these departures do not coincide exactly in the two curves, and the fall in estimated purity around 2004 is not reflected in the positivity curve. Overall, however, the degree of concordance is strong.

The price-positivity correlation coefficients in Table VI-10 are weak. The time series for expected price is fairly flat, while that for the associated positivity is generally decreasing (especially at the higher quantity levels). The potential impact of different marijuana varieties (as noted in Section A above) on the correlation statistics is unknown.

Figure VI-1. Predicted Price per Expected Pure Gram of Cocaine (Powder Cocaine @ 0.75 grams and Crack Cocaine @ 0.30 grams) and Quest Diagnostics General Workforce Random Testing Positivity Rate for Cocaine

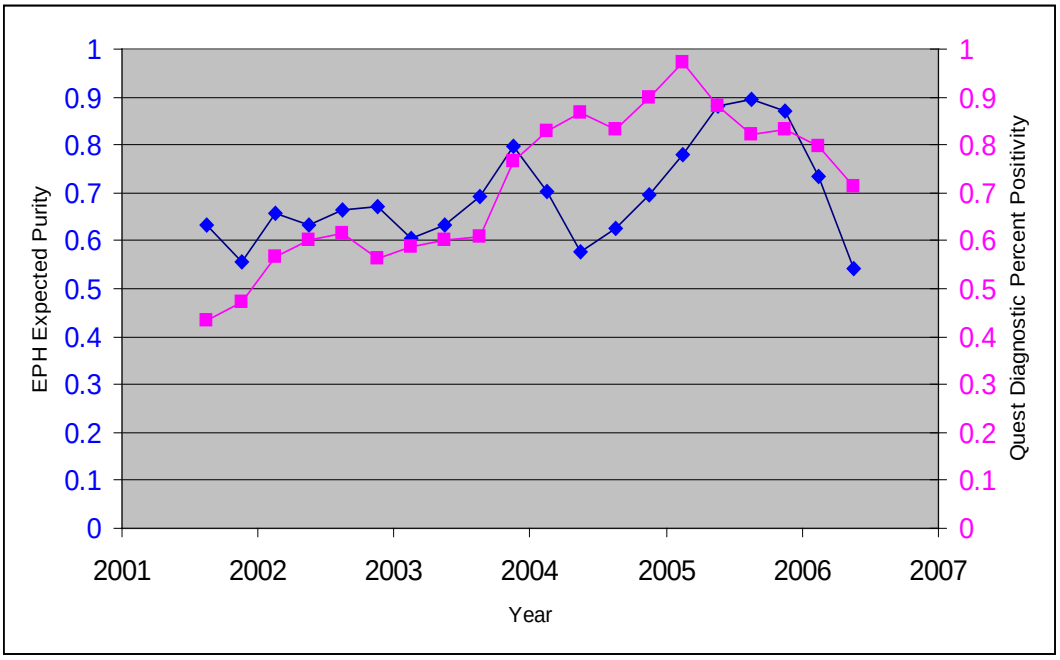

Figure VI-2. Expected Purity of d-Methamphetamine (@ 0.75 grams) and Quest Diagnostics General Workforce Random Testing Positivity Rate for Amphetamines

APPENDIX A

ACRONYMS

APPENDIX A
ACRONYMS

CPI	Consumer Price Index
DEA	Drug Enforcement Administration
DMP	Domestic Monitor Program
IDA	Institute for Defense Analyses
NDIC	National Drug Intelligence Center
ONDCP	Office of National Drug Control Policy
STRIDE	System To Retrieve Information from Drug Evidence

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
31-10-2008	IAD Final	March 2006 – July 2008

4. TITLE AND SUBTITLE

Technical Report for the Price and Purity of Illicit Drugs: 1981 – 2007

5a. CONTRACT NUMBER
DASW01-04-C-0003

5b. GRANT NUMBER
— — —

5c. PROGRAM ELEMENT NUMBER
— — —

6. AUTHOR(S)

Fries, Arthur; Anthony, Robert W.; Cseko, Jr., Andrew; Gaither, Carl C.; Schulman, Eric

5d. PROJECT NUMBER
— — —

5e. TASK NUMBER

5f. WORK UNIT NUMBER
— — —

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

8. PERFORMING ORGANIZATION REPORT NUMBER
P-4370
H08-0001216

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of National Drug Control Policy
750 17th Street, Room 708
Washington, DC 20503

10. SPONSOR/MONITOR'S ACRONYM(S)
ONDCP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
— — —

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
— — —

14. ABSTRACT

This Technical Report provides material that both complements and supplements the companion Results Report, IDA Paper P-4369. Topics addressed include IDA modifications to the data processing and statistical implementation steps undertaken in the 2004 ONDCP report (i.e., EPH methodology); expanded discussions and analyses of the challenges associated with STRIDE-centric studies; quantity discount curves; a median-based methodology for conveniently portraying STRIDE data and facilitating direct simple comparisons to the EPH model results – to provide a context for assessing the EPH formulation and to motivate potential future enhancements; and correlation analyses comparing STRIDE estimates to general workforce drug testing results and to price compilations reported by the National Drug Intelligence Center.

15. SUBJECT TERMS
llicit drugs, price, purity, powder cocaine, crack cocaine, heroin, d-methamphetamine, marijuana, STRIDE, correlation analyses, workforce drug testing

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Michael A. Cala
Unclassified	Unclassified	Unclassified	Unlimited	72	**19b. TELEPHONE NUMBER** *(include area code)* 202-395-5205

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18