UNITED NATIONS
Office on Drugs and Crime

BULLETIN ON NARCOTICS

Volume LVI, Nos. 1 and 2, 2004

Illicit drug markets

UNITED NATIONS OFFICE ON DRUGS AND CRIME
Vienna

BULLETIN ON NARCOTICS

Volume LVI, Nos. 1 and 2, 2004

Illicit drug markets

UNITED NATIONS
New York, 2006

UNITED NATIONS PUBLICATION
Sales No. E.06.XI.6
ISBN 92-1-148211-9
ISSN 0007-523X

Sandeep Chawla, Editor,
Suzanne Kunnen, Editorial Assistant

United Nations Office on Drugs and Crime
Vienna International Centre
P.O. Box 500
1400 Vienna, Austria

Telephone: (+43-1) 26060-5799
Fax: (+43-1) 26060-5866

The *Bulletin on Narcotics* is available on the World Wide Web at
http://www.unodc.org/unodc/en/ bulletin_on_narcotics.html.

The Office for Drug Control and Crime Prevention became the United Nations Office on Drugs and Crime on 1 October 2002. The Office on Drugs and Crime includes the United Nations International Drug Control Programme (UNDCP).

PREFACE

The *Bulletin on Narcotics* is a United Nations journal that has been in continuous publication since 1949. It is printed in all six official languages of the United Nations—Arabic, Chinese, English, French, Russian and Spanish.

The *Bulletin* provides information on developments in drug control at the local, national, regional and international levels that can be of benefit to the international community.

The present double issue of the *Bulletin* (vol. LVI, Nos. 1 and 2) contains a selection of papers that were originally prepared for an expert consultation on technical challenges to the drug community, held in Vienna from 3 to 5 March 2002. The United Nations Office on Drugs and Crime wishes to thank the United States Office of National Drug Control Policy, in particular Barry D. Crane, former Deputy Director for Supply Reduction, for facilitating the consultation. Thanks also go to Melissa Tullis of the United Nations Office on Drugs and Crime for editorial assistance in preparing the present double issue.

EDITORIAL POLICY AND GUIDELINES FOR PUBLICATION

Individuals and organizations are invited by the Editor to contribute articles to the *Bulletin on Narcotics* dealing with policies, approaches, measures and developments (theoretical and/or practical) relating to various aspects of the drug control effort. Of particular interest are the results of research, studies and practical experience that would provide useful information for policymakers, practitioners and experts, as well as the public at large.

All manuscripts submitted for publication in the *Bulletin* should constitute original and scholarly work that has not been published elsewhere or is not being submitted simultaneously for publication elsewhere. The work should be of relatively high professional calibre in order to meet the requirements of a United Nations technical publication. Contributors are kindly asked to exercise discretion in the content of manuscripts so as to exclude any critical judgement of a particular national or regional situation.

The preferred mode of transmission of manuscripts is the Word format. Each manuscript submitted should consist of an original hard copy and an electronic version, in Word for the text and Excel for charts and tables, in any of the six official languages of the United Nations—Arabic, Chinese, English, French, Russian or Spanish. The manuscript should be accompanied by an abstract of approximately 200 words, a complete set of references numbered in the order of their appearance in the text and a list of key words. The manuscript should be between 10 and 20 double-spaced typewritten pages, including tables, figures and references. Tables should be self-explanatory and should supplement, not duplicate, information provided in the text.

Manuscripts, together with brief curricula vitae of their authors, should be addressed to the Editor, *Bulletin on Narcotics*, United Nations Office on Drugs and Crime, Vienna International Centre, P.O. Box 500, 1400 Vienna, Austria. A transmittal letter should designate one author as correspondent and include his or her complete address, telephone number, fax number and electronic mail (e-mail) address. Unpublished manuscripts will be returned to the authors; however, the United Nations cannot be held responsible for loss.

Views expressed in signed articles published in the *Bulletin* are those of the authors and do not necessarily reflect those of the United Nations Secretariat. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat concerning the legal status of any country, territory, city or area, or its authorities, or concerning the delimitation of any frontiers or boundaries.

Reprints, purchases and subscriptions

All issues of the *Bulletin on Narcotics* (from vol. I, No. 1 (1949), to the present double issue) are available on the home page of the United Nations Office on Drugs and Crime (http://www.unodc.org/unodc/en/bulletin_on_narcotics.html).

The following special issues of the *Bulletin* are also available as United Nations publications:

1993

Policy issues relating to drug abuse and the human immunodeficiency virus (HIV) (vol. XLV, No. 1)

Drug testing in the workplace (vol. XLV, No. 2)

1994

The family and drug abuse (vol. XLVI, No. 1)

General issue on drug abuse (vol. XLVI, No. 2)

1995

Special issue on gender and drug abuse (vol. XLVII, Nos. 1 and 2)

1996

Special issue on rapid assessment of drug abuse (vol. XLVIII, Nos. 1 and 2)

1997 and 1998

Double issues on cannabis: recent developments (vol. XLIX, Nos. 1 and 2, and vol. L, Nos. 1 and 2)

1999

Occasional papers (vol. LI, Nos. 1 and 2)

2000

Economic and social costs of substance abuse (vol. LII, Nos. 1 and 2)

2001

Dynamic drug policy: understanding and controlling drug epidemics (vol. LIII, Nos. 1 and 2)

2002

The science of drug abuse epidemiology (vol. LIV, Nos. 1 and 2)

2003

The practice of drug abuse epidemiology (vol. LV, Nos. 1 and 2)

Requests for permission to reprint signed material should be addressed to the Secretary of the Publications Board, United Nations, New York, NY 10017, United States of America.

Correspondence regarding the purchase of copies of and subscriptions to the *Bulletin* should be addressed as follows:

For Asia, North America, Oceania and South America:

The Chief
Sales and Marketing Office in New York
United Nations Publications
United Nations Headquarters
New York, NY 10017
United States of America

For Africa, Europe and the Middle East:

The Chief
Sales and Marketing Office in Geneva
United Nations Publications
United Nations Office at Geneva
Palais des Nations
1211 Geneva 10
Switzerland

CONTENTS

Empirical modelling of narcotics trafficking from farm gate to street*

R. ANTHONY and A. FRIES**

Research Staff Members, Institute for Defense Analyses,
Alexandria, Virginia, United States of America

ABSTRACT

The mass distribution of illicit drugs requires large organizations and industrial-scale activity. In the present article, the worldwide cocaine distribution system is analysed and modelled with a view to identifying "business" patterns common to all levels of trafficking and comparing them with those for other drugs of mass distribution. Three models are developed and discussed: the user function (the pattern of cocaine consumption among users in the United States of America), the price-quantity scaling relationship for cocaine (the near-perfect correlation of observed prices per pure gram to transaction quantities in pure grams at each level of the cocaine trade) and a generic model for mainstream cocaine trafficker organizations that applies to all levels, from the source zone to United States streets. The user function illustrates both methodological and consumption pattern issues. The form of the price-quantity scaling relationship is explained by the influence of three constraints common to all levels of drug trafficking: lack of trust, competition and law enforcement pressure. The trafficker organization model incorporates the three constraints, the parameters of the scaling relationship and the results from the analysis of seizure rates across trafficker levels. Together, the scaling relationship and the trafficker organization models explain the observed price dynamics associating source-zone counter-cocaine operations with price swings and user responses in the United States. The three models provide insight into drug use and the complex narcotics trafficking business, point to the vulnerabilities of

---

*The views expressed in the present article are those of the authors and do not represent an official position of the Institute for Defense Analyses or of its sponsors.

**The present article synthesizes more than a decade's worth of empirical analysis and model-building by researchers at the Institute of Defense Analyses. The invaluable contributions of Arthur (Rex) Rivolo, Stephen Hanson, Samir Soneji, Larry Lyons and Barry Crane are especially acknowledged. Rex Rivolo laid the theoretical foundations underpinning the analyses of the System To Retrieve Information from Drug Evidence (STRIDE) database, the cocaine user function and the indicators of cocaine abuse. While working as a summer intern, Stephen Hanson collected extensive and essential data from a wide variety of sources in the United States and Peru that enabled the detailed analysis of the air bridge interception efforts. Also working as a summer intern, Samir Soneji conducted comprehensive time series analyses of the Peru air interceptions and other source-zone events relative to features in the United States price index. Larry Lyons provided law enforcement data and insights that supported and illuminated the analyses. Finally, during his tenure at the Institute, the remarkable leadership, drive and expertise of Barry Crane spearheaded the individual and collective research efforts.

the cocaine production and distribution system, map the limits of trafficker adaptability, offer a basis for measuring supply-control performance at each market level and create a common framework within which experts can exchange knowledge.

Keywords: Drug trafficking; cocaine; price; purity; interception; seizures; international; drug transactions; time series.

Introduction

The present article synthesizes over a decade's worth of empirical analysis and model-building by analysts of the Institute of Defense Analyses. Several methodological innovations were necessary to extract coherent patterns and create verifiable models from over a dozen major data sets on cocaine trafficking, counter-cocaine operations and indicators of damage to the illicit cocaine distribution system. The results measure the effectiveness of supply-control efforts as a complement to other counter-cocaine efforts such as education, demand reduction and treatment.

There are many reasons to construct and to validate models of narcotics usage and trafficking: *(a)* to provide a shorthand for thinking about narcotics production, trafficking and usage so that any overall patterns and the conditions that shape them may be recognized; *(b)* to find the vulnerabilities of the narcotics production and distribution system, as the basis for focusing supply-control efforts; *(c)* to understand the options and mechanisms by which narcotics businesses can adapt to threats from counter-drug activities, as well as the changing business environment, so that traffickers' reactions can be anticipated and countered; *(d)* to measure and report on the effectiveness of supply control at every level so that the benefits of counter-drug operations can be judged realistically and resources allocated effectively; and *(e)* to create a common framework within which diverse experiences and expert views can contribute coherently.

In the present article, the use of the term "modelling" entails more than an abstract exercise to test preconceptions of how the drug business works. It is guided by and relies critically on real data and the actual experiences of those working to reduce the supply of illicit drugs. While the illegal drug business is complex and barriers are intentionally created to stop people obtaining information about it, several practical constraints apply to those in the business, regardless of the level at which they operate. These restrictions are sufficient to produce observable regularities that can be used as a basis for creating three insightful and useful models: the first is the user function describing relative intensities of consumption; the second is a representation of cocaine distribution from farm gate to street users worldwide and its associated market dynamics in response to major source-zone counter-drug operations; and the third is a generic "business model" common to all levels of trafficking. Although the models used necessarily collapse a great deal of variation into caricature representations, they are nonetheless sufficient to address our objectives.

After commenting on the complexity of the narcotics trafficking business sector, the user function of cocaine abuse is described, followed by three simplifying constraints that facilitate the development of the common trafficker business model. Next, data from the System To Retrieve Information from Drug Evidence (STRIDE) of the United States Drug Enforcement Administration are employed to reveal remarkably consistent scaling relationships across five major narcotics. For cocaine, a variety of additional data are employed to extend the scaling back towards the top of the trafficker hierarchy in the source zone, that is, to the major drug trafficker organizations, and from there back down another collection hierarchy to the farm gate. Combining early Peruvian data with the sweeping market changes of the 1980s, the reason for the drop in the price of cocaine in the United States before demand abated is explained. Finally, the dynamics of the cocaine business up to the 1990s is analysed in detail, establishing that successful major counter-cocaine operations in the source zone can substantially affect the purity and price of domestic cocaine and cause significant shifts in leading indicators of cocaine abuse.

Complexities and constraints

Considering the diversity of conditions and factors influencing drug trafficking, one might wonder whether analysis and modelling can offer insights. Traffickers operate in several countries and geographical environments. Most United States counter-drug efforts have focused on cocaine, but cocaine is only one of several narcotics of mass distribution. Even for cocaine moving from farm gate to the United States streets, traffickers must perform several tasks in addition to buying and selling drugs. For example, they must reprocess or repackage the drugs, maintain security, launder money, communicate, smuggle through hostile environments and avoid both competitors and law enforcement. Many of these tasks are performed by quasi-independent contractors to the trafficker organizations who own the drugs. Furthermore, narcotics traffickers are secretive about their operations and decision processes. Thus, only limited types of information are available to guide, verify or calibrate any modelling effort. Lack of access to information also means that analysts and modellers have far fewer case studies of illicit businesses than licit businesses. This creates a risk that theories and models of licit businesses might be adopted inappropriately to analyse and model illicit drug businesses.

These complexities of business practices give rise to a number of methodological problems. One of the most challenging is the recurrence of extremely broad and fat-tailed statistical distributions of prices and other transactional characteristics. To illustrate this problem with an important related example, a comparable fat-tailed statistical distribution representing the amounts of cocaine consumed by users is explained.* Then, once again, the trafficking business is

*It will be seen that this fat-tailed distribution has a divergent integral for its theoretical variance; thus, the Central Limit Theorem, which is the theoretical basis for interpreting mean values and using regression analysis methods, fails. Instead, the distribution of the number of users by the amounts they consume appears to follow an inverse power distribution.

discussed and three important simplifying constraints imposed upon narcotics traffickers at all levels of the distribution hierarchy (lack of trust, competition and law enforcement) are considered.

The user function for cocaine abusers in the United States

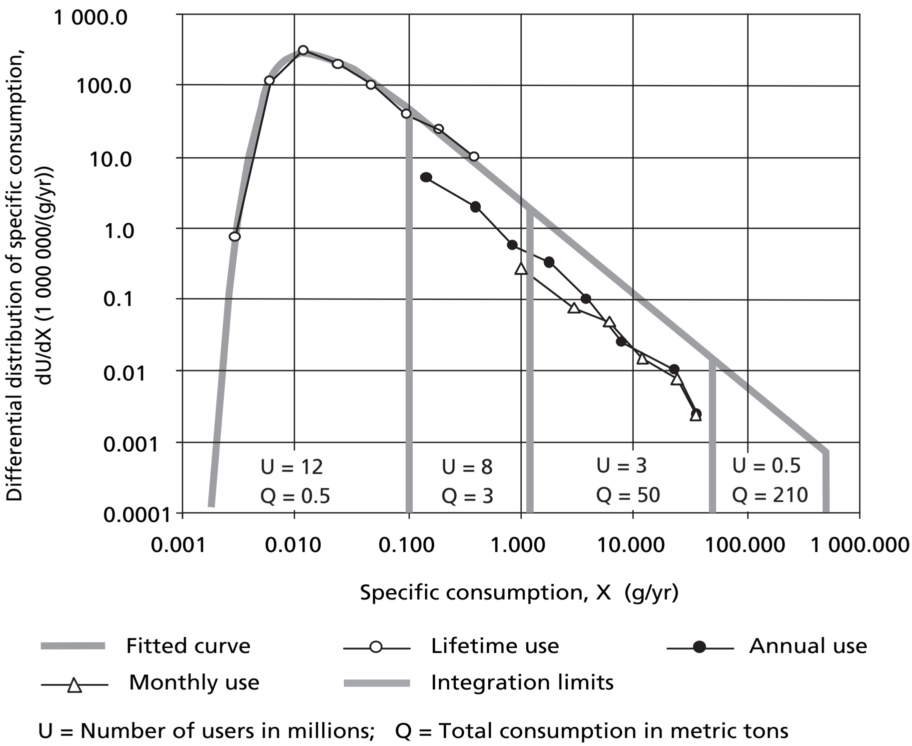
Data from the United States National Household Survey on Drug Abuse were examined to study the statistical distribution of the numbers of users consuming specific amounts of cocaine [1]. Figure I shows the differential distributions of specific consumption for users aged 12 to 35, expressed as annualized consumption in grams (g) per year, for the three categories of usage defined by the survey: lifetime, last-year and last-month use. This plot combines and renormalizes data from the 1990-1994 household surveys, at which time consumption was at roughly half the level of the 1985-1987 peak of the "cocaine epidemic" in the United States. Since the survey asks respondents the "number of times" they used a drug rather than the quantity they used, for the purpose of this research 100 milligrams was taken as a typical dose to obtain grams per year. Thus, the figures for specific consumption of less than 0.1 g per year, given in figure I, represent a lifetime frequency of less than one dose per year. The change of slope and drop in the curve below 0.01 g per year indicates the large number of people who experimented with cocaine, for example, using it once over a 10-year period.

The three empirical distributions of consumption in figure I do not match in overlap regions and this calls for explanation. For example, the lifetime specific consumption ends at a level nearly five times greater than that of the beginning of the annual reported consumption. While the twofold drop from the peak of the cocaine epidemic might explain double the number of lifetime users at a given rate of consumption, it seems implausible that the number of users consuming between 1 and 30 g a year was five times greater during the cocaine epidemic of the 1980s than in the 1990-1994 samples. A plausible explanation might be that users who quit were more willing to report their consumption than those who continued to abuse cocaine. A similar consideration might apply to the greater reported consumption during the last year compared to the last month.

Additional biases may apply to current heavy users. Those reporting the highest levels of consumption during the last year or month would have to remember over 300 doses for the annual question and 30 doses for the monthly question. At a level of abuse averaging about once a day, one might expect such users to have difficulty accurately remembering true counts or to be unavailable as a result of their addiction. Thus, in figure I, the downward slope of −1.3, which corresponds to the exponent of an inverse power-law, might exaggerate the true decline in numbers of users.

However, the discrepancy most difficult to explain is the gross inconsistency between the cumulative consumption estimated from the reported yearly or monthly distribution compared with known levels of trafficking. Total

Figure I. Cocaine distributions of specific consumption



Source: United States National Household Survey on Drug Abuse, 1990-1994.

consumption, obtained by integrating the fitted distribution weighted by amounts consumed, sums to barely 25 metric tons, less than 10 per cent of the total consumption estimated from agricultural surveys and trafficking volumes.*

To reconcile total consumption with these reported rates of consumption, the data were fitted to a power-law distribution truncated at the low end by a Fermi step function and at the high end by a sharp cut-off.** To arrive at a smooth, continuous function, the segments derived from "annual" and "monthly" consumption were scaled up or down as necessary as part of the fitting procedure. The exponent value was also allowed to float and the data were fitted by a chi-squared minimization. This fitted distribution was called the "user function". Total consumption for this fitted curve, as indicated in figure I, still would not exceed 50 metric tons if the high side cut-off to the integral were placed where the data ends. This is less than one quarter of what is believed

*All contemporaneous estimates of United States consumption during this time frame were approximately 250 metric tons and, even in 2000, a good estimate would be 210 metric tons, as found in table C.1 of annex C.

**A Fermi step function is a logistic curve with a very narrow transition from near zero to near maximum values.

to have been consumed during this period. Extending the high side cut-off to 500 g per year, a reasonable physiological limit, increases the total consumption to approximately 263 metric tons, which is a value consistent with estimated consumption from other sources. This calculation implies that the great bulk of consumption occurred in a population not sampled by the National Household Survey on Drug Abuse.

The key methodological implication of this analysis is that the total number of lifetime users (U) is almost entirely driven by the lower specific consumption cut-off of 0.01 g per year, while the total annual consumption (Q) is almost entirely driven by the upper limit cut-off of 500 g per year. Fully 80 per cent of consumption results from abusers who consume 50-500 g per year and constitute a population apparently unseen by the household survey. This population, if it exists, might be found by surveying treatment centres. Since it is difficult to conceive of individuals consuming 10 g per week for 52 consecutive weeks, more effort needs to be focused on measuring this user function with increased accuracy. Ultimately, new evidence might relax one or more of the constraints imposed upon the user function such as a reduced estimate of total annual cocaine consumption, an increased estimate for heavy user responses or a better understanding of abuse patterns by very heavy users.

The estimated user function shown in figure I shows that the "average" consumption for this period was 263/23 = 11 g per year, while the median or "typical" user consumed only about 0.1 g, a single dose, per year. Here the average is a poor indicator of usage, between 5 and 50 times below the heavy use level and 10 times the typical use level. Worse, the average is driven entirely by the upper cut-off at 500 g per year, which is the least understood portion of the overall distribution. In general, the mean value of a power-law distribution or similar long-tailed distribution is essentially determined by the upper limit and poorly characterizes the distribution or changes to the distribution. Better measures are the median and other percentiles or some set of the parameters of the distribution itself, such as the exponent and cut-off points.*

Since this user function for cocaine has the same shape (that is, it is dominated by a power-law fat tail) as the user function for the legal drugs alcohol and nicotine, it must reflect a fundamental property of human susceptibility to substance abuse across a large population. If so, deriving and comparing the user functions for several legal and illicit drugs should provide important a priori knowledge about prevalence estimates and abuse patterns. This might also compensate for survey biases when abusers are inaccessible or only partially sampled. Furthermore, by deriving and comparing the user function for several narcotics, it would be possible to place the abuse of illicit drugs into the continuum of legal prescribed uses of psychotropic substances.

---

*If one attempts to compute a variance for this user function, one obtains a divergent integral because $U^2 . U^{-1.3} = U^{0.7}$, which diverges when integrated to infinity. Thus, the Central Limit Theorem does not apply to the mean of this power-law distribution.

Three common constraints on all traffickers

Lack of trust, competition and risks from law enforcement interception commonly constrain all traffickers. Lack of trust applies to relationships within business organizations and to transactions with outsiders, including trade in narcotics and contracting for essential services. In fact, interviews with smugglers reveal that their primary concern is with informants [2, 3]. As a reaction to low levels of trust, personal familiarity and constant monitoring are essential to survival. This need for personal familiarity limits organizational size to about 50 to 80 people, with embedded cells of about 5 to 10 people for specialized high-risk tasks such as arranging transactions, providing security or managing untraceable communications. A central individual, or at most a few people, controls the organization and those in charge keep subordinates from knowing enough to become a threat. Withholding key information and forbidding collateral relationships among subordinate cells are necessary practices to suppress threats of displacement by ambitious insiders. Often, outside contractors provide cocaine-processing, smuggling, money-laundering and other services.* These considerations fragment organizations into loosely connected webs of small, specialized cells and constrain horizontal and vertical integration.**

Competition among traffickers is intense because criminals operate outside the mitigating effects of strong social sanctions and a system of justice. Individuals and organizations are under the constant threat of being supplanted by competitors. Encroachment is possible from a supplier, customer or collateral organization. The diversity of sources, buyers, specialized contractors and trafficking options seldom offers any business the leverage necessary to exercise monopoly control.***

Law enforcement acts as a direct threat by seizing narcotics, assets or investments and sometimes by arresting traffickers. It also acts indirectly by employing informants motivated by rewards, exposing a competitor, seeking revenge or bargaining for a lighter sentence. Each business activity must hide from detection or pay for protection from corrupt officials. All of these threats raise the cost of every aspect of the narcotics trafficking business.

For a multitude of trafficking enterprises, each subject to these three constraints, to form a network capable of distributing narcotics to millions of customers each month from sources in remote areas of distant countries, they must be organized into a multi-level hierarchy. Nonetheless, the lack of trust

---

*The law enforcement experience of Larry Lyons, Drug Enforcement Administration station chief in Colombia for several years, provided this and other important information about trafficker operations.

**During the cocaine epidemic phase of the late 1970s to late 1980s, profits were truly excessive and expanding trafficker operations could support hirelings to organize distribution. But afterwards, competitive pressures reduced profit margins and law enforcement focused on the largest trafficking operations. Together, these forces suppressed vertical and horizontal integration.

***Although the cocaine cartels exercised near-monopoly control of the laboratory-processing of cocaine base into hydrochloride, they were not capable of sustaining monopoly prices as the cocaine supply caught up with the demand during the 1980s. Without the extreme profits of the epidemic, belonging to a vertically and horizontally integrated organization became more of a liability than an advantage.

prevents these trafficking enterprises from merging into a vertically integrated organization. Since all three restrictions apply to all trafficking enterprises and to all levels of the hierarchy, it should be possible to model each of them as a "typical" trafficking enterprise operating with a common set of strategies. For a hierarchy of such enterprises, these common strategies should lead mathematically to a scaling relationship across the distribution levels.

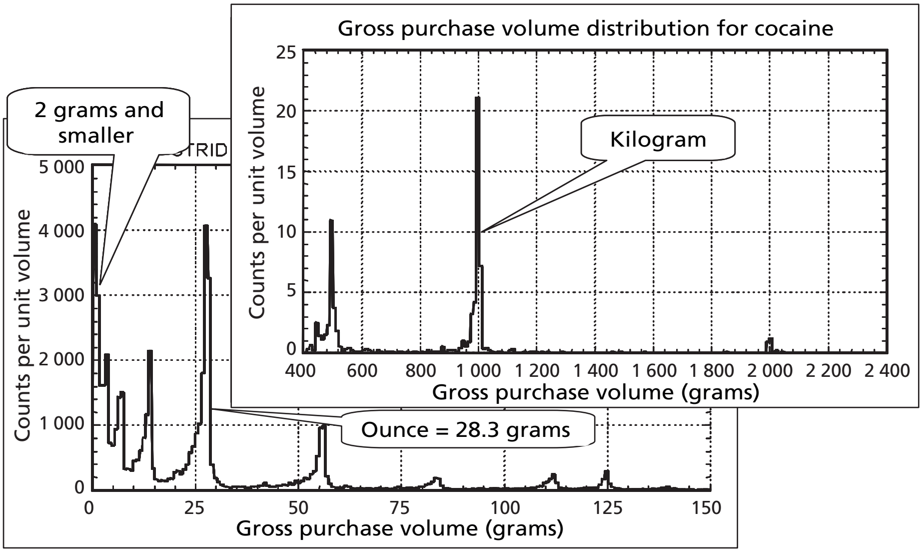
Price-quantity scaling relationship in United States transactions

Most people would expect that traffickers buying in large quantities would get big discounts; however, it is not so obvious that this discounting would lead to a price-quantity scaling relationship. Before explaining this, the data and the analytical methods used in the present research must be introduced.

The STRIDE database of the Drug Enforcement Administration provided the first and best evidence of scaling among trafficker transactions [4]. For two decades, the Drug Enforcement Administration employed undercover agents to make purchases of cocaine (as well as other illicit drugs such as heroin, methamphetamine, marijuana and Ecstasy) and forward the samples to their laboratory for analysis. As shown in figure II, most STRIDE purchases cluster tightly into a spectrum of discrete quantities (weights): kilograms (kg), ounces or grams (g) and fractions thereof. This corresponds to an approximate stratification of traffickers into wholesalers buying kilograms and selling ounces and dealers buying ounces and selling in grams. Although many transactions take place in fractions or multiples of kilograms, ounces or grams, for the purposes of the present research, the model of the distribution system used is simplified by focusing only these principal quantities: kilogram, ounce and gram. However, the model is not so precise that it excludes a wide range of quantities and prices at each step. This restores realism and demonstrates that the results of the research remain valid if generalized to actual ranges of variability. It should be noted that, because one ounce is 28.3 g, the hierarchy of trafficking branches with a ratio of about 32, since $1{,}000/28.3 \approx 35$, $28.3/1.0 \approx 28$ and, also, the square root of 1,000 is $31.6... \approx 32$.

Licit businesses offer opportunities for one level to purchase companies at other levels to create a vertically, as well as horizontally, controlled distribution process, but illicit businesses encounter many impediments to integration. While some criminal organizations might buy kilogram quantities and, for example, pay for services in gram quantities, such cases of absorbing middle-level dealing are probably rare exceptions. A wholesaler attempting to sell directly to 1,000 one-gram street customers without being revealed and arrested would, of course, be overwhelmed by the task. Alternatively, if a cocaine wholesaler attempted to force his dealers to work for him, he would be at great risk because controlling dealer cheating would also involve many more contacts than 32 sale transactions.

Figure II. Spectrum of cocaine raw purchase volumes, 1980 to mid-1996



Source: System To Retrieve Information from Drug Evidence (STRIDE) database.

Since the STRIDE data collection process reflects law enforcement priorities rather than an effort to orchestrate a rigorously balanced statistical survey of all United States markets, it presents a number of methodological problems to overcome. For example, there are variations in the number of transactions across different drugs, transaction volumes, geographical areas and seizure rates. Also, regions are not sampled in proportion to their population or level of drug trafficking activity; therefore, aggregate STRIDE prices do not represent a true national average [4]. Despite these sampling distortions, annex A establishes that such effects have only a minimal impact on the utility of STRIDE for analysing features and trends of relative prices. There are usually large price variations for a given quantity of narcotic at the same time within a neighbourhood and purchase purity at the ounce and gram levels of the distribution system varies even more than price, ranging down to zero for a complete swindle. The statistical distribution of sale prices per pure gram is so skewed that, for the smallest volumes, an average is a very misleading measure and, in principle, might not even be mathematically defined for the true distribution function [4]. Within this context, annex B illustrates that medians are much more informative than averages when characterizing very broad and long-tailed distributions.

Cocaine price-quantity relation

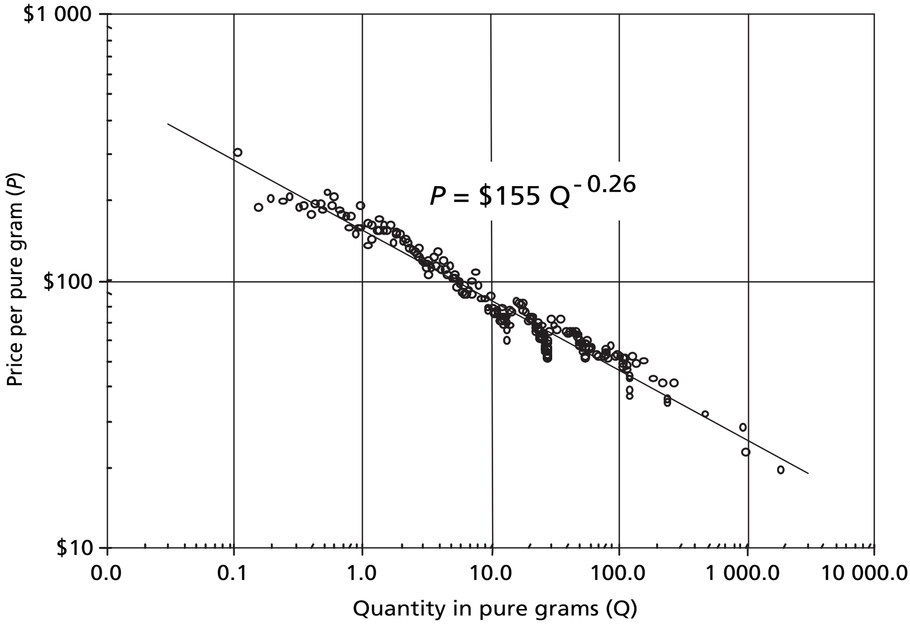
The most direct way to analyse any scaling relationship across trafficker distribution levels for cocaine is to plot the logarithm of the price of cocaine against the logarithm of the quantity of cocaine being transacted. Since multiplicative factors become constant intervals on logarithmic scales, a common price mark-up would appear as the same-sized interval on such a plot, no matter whether it applied to gram quantities or kilogram quantities. Similarly, the trade-off between risk and profit leading to typical number of customers generates a common ratio for the quantity bought divided by the quantity sold for each trafficking level. This common ratio is also represented by a fixed interval on the logarithmic scale of transaction quantities. Thus, ideally, a scaling relationship between price and quantity should appear as a simple linear trend on such a price-quantity plot. Because cocaine purity varies widely across and within transaction levels, the analytical procedure must be somewhat more complicated.

To determine the price-quantity scaling relationship, the following procedure overcomes the previously identified data shortcomings: for each transaction recorded in STRIDE, the purchase price per pure gram (referred to as the "normalized purchase price") is calculated, thereby properly including variations in purity among transactions. For the same reason, the transaction quantity is normalized to pure grams. Next, the data are ordered by increasing purchase quantity and the median of their normalized purchase prices is computed for each successive 200 or so quantities throughout the entire data set [4]. Finally, the logarithms of the respective medians of purchase quantity and normalized price for each sample point are taken. It should be noted that this procedure is also insensitive to the lower sampling rate of STRIDE for the gram- and kilogram-transaction quantities relative to ounces. Although the median prices are biased by the non-uniform sampling of regions of the United States, the sampling bias should apply to all quantities alike, therefore preserving the accuracy of the slope of the data trend, that is, the scaling relationship itself. Examination of regional differences among median prices indicates that a true median would fall within 10 per cent of the STRIDE median.

Figure III shows median normalized purchase prices for a spectrum of purchased amounts represented by their pure cocaine content. It should be noted that, in the range between 14 and 28 g, the declining diagonal trend of sample points represents the variation in price and purity for one-ounce transactions. Overall, the remarkably uniform downward linear trend of these medians when plotted as the logarithm of unit price versus logarithm of normalized quantity reveals a scaling relationship with slope −0.26.* Thus, for two different scales of transaction, Q_L for large volume and Q_S for small volume, the corresponding price ratio is $P_S/P_L = (Q_S/Q_L)^{-0.26}$. Supposing $Q_S/Q_L = 32$ (for the levels of the simplified distribution model), then $P_S/P_L = 2.5$, that is, the price mark-up is a factor of 2.5. To gain larger mark-ups while constrained by this scaling

*By 1989, cocaine prices had reached a floor that was sustained until late 1999, with the exception of several rises typically less than a doubling. Medians of 200 successive purchases ranked by pure quantity were taken. The regression line slope is −0.262 ± 0.003 and the adjusted R-squared is 0.95. A hypothesized quadratic curvature of this trend yielded a regression coefficient less than its standard error. The quality of this fit to a simple regression argues strongly that it represents an important regularity of the cocaine trade; a similar mathematical relationship was reported previously in the literature (J. Caulkins and P. Rema, "Quantity discounts and quality premia for illicit drugs", *Journal of the American Statistical Association*, vol. 88, No. 423 (1993), pp. 748-757).

Figure III. Cocaine purchase price-quantity relation, 1989-2003



Source: System To Retrieve Information from Drug Evidence (STRIDE) database.

relationship, traffickers would have to make more sales for every purchase, but more sales mean more risks in arranging those additional transactions.

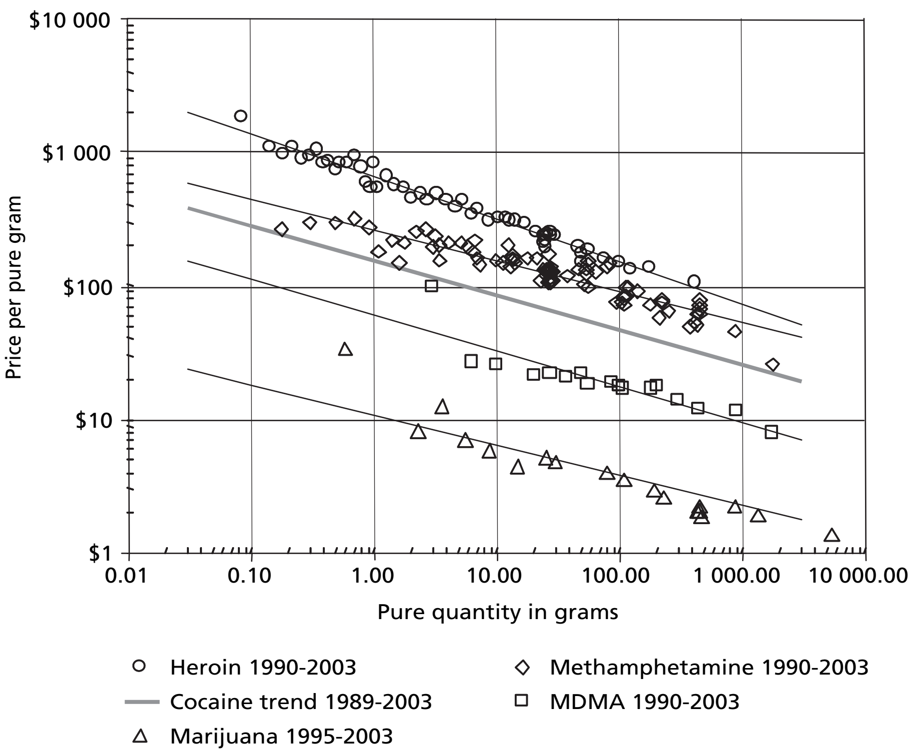
The following idealized example based on pure quantities illustrates how competition maintains this linear trend. Let us suppose that a wholesaler buying kilograms and selling ounces was able to find dealers who would buy at a unit price of $100 per pure gram rather than the scaling relationship value of $64. While the wholesaler would achieve a mark-up of 3.9, the dealers selling grams to users would find it difficult to sell at much above $155 per gram and therefore difficult to survive with a mark-up of only 1.6. Encroachment could come from any direction: from collateral wholesalers selling at the $64 unit price, from dealers setting up business as a wholesaler to compete or from a spin-off member of the distributor network moving into a lucrative niche as a wholesaler. Alternatively, if a wholesaler were forced to sell to dealers for a much lower mark-up of 1.6, that wholesaler would have little incentive to stay in business, would have trouble covering costs and would be a threat to competing wholesalers.

Thus, the three constraints lead to a simplification. Lack of trust restricts the number of transactions much beyond dealing with familiar individuals; for cocaine this is about 32 customers. Large business risks also require large price mark-ups compared with licit businesses; for cocaine this is a mark-up of 2.5 times purchase price. Finally, competition enforces this "business model" across all major traffickers at all levels. Hence, the scaling relationship expresses the equality of risk-taking per mark-up gain for each trafficking level.

Quantity discounts for other narcotics

If risks, lack of trust and competition among traffickers cause a uniform price-quantity scaling relationship for cocaine, they should also cause scaling for other illicit narcotics. This is illustrated in figure IV, which depicts the price-quantity relationships for four other major drugs: heroin, methamphetamine, marijuana's active ingredient tetrahydrocannabinol at an assumed 1 per cent concentration of purchased weight and 3,4-methylenedioxy-N-methylamphetamine, also known as Ecstasy (MDMA). The data for cocaine, which were shown in figure III, have been omitted from the trend line for cocaine to reduce clutter. Both MDMA and marijuana show an abrupt doubling or greater price mark-up for the smallest quantities sold, possibly resulting from sales in entertainment clubs. Since the Drug Enforcement Administration collects more data on heroin and cocaine and their respective markets appear more stable over time, they fit the regression trends better than the remaining three narcotics (as indicated by the adjusted R-square values in table 1). Consequently, their respective slopes are known more precisely than the others.

Figure IV. Price-quantity discount for all major United States narcotics



Source: System To Retrieve Information from Drug Evidence (STRIDE) database.

Table 1. Price-quantity relationship parameters for major United States narcotics

Narcotic	*Purchase per median*	*Exponent*	*Adjusted R-square*	*Price per pure gram*	*Customers per 2.5 mark-up*	*Revenue*
Heroin	200	–0.32	0.97	$661	17	$11 200
Methamphetamine	100	–0.23	0.84	$266	52	$13 800
Cocaine	200	–0.26	0.95	$155	33	$5 100
MDMA (Ecstasy)	50	–0.27	0.77	$61	30	$1 830
Marijuana	50	–0.23	0.74	$11	57	$627

Source: System To Retrieve Information from Drug Evidence (STRIDE) database.

Since all these drugs compete with the other drugs at one or more levels of the narcotics distribution hierarchy and also compete for heavy users in the poly-drug abuse environment, the profitability of each drug relative to its risks must be comparable. This competition among alternative criminal "investments" also implies that there should be a generic narcotics distribution network ready to distribute any new promising drug to a mass market. If competition led all traffickers to base their business models on a mark-up of 2.5 times the purchase price, then the price-quantity exponent for each drug would imply different numbers of sales transactions, as indicated in the column headed "Customers per 2.5 mark-up" in table 1. Heroin dealers would need to conduct only half as many transactions as cocaine or Ecstasy dealers, who in turn would make roughly half as many sales as methamphetamine* and marijuana dealers. With the exception of methamphetamine, this sequence corresponds to higher returns for those drugs with greater trafficking risk and to risk moderation for selling to fewer customers.

Scaling extended to the source zone

If a lack of trust, high risk and competition enforce a common scaling for all major United States narcotics businesses, these constraints might apply with equal force to the trafficking levels leading back to the source zone.

Farm gate to United States streets in the mid-1990s

Table 2 shows the transaction quantities and prices for cocaine distribution during the 1990s from farm gate to the United States streets. The source-zone data for the consolidation steps represent depressed cocaine base prices and the

*The methamphetamine market in the United States is very volatile and would require detailed analysis to explain this anomaly.

large mark-ups for those few remaining pilots willing to smuggle from Peru to Colombia following the implementation of the "force-down or shoot-down" policy in 1995 [5]. Pilots' fees rose at the expense of coca growers' profits because Peruvian cocaine base arriving in Colombia had to sell for competitive prices [5, 6]. Transaction volumes and prices for the major drug trafficker organizations are severely constrained by the juncture between the concentration and distribution steps. Transactions between Colombian major drug trafficker organizations and Mexican or Caribbean transit-zone traffickers may take place at sea or at other intermediate locations; thus, the Colombian "wholesaler" prices might be lower as a result of diversion closer to the source. It is known that about equal quantities of narcotics are smuggled by "go-fast boats" (speedboats) and fishing vessels, with air trafficking and other modes of transport carrying a minor fraction of the total. While fishing vessels might carry 4 metric tons, often several major drug trafficker organizations pool shipments to make up large loads; hence the transaction volume is many times less than the smuggler load size.

Data sources for the source-zone and transit-zone distribution steps generally do not associate a price with a quantity and it was found that serious methodological distortions resulted from attempting to assign broadly reported typical prices to similarly reported quantities.* For example, access-zone smuggling into the United States spans a very diverse range of methods and quantities. Therefore, wide uncertainty ranges were adopted for transaction quantities; table 2 gives approximate ranges for these. As mentioned in the section above on the price-quantity scaling relationship in United States transactions, the lack of systematic sampling in the accumulation of STRIDE data implies that both raw and normalized price medians for the United States as a whole might differ from actual values by 10 per cent. Thus, a generous uncertainty range was adopted for prices: plus or minus 30 per cent. These wide ranges also span the few values available from classified sources, which also support the findings of the research.

Because transaction quantities from source zone to consumer zone range from metric tons to fractions of a gram, even large uncertainties for individual transactions do not seriously compromise the main result. Four steps, each with a 2.5 mark-up, separated by five transactions, break down a metric ton of cocaine into gram quantities sold on the street; ideally, each trafficker would sell to 32 customers. Figure V depicts this adjusted version of the STRIDE price-quantity trend extrapolated from the consumer zone into the source zone. For the 1990s, the price-quantity model becomes $P = 145 \cdot Q^{-0.265}$ rather than $P = 155 \cdot Q^{-0.262}$; lower 1990 prices lead to the \$145 factor and the difference in the exponents is less than the statistical uncertainty. This trend passes through the uncertainty ranges for Mexican and Caribbean transit-zone markets, the major drug trafficker organization to transit-zone trafficker exchange and eventually connects with the size

*This problem arises even with a large quantity of data from the STRIDE database: taking medians of quantity and price individually and then combining those medians gives different results from the proper method of normalizing price and quantity for individual transactions and then taking the medians.

Table 2. Cocaine transaction quantities and prices, 1990-1999

Production/ concentration/ distribution step: activity at point of transaction	*Quantity of transaction*			*Approximate prices for each transaction (unit price per gram)*			
	Range	*Nominal raw weight*	*Pure grams*	*Transaction price (dollars)*	*Raw unit price (dollars per kg)*	*Price of pure gram (dollars)* Observed	*Price of pure gram (dollars)* Model
Leaf harvesting, selling by *arroba*[a]	A few *arroba* of leaf	11.5 kg	82	12	12 per 11.5 kg	0.15	0.18
Base[b]	500 g to 2 kg	1.0 kg	800	320	320 per kg	0.40	0.29
Transport to cocaine hydro chloride laboraory[c]	250 kg to 1.0 metric-ton	500 kg	400 000	400 000	800 per kg	1.0	1.1
Major drug trafficker organization laboratory prices are artificial[d]	6-20 metric tons	10 metric tons	8 000 000	16 000 000	1 600 per kg	2.0	2.1
Transit zone smuggling[e]	0.5-4 metric tons	1.0 metric ton	800 000	2 400 000	2 400 per kg	3.0	3.9
Access zone smuggler or distributor[f]	10-100 kg	30 kg	24 000	210 000	7 000 per kg	8.8	10
Wholesaler[g]	500 g-2 kg	1.0 kg	800	20 000	20 000 per kg	25	25
Dealer[h]	05.-2 oz	29 g	20	1 360	47 per g	67	65
Consumers or consumer/dealers[i]	0.07-5 g	1.0 g	0.65	110	110 per g	169	163
One hit of crack cocaine[j]	100 mg	0.1 g	0.07	18	180 per g	260	290

Note: 1 *arroba* = 12.5 kilograms.

[a]Although there is a very limited market for leaf even in Peru, a few *arrobas* from small producers may complete a batch or add profit when base prices are up (leaf is about 0.7 per cent cocaine).

[b]Farmers process leaf and produce from 1 to a few kilograms of base of about 80 per cent purity. Prices are for the period after the force-down or shoot-down policy and before the 1999 rise.

[c]Without interception, $20 000 per load; for non-lethal risk, $90 000; and with lethal interception pilot fees rise to $200 000 [5].

[d]Major growing regions produce enough leaf for about 20 metric tons per quarterly harvest, but some small producers do not transact that amount. At the time, laboratories produced about 80 per cent pure cocaine hydrochloride.

[e]Planes and go-fast boats (speedboats) carry about 500 kg to 1 metric ton, fishing vessels about 4 metric tons and container freight that much or more.

[f]These volumes are about the typical size of cross-border smuggling, which ranges widely. They also characterize mid-level distributors' loads in the United States.

[g]Local wholesalers typically buy 1 kg quantities, grind up the bricks, dilute the powder about 10 per cent and resell in ounce quantities.

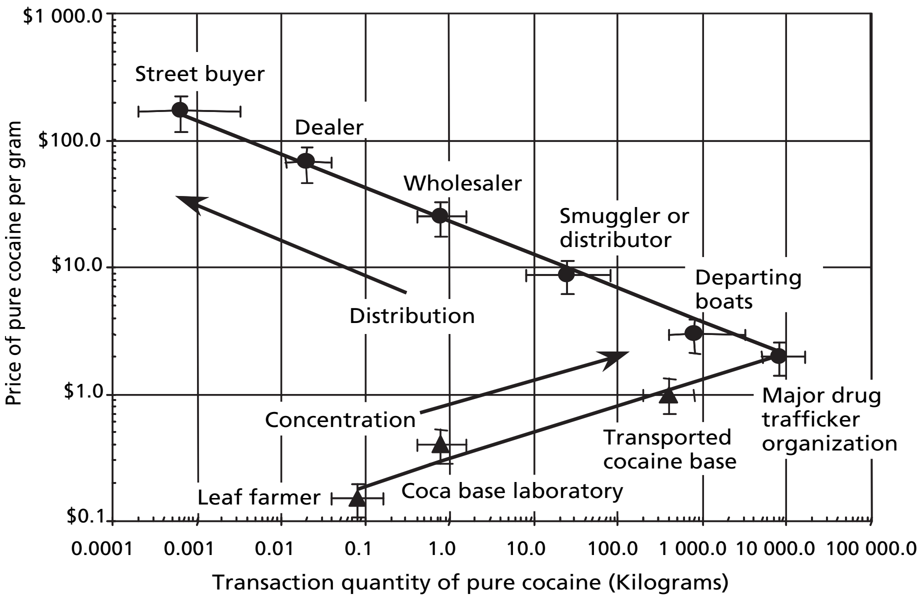
[h]Dealers purchase about an ounce, dilute it about 10 per cent and resell it in gram or partial gram quantities. Crack dealers cook the cocaine into hard, nearly pure crystals and resell in rocks or vials.

[i]Users and buyers range widely in consumption and transaction amounts. The median user may consume a gram per year, but most consumption is by heavy users absorbing over a gram per week. Some of these are user-dealers, who pay for their heavy use by selling to others.

[j]Crack cocaine is sometimes sold as a single "hit" to those with little money.

of the largest cocaine hydrochloride laboratory complex ever found. The Colombian National Police raided this laboratory complex in December of 1996 and January 1997. Judging by the equipment and chemicals found on site, it had an estimated capacity to process 100 metric tons of cocaine per month, although it did not operate continuously. A total of 7 metric tons of cocaine hydrochloride were seized, along with 1.5 metric tons of cocaine base. The laboratory complex had access to several growing regions that together produced about one third of the cocaine destined for the United States. The processing batches, or "transaction sizes", were probably comparable to the production of an entire growing region, that is, equivalent to 5-16 metric tons of pure cocaine.

Figure V. Late-1990s price-quantity trends for consolidation and distribution segments



Since the 1996-1997 raids, cocaine hydrochloride laboratories have become smaller and more dispersed and most laboratory owners contract their services to major drug trafficker organizations, which own the cocaine [7]. This spreads the risk both for the major drug trafficker organization and the laboratory owner, who most often also has several laboratories. Let us estimate a major drug trafficker organization's risk based on their number of transactions. Let us suppose that the organization buys 20 cocaine base loads of 500 kg each, processes it in five laboratories and sells metric ton loads to 10 transit-zone buyers. This represents 20 + 10 + 5 = 35 transactions, comparable to the standard business model "risk" of 32 transactions.

Ascertaining the number of major drug trafficker organizations and their potential impact on the highest levels of the cocaine business is important to law enforcement planning. Since the price-quantity trends constrain the typical major drug trafficker organization to arrange for the processing of 5-16 metric tons of pure cocaine at each harvest, four or five times a year, it is possible to estimate their number. In 1999, Colombian cocaine hydrochloride laboratories could have produced as much as 560 metric tons of pure cocaine, 520 from Colombia itself and another 40 smuggled in from Peru. At one extreme, each major drug trafficker organization transacts 80 metric tons a year; thus, there would be 560/80 = 7 such organizations. At the opposite extreme, each transacts only 20 metric tons a year; hence there would be 28 such organizations. In reality, there must be a distribution of major drug trafficker organization sizes, with the largest possibly controlling about a third of the market.

Figure V also shows the price-quantity relationship, beginning with Peruvian coca growers on up to Colombian cocaine hydrochloride laboratories. These steps involve physical processing costs in addition to price mark-ups of about 2.5. However, traffickers collecting cocaine base from farmers in 1 or 2 kg quantities, consolidating it into 500 kg loads, smuggling it to Colombia and selling it to major drug trafficker organizations represents a concentration ratio of 250 to 500:1. These traffickers can deal with hundreds of farmers in contrast to the typical 32 customers of distributors because there is very little risk of interception or theft in the remote growing areas. Nonetheless, their mark-up remains 2.5.

Rigid adherence to the price-quantity scaling relationship

The price-quantity relationship represents the quantity discount rate for the typical trafficker's business models. Two unusual situations illustrate how rigidly the three constraints enforce this business model. First is the 1995 interception of the vessel *Natalie I* coming into Los Angeles carrying 20 metric tons of cocaine, fully 100 times more than a typical "large" 200-kg load. Even had it arrived safely, it would not have been a wise business strategy for the traffickers: either they would have had to store the bulk while parcelling it out to a limited number of wholesalers, or they would have had to sell to many more than 32 customers. Both options entail extremely high risks relative to the option of operating within the bounds of the price-quantity relationship.

Second is the 2001-2002 spate of smuggling 10 or more metric ton loads from Colombia to Mexico through Eastern Pacific routes. When the United States Coast Guard resumed interception operations following the stand-down from the terrorist attacks of 11 September 2001, they intercepted several such loads on large fishing vessels.* These load sizes were almost 10 times larger than those on the price-quantity trend. Upon closer investigation, however, police learned that these large loads were actually owned by several major drug trafficker organizations in Colombia and were being delivered to several traffickers in Mexico. This finding moves the actual transaction sizes onto the trend line. Yet it is still

*A private communication emanating from the United States Joint Interagency Task Force in 2002 recounted these events.

of concern that traffickers in both Colombia and Mexico were willing to deal with a common smuggler, apparently discounting the increased risk of betrayal from this single contact. Certainly, the smuggler would have to have detailed knowledge of all those contributing to or receiving from the load. Increased international cooperation and more effective law enforcement soon brought an end to these bulk shipments.

Cocaine price mark-ups worldwide

The United Nations listed cocaine price and purity at the wholesale (kilogram) and retail (gram) levels for countries reporting a cocaine trafficking or abuse problem in 2000 [6]. These data enabled an analysis to be carried out of the mark-up in price as cocaine moves from the Andean source-zone nations through the several transit-zone countries and finally to the four major consumer zones. These mark-ups reveal consistent patterns worldwide confirming the belief that three constraints apply across geographical barriers, ethnic differences and various law enforcement systems and systematically increase prices along each route. Since the distribution chain to the United States has a trafficker level for each 2.5 price mark-up, it is hypothesized that this is true for distribution chains worldwide. If so, the number of middlemen in these distribution chains can be estimated. In principle, this could be corroborated by law enforcement observation; however, only limited information is available for such validation.

Table 3 lists the cocaine prices and purities at both wholesale and user levels for the countries reporting to the United Nations and these are grouped into source-zone, transit-zone and consumer-zone regions for various distribution avenues. Since many countries do not report all of the data elements and some regions consist of several countries, notes below the table explain the regional aggregations and the assumptions necessary to restore missing data. To estimate the price mark-ups consistently, the prices per pure quantity are calculated in two additional columns.*

Although the United Nations reported $1,900 per kg for wholesale cocaine hydrochloride in Colombia, this price probably applies to cocaine in a major city or on its way to a coast for smuggling abroad.** For the purposes of the research, it was desired that the source-zone step would begin as cocaine base was sold for processing to cocaine hydrochloride, yet would include the processing costs. To estimate this price, $200 for laboratory processing was added to the $938 price of cocaine base to obtain $1,100 for the baseline source price. At 80 per cent purity, this implies $1,375 per pure kg. For smuggling to other South

*While several countries reported 80 per cent or greater purity in 2000, the United States Drug Enforcement Administration found kilogram bricks to have a much lower, 70 per cent median purity. Nonetheless, the United Nations reported values have been used in this analysis.

**While in table 2, $2,400 per kilogram for cocaine transactions with transit-zone traffickers was reported, this was 80 per cent pure in the 1990s, versus the 70 per cent purity in 2000. Furthermore, it is believed that the $1,900 represents a discounted wholesale price for diversion well before the transaction point with transit-zone traffickers.

American consumer-zone countries, Bolivian prices are appropriate. Even Peruvian cocaine must transit Bolivia on the way to most of these markets, although some Peruvian high-grade cocaine base goes to Colombia and a very small fraction is converted to cocaine hydrochloride for smuggling off the Pacific Coast directly to consumer-zone markets.

Table 3. Cocaine prices and purities at the wholesale and user levels

	Wholesale level			*User level*		
Country or region	*Dollars per kilogram*	*Purity*	*Dollars per pure kilogram*	*Dollars per gram*	*Purity (percentage)*	*Dollars per pure gram*
Colombia source	1 100	80	1 375	. .	. .	. .
Colombian cities	1 900	80	2 375	2.50	60	4.17
Peru[a]	. .	. .	. .	1.10	73	1.51
Bolivia	1 000	100	1 000	0.60	60	1.00
Central America	7 000	80	8 750	10	80	12.50
Mexico[b]	7 000	80	8 750	. .	. .	. .
Venezuela	4 100	80	5 125	5.00	60	8.33
Ecuador	2 000	96	2 083	1.00	70	1.43
Caribbean[c]	8 500	80	10 625	10	60	16.67
Brazil	2 500	90	2 778	4.90	60	8.17
Chile	6 000	95	6 316	2.40	60	4.00
Argentina	2 100	90	2 333	2.60	60	4.33
Africa[d]	40 000	80	50 000	12	10	120.00
W. Europe[e]	40 000	75	53 333	100	56	180.00
Canada	32 000	80	40 000	89	75	118.67
United States[f]	22 300	80	27 875	67	63	106.35

Sources: *Global Illicit Drug Trends 2001*, United Nations Office for Drug Control and Crime Prevention, 2001. Shaded numbers are from sources other than the United Nations, while missing purities were assigned consistent values.

Notes: Two dots (. .) indicate that data are not available or are not separately reported.

[a]Peru did not report a supply price; however, most Peruvian trafficking passes through Bolivia and since Peruvian quality is as good as Colombian quality, prices comparable to those in Colombia are expected.

[b]Because Mexico is not included in the United Nations tables, the wholesale price was obtained through private communications with the United States Embassy.

[c]The Caribbean is represented by the islands listed by the United Nations as having the most trafficking: the Bahamas, the Dominican Republic and Saint Lucia.

[d]Africa is represented by Nigeria because it has by far the largest user population. Ten per cent purity was assumed to achieve some mark-up, since Nigeria's users pay a very low user price relative to other African countries.

[e]European countries that are points of entry for cocaine, such as the Netherlands, have prices per kilogram in the range of $23,000 to $30,000. But these prices might be for multiple kilogram quantities and not typical of Europe as a whole. The destination countries were therefore taken as typical of European prices.

[f]The United Nations reports a United States wholesale purity of 55 per cent, which must be a typographical error.

The number of distribution steps from source to in-country wholesale level were estimated by computing the fractional number of 2.5 mark-ups, N_S, for a total mark-up, M, using the relationship $M = 2.5^{Ns}$ also written $N_s = \log(M) / \log(2.5)$. Finally, the computed steps, N_s, are rounded off in the "Transit" column of table 4.*

Table 4. Cocaine price mark-up and number of steps for transit and internal flows

Country or region	Transit market		Internal market		Steps rounded off		
	Mark-up	Steps	Mark-up	Steps	Transit	Internal	Total
Colombia	1.73	0.6	1.75	0.6	1	1	2
Peru	. .	. .	. .	. .	0	. .	. .
Bolivia	. .	. .	1.00	0.0	0	1	1
Central America	6.36	2.0	1.43	0.4	2	1	3
Mexico	6.36	2.0	. .	. .	2	1	3
Venezuela	3.73	1.4	1.63	0.5	1	1	2
Ecuador	1.82	0.7	0.69	–0.4	1	1	2
Caribbean	7.73	2.2	1.57	0.5	2	1	3
Brazil	2.50	1.0	2.94	1.2	1	2	3
Chile	6.00	2.0	0.63	–0.5	2	1	3
Argentina	2.10	0.8	1.86	0.7	1	1	2
Africa	36.36	3.9	2.40	1.0	4	1	5
Western Europe	36.36	3.9	3.38	1.3	4	2	6
Canada	29.09	3.7	2.97	1.2	4	2	6
United States	20.27	3.3	3.82	1.5	3	2	5

The mark-up ratios for the internal markets given in table 4 reveal several inconsistencies. Since the United Nations data are complete for Bolivia and Ecuador, simple arithmetic leads to the apparent conclusions that wholesale to retail trafficking in Bolivia is at cost (that is, for no profit) and similarly that trafficking in Ecuador operates at a loss (that is, less than a 1.0 mark-up). For the United States, the reported United Nations user price per pure gram is only \$106, but the lowest found for that quantity from STRIDE was \$134. The mark-up to \$106 is only 3.82, which is just 1.5 steps, rather than the 5.8 for 2.0 steps obtained from the previous, more precise, scaling relationship analysis. In general, the "internal market" steps are all smaller than expected. One possible reason for these discrepancies is that reporting countries have not learned to

*The estimated numbers of steps, excluding the artificial estimate for Colombia, tend to cluster about integral values. A standard chi-square test of the null hypothesis of uniformly spread first decimal digits yields a p-value of 0.08. A more specialized likelihood ratio test comparing this null hypothesis against an alternative of a geometrical decline away from integer values gives a p-value of 0.06.

deal with the large price and purity spreads typical of user transactions. Therefore, a minimum of one step is assumed from supply to street for source-zone and transit-zone countries and two steps for Brazil, Europe, Canada and the United States. All of this is equivalent to adding about 0.5 to all the internal market "step" estimates. Overall, with these adjustments, the total number of steps from the simplified model agree remarkably well with the presumed relative law enforcement pressures on cocaine trafficking in each of these countries and regions.

Using the United States as a template, the analysis yields a total number of trafficking steps from source to users for each country or region. It suggests that there is an extra middleman leading to European and Canadian users. For other countries and regions, the numbers of steps are plausible, given what is known about law enforcement and drug abuse. Therefore, the analysis provides a plausible profile of worldwide cocaine trafficking to compare with observations made by law enforcement officers.

Seizures and trafficker risks

For traffickers, seizures represent both a business risk, that of losing revenues, and some personal risk of arrest. Analyses at the United Nations discovered that seizures of cocaine and heroin increase in proportion to the total volume of trafficking in these narcotics [8]. There is little if any evidence that seizures alone inhibit trafficking, as might be expected. Therefore, the detailed United Nations tabulation of seizures for 2000 were analysed to determine whether they simply provide an indicator of risk; that is, whether seizure rates, like price mark-ups, are comparable for each step in the distribution system. Because experienced smugglers believe they face about a 1:10 to 1:20 chance of being intercepted, it seems plausible that experienced traffickers also anticipate seizure losses and build these into their business models' expenses to be covered by the high profits [3, 5].

Before the seizure rate can be derived, total production must be balanced against seizures and consumption, then the total traffic in the source, transit and consumer zones can be estimated. A small adjustment for the reduced purity of seized cocaine at the consumption level improves model realism. With United Nations prevalence data by country, local consumption could be derived from population statistics if the amount consumed by the "average" abuser was known [6]. Since it has already been explained that total consumption is driven by the number of heavy users, that heavy use is limited by human physiology and that heavy use is proportional to total usage according to the user function, it follows that total consumption should be proportional to the number of users. Annex C employs this condition to derive the consumption, seizures and flows for each trafficking zone. Although this analysis depends upon a very simplified flow analysis and there are many well-recognized problems with estimating total flows, it will also be verified that the results are quite insensitive to total flow.

Seizures as a fraction of flows

Table 5 presents a simplified summary of flows from source zone to consumer zone, in which each group of rows gives information on seizures, the seizure rate, consumption and exported quantities. The table includes a group of rows for the redistribution, balancing source-zone contributions with flows leading to major consumer zones.* Along with consumption, the consumer-zone group of rows includes "adjusted seizures" expressing the reduced purity at ounce and gram levels. Finally, seizure rates are calculated from seizures divided by incoming flow before consumption.

Table 5. Balance of cocaine flows to major consumer zones showing seizure rates

Flow or rate	*Unit*	*Colombia*	*Peru*		*Bolivia*	*Total*
Source zone						
Production	Metric tons	520	175	. .	70	765
Seizures	Metric tons	64	11	. .	8	83
Seizure rate	Percentage	12	6	. .	11	. .
Consumption	Metric tons	13	7	. .	1	21
Exported	Metric tons	443	157	. .	61	661
Redistribution of flows	*Unit*	*United States and Canada*	*Europe*	*Africa*	*Other South America*	*Total*
Colombia	Metric tons	–22	22	. .	. .	0
Peru	Metric tons	. .	–59	43	16	0
Bolivia	Metric tons	. .	. .	. .	. .	0
Net exports to consumer zone	Metric tons	421	120	43	77	661
Transit zone						
Seizures	Metric tons	64	17	6	. .	86
Seizure rate	Percentage	14	14	14	. .	. .
Consumption	Metric tons	29	8	3	. .	40
Exported	Metric tons	328	95	34	77	534
Consumer zone						
Seizures	Metric tons	134	44	1	12	191
Adjusted seizures	Metric tons	112	39	1	12	163
Seizure rate	Percentage	34	40	1	16	. .
Idealized rate	Percentage	32	40	. .	12	. .
Consumption	Metric tons	216	57	34	65	371

Note: Two dots (. .) indicate that cells have been intentionally left blank because they have no defined content.

*For simplicity, Asia and Australia were left out as consumer zones. This drops about 7 metric tons from total consumption, which would adjust other consumption downward by about 1.5 per cent.

In the source and transit zones, seizure rates range from 11 per cent for Bolivia to 14 per cent for the transit zone. The sole exception is 6 per cent for Peru, which exports a large fraction of production as cocaine base rather than cocaine hydrochloride and therefore has less interception exposure. Since there are some seizures of the small portion of consumed cocaine, the transit-zone countries would have a 14 per cent seizure rate if the underlying baseline rate per step were 12 per cent.

In the consumer zone, seizure rates rise significantly: 34 per cent for the United States and Canada, 40 per cent in Europe and 16 per cent in countries in South America other than Bolivia, Colombia and Peru. Only Africa has a low rate of 0.5 per cent, suggesting there is little law enforcement effort to intercept cocaine once it arrives on the continent. The consumer-zone seizure rate for countries in South America other than Bolivia, Colombia and Peru in aggregate indicates a relatively low level of law enforcement comparable to transit-zone countries.

To understand the seizure rates in the United States and Europe requires some modelling. Table 4 indicates that for Europe there are two additional steps to arrive at the wholesale level and two more to reach users. The idealized seizure rate for Europe based on 12 per cent loss per step is calculated as $1 - (1 - 0.12)^4 = 0.40$. That is, four interception steps with a 12 per cent compounded loss at each result in a 40 per cent loss, rather than the simple linear approximation of 48 per cent. Because the United States as a consumer zone includes three additional steps, one to get in and two from wholesaler to user, a similar model yields a 32 per cent aggregate seizure rate.*

It has been assumed that the initial total flow of 80 per cent pure cocaine is approximately that given by the United States estimate of Colombian production of 520 metric tons. However, the joint United Nations and Colombian estimate is 680 metric tons. The simple scaling result was re-examined to determine whether it was undermined if total flows were 160 metric tons larger. It was also assumed that the United States and Canada, as well as Europe, effectively seized less cocaine of 80 per cent purity because a significant portion of seizures would be of a lower purity, street-grade product. Again, the scaling result was checked to determine whether this assumption might have undermined it. The seizure rates for all four combinations of the two assumptions were recalculated and it was determined that the simple scaling works in all cases. The estimated magnitude for the common seizure rate does vary, however, from case to case, ranging between 10 and 15 per cent, compared to 12 per cent for the baseline case.

Connection between seizure rates and price mark-ups

Although this analysis of price mark-ups and seizure rates worldwide is highly stylized, the model appears to be robust with regard to alternative sets of assumptions and it reveals several important findings about the cocaine traffick-

*Although Canada has an additional step, it has only 2.5 per cent of the United States flow. Canada cannot be analysed independently since an unknown fraction of Canadian consumption flows through the United States.

ing industry. Comparable seizure rates for comparable price mark-ups all along the international distribution chains for cocaine is strong evidence that both are built into traffickers' business models. If seizure rates represent trafficker risks, then higher seizure rates should cause either a higher price mark-up or the insertion of additional trafficker levels in the distribution hierarchy. Furthermore, increased seizure rates at all levels should have a compound effect on access to illicit drugs, first by reducing the supply itself and second by greatly increasing the price to users. If, however, counter-drug interception operations are to have a strategic impact rather than simply continuing to exact a tax, either significantly more than 15 per cent must be seized or traffickers must be arrested to obtain a deterrent effect [5, 9].

The findings for cocaine suggest that analysts should apply this model to the distribution systems for other major drugs of abuse. By so doing, different counter-drug organizations and countries might recognize their individual roles in the context of reducing the trafficking worldwide and see where mutual efforts might have strategic consequences. For example, countries with large traffic in several narcotics might be critical hubs enabling drug substitution or the introduction of new drugs that blunt the impact of periodic counter-drug successes against an individual narcotic.

Market dynamics in response to changing conditions

Up to this point, a static model of the cocaine distribution system has been presented, from farm gate to consumer-zone users, in which price mark-ups and seizure losses reveal distinct levels of trafficking. It must now be asked how this hierarchical distribution system adjusts to changing conditions: for example, how major events have affected business levels, how major source-zone counter-cocaine operations affect prices throughout the distribution hierarchy and how the quantity-discount relationship adjusts to market shocks.

For purposes of tracking substantive United States cocaine price trends over time, a simple two-step methodological procedure can overcome the inherent variability of the STRIDE data: *(a)* for each transaction, the purchase price is calculated per pure gram and *(b)* for a sequence of 100 or so of these nearly time-coincident normalized prices, their median is computed [4]. These medians are then called the "United States price index" and these data are often plotted along with their nine-point moving average.* In this way, a sample measure, the median, is obtained, which avoids the divergence of statistical averages or moments and includes information about cheated customers, who pay infinite unit price. Further, as annex B explains, the median provides a stable statistical measure, which also has finite variance amenable to traditional statistical methods such as informative data plotting, regression analyses, time

*The price given by the United States price index is artificial in that it corresponds roughly to the median of STRIDE sampled quantities. Since the distribution of purchase quantities in STRIDE remained stable over the period of analysis, the United States price index reveals relative shifts of normalized price with a high degree of accuracy.

series procedures, and so forth. Most importantly, data features exhibited by appropriately aggregated STRIDE statistics are remarkably robust: they are insensitive to data perturbations, internally consistent across natural partitions of the data, strongly correlated with external data sets portraying market conditions and logically compatible with the incidence and timing of specific counter-cocaine operations [4, 5, 10-15].

The end of the cocaine epidemic and business-level adjustments in Peru

Cocaine abuse began in the United States as an epidemic wave in the late 1970s, crested in the mid-1980s, receded rapidly in the late 1980s and ebbed gradually thereafter. Initially, supply was short, prices surged, purity was low, coca cultivation exploded in the Andean producer nations and vertically integrated cartel organizations expanded through ethnic connections between Colombia and the United States. While Bolivian cultivation expanded somewhat early on, its greater distance from Colombia put Bolivian farmers at a competitive disadvantage relative to those in Peru. By the early 1990s, most Bolivian cocaine was supplying South American, European or other non-United States markets. Figure VI portrays the entire epidemic cycle of United States user demand from pre-epidemic 1970 to 2000. It also shows the falling United States price index from 1980 onwards and the rising and shifting cultivation in the Andean countries [4, 5].

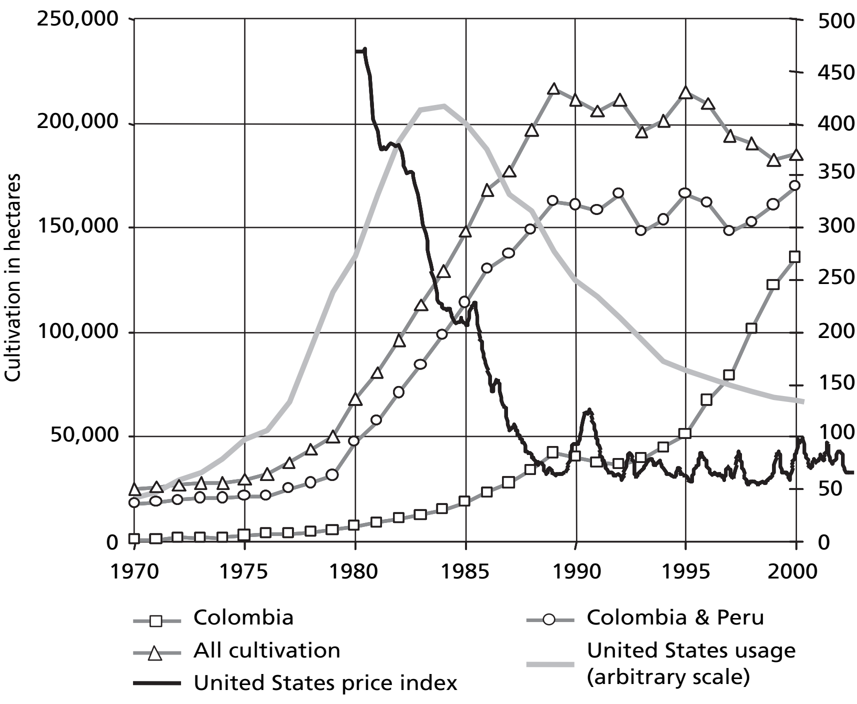
From a United States perspective alone, it is strange that cocaine prices plummeted from 1980 to 1983, while the United States household surveys indicated demand was still rising [1]. The Medellín cartel leaders surrendered in 1991, long after prices had fallen and the Cali cartel leaders surrendered three months after Peruvian cocaine base prices had collapsed in 1995 as a result of operations to sever the Peru-to-Colombia trafficking air bridge [5]. Thus, the fall of the cartels did not cause price drops.

Comparing the United States price drop with the rise in Peruvian cultivation, however, reveals a remarkable congruence. Figure VI shows the cumulative Andean cocaine production, with the stacking of contributions from Colombia, Peru and Bolivia respectively [16, 17]. The Peruvian cultivation explosion began in 1979, which was the year prices began to fall in the United States. Prices continued to fall at a rapid rate until 1989, which was the year cultivation expansion abruptly stopped in Peru and throughout the Andean region. After 1989, United States prices bottomed out at a floor that continued to hold throughout the 1990s, with only episodic surges that never approached pre-1985 levels. Demand, on the other hand, continued a smooth, uninterrupted decline through the 1989 price transition.

If the cocaine supply from Peru caught up with United States demand, leading to a price index drop of greater than a factor of six, then there should also be some evidence of this in Peru. There is indeed such evidence, as the following story explains: epidemic demand for cocaine in the United States caused

Peruvian coca leaf prices to jump from \$0.60 to \$3.00 per kilogram, six times higher than their traditionally established price baseline [16, 18]. This induced heavy migration into remote and lawless coca cultivation areas and a rapid expansion of the crop beyond the pre-epidemic 17,000 hectares [17]. Farmers with about 200 kg of dried leaf worked with cocaine paste processors, thereby reducing its mass by 200:1. Paste collectors either sold it for export to Colombia or brought it to cocaine base processors, who would make 1 kg of cocaine base from 2.5 kg of paste. Cocaine base collectors either sold that base for export to Colombia or, occasionally, brought it to Peruvian cocaine hydrochloride processors. Peruvian cocaine hydrochloride was sold for local consumption or trans-shipped to South American or European markets.

Figure VI. Cocaine epidemic and afterwards, 1970-2000



By the end of the 1980s, cocaine profits had dropped dramatically and Peruvian cultivation had stabilized at around 120,000 hectares. Most coca growers were now producing their own cocaine base to sell directly to trafficking organizations for export to Colombia. Some leaf was sold on a spot market, but this was by now a peripheral aspect of the business, since there was no longer a profit in making cocaine base from purchased leaf. Cocaine paste had virtually disappeared from the marketplace, as paste now sold for a price equal to that of the farmer's cost of growing the leaf to make it.

Figure VII shows the price series for all levels of the Peruvian cocaine business in the 1980s: dried coca leaf, cocaine paste, cocaine base and cocaine hydrochloride [4, 16, 18-21]. All prices are for a one-kilogram equivalent of cocaine. Since figure VII depicts prices on a logarithmic scale, the typical trafficking mark-up of 2.5 is a constant interval for all price ranges, as can be seen from the broad vertical bars at the left. The large crosses indicate two middlemen being forced to drop out by the late 1980s. Along the top of the figure is the United States price index developed during the research, scaled up to kilogram quantities. Immediately, the parallel decline in United States street prices with Peruvian source-zone cocaine base and hydrochloride prices becomes visible and both declines include a plateau feature between 1983 and 1985.* Since United States prices are more than 100 times those in Peru, this lock-step decline must result from a multiplicative market in which source-zone price changes are amplified as comparable percentage changes in the distant consumer zone.

Along the bottom of figure VII are two artificially constructed price series for leaf. The lowest is a flat baseline representing the cost of 500 kg of pre-cocaine-epidemic leaf. The next up from the baseline includes the investment cost necessary to expand cultivation at the rates observed during the 1980s.** All mark-ups would be measured relative to these baseline costs.

The number of researchers employing different data sources and analytical methods over different periods, coupled with the variety of trafficker practices, complicate the interpretation of the many price series depicted in figure VII [15]. Nonetheless, it can be seen that most series drop by a factor of two between 1982 and 1983 and by a factor of three between 1985 and 1989. By the end of the decade, these price drops had squeezed out the independent cocaine paste or base producers as well as the paste collectors. Instead, growers had been forced to learn how to make their own cocaine base. It should be noted that, in 1989, the United Nations began systematically collecting cocaine base prices from a wide range of growing areas, thereby causing a sharp discontinuity in reported price relative to previous series. Nonetheless, relative changes in the previously reported series should reasonably correspond to real changes in the Peruvian markets.

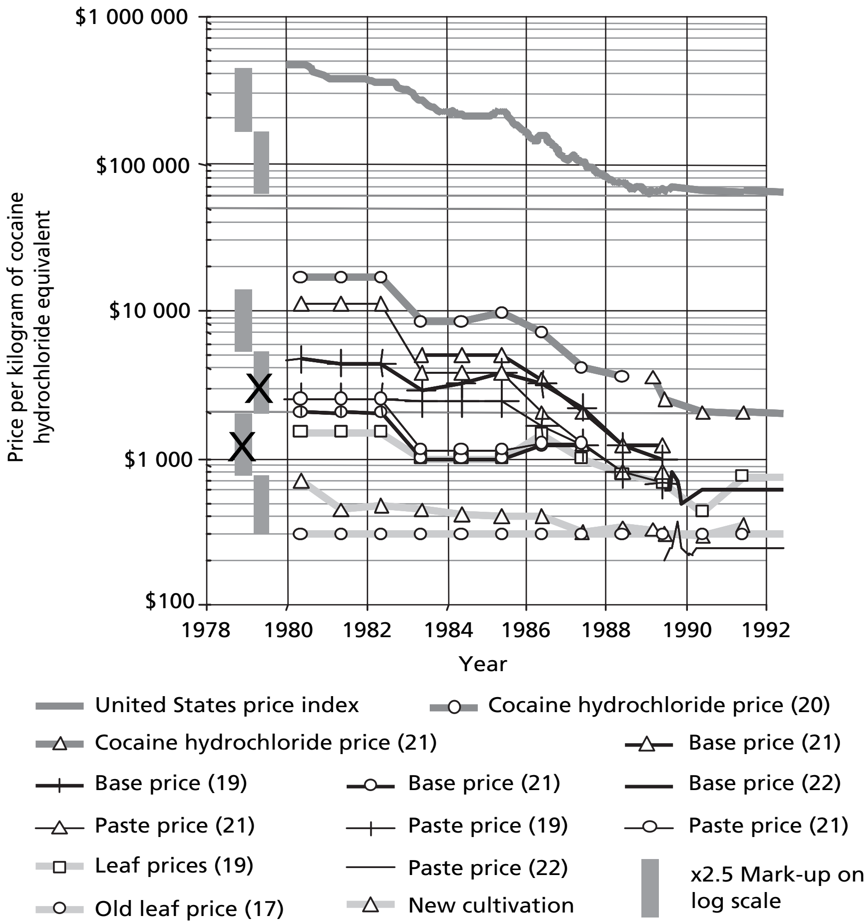
These results reinforce and extend the hierarchical distribution model. A reduction in price of a factor of six caused a structural change, squeezing out two trafficking levels, each with the standard mark-up of 2.5, that is, $2.5^2 = 6.25$. Leading to street users, a five-step mark-up amplified price changes in the source zone by 100, with virtually mechanical precision throughout the transition. And the source-zone traffickers, rather than others further up the distribution chain, lost their market niches, suggesting that concentration steps are the most economically vulnerable.

---

*This plateau was probably the result of Operation CHEMCON in Colombia, in which chemical drums were bugged with radios, leading to several major cocaine hydrochloride laboratory raids [4].

**During the 1980s, it cost $2,587 to plant a hectare of coca that, two years later, would produce 1,790 kg of leaf and it took 500 kg of leaf to make 1 kg of cocaine base. With these data and the relative rate of increase in production compared to the producing mature crop, the additional investment per kilogram of cocaine was computed. For example, this was $404 in 1980, dropping to $100 in 1985.

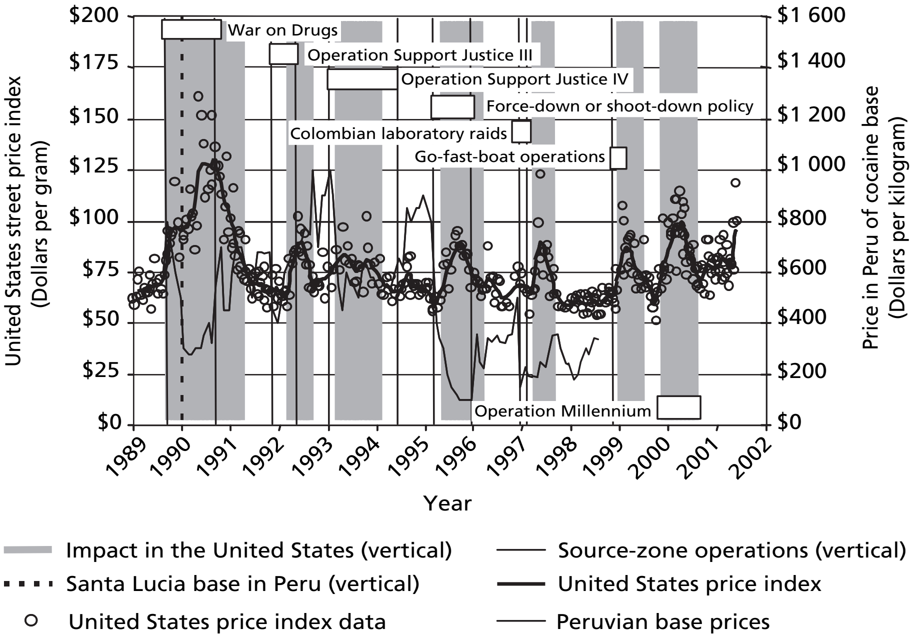
Figure VII. Decline in Peruvian cocaine prices, 1980-1992



Price reactions to source-zone and transit-zone operations in the 1990s

During the 1990s, the Peruvians, the Colombians and the Government of the United States conducted several major cocaine-interception operations targeting the strategic vulnerabilities of the cocaine-distribution system and of a scale large enough to have the potential of damaging the entire system. Figure VIII presents two time series, the United States price index and Peruvian cocaine base prices, overlaid with thin vertical lines representing the beginning and ending of each operation. Vertical grey bands represent the time-lagged intervals of significant price increase on United States streets. It should be noted that, during the 1990s, the United States price index settled to a floor of $55 to $60 whenever there were no disturbances arising from source-zone interception operations.

Figure VIII. Drops in price of cocaine base in Peru and rises in United States price index following each major source-zone or transit-zone operation of the 1990s



The following short descriptions of these operations cover only the highlights [5]:

(a) The "war on drugs" consisted of coordinated operations against the Medellín cartel in Colombia and an extensive blockade of cocaine traffic leaving Colombia to cross the Caribbean. It also included the establishment of the Saint Lucia base, supporting air interceptions against flights originating in the Upper Huallaga Valley of Peru, indicated by the dashed vertical line. Flights from Peru were reported as dropping from 50-60 per week to 5. Cocaine prices rose immediately in the United States and Peruvian cocaine base prices plunged with the drop in air traffic. The war on drugs ended with the Government of Colombia focusing on Pablo Escobar, leader of the Medellín cartel, and the United States diverting naval assets from the Caribbean for the build-up to the first Gulf War. Thereafter, United States cocaine prices returned to the price floor;

(b) Operation Support Justice III was an air interception operation in which the United States provided detection and monitoring support to the Peruvian air force.* On occasion and without sanction from the United States, the Peruvian military interceptors would fire upon trafficker aircraft. While successful for a short period of time, United States support abruptly ended when a

*Operations Support Justice I and II were short training exercises that were not expected to cause noticeable impacts.

Peruvian interceptor mistakenly fired upon a United States C-130, killing one crew member. The vertical lines in figure VIII indicate the beginning and end of Operation Support Justice III, during which time cocaine base prices dipped in Peru. The vertical grey band indicates the sharp price rise in the United States five months later;

(c) Operation Support Justice IV was also an air interception operation, but with stringent controls on the use of deadly force. Air smuggler fees rose from $20,000 to $90,000 for those flying 500 kg of cocaine base to Colombia. Traffickers took the higher smuggling fees out of farmers' pockets as cocaine base prices fell to the level of estimated production costs. Three months later, United States prices rose noticeably, earlier than five months because the Colombians began Operation Support Justice IV two months before the Peruvians. Since the interception threat was insufficient to deter traffickers, cocaine moved unabated to Colombia and United States prices approached their floor even before Operation Support Justice IV ended;

(d) Approval of the force-down or shoot-down policy by the United States Congress and President Clinton gave United States intelligence and tracking support to the Peruvian air force under a strict set of conditions and controls. The use of lethal force was justified by the insurgent and drug trafficker threats to Peru's national sovereignty. The force-down or shoot-down operations achieved a quick success in March 1995 and soon thereafter had deterred more than 80 per cent of the air trafficking. These operations continued throughout the 1990s, with only a brief interruption in December 1995, at which point cocaine base prices abruptly recovered to a less depressed level of $320 per kilogram. Severing the air bridge caused Peruvian farmers to abandon over 66 per cent of their coca crop by 1999. While some riverine and overland smuggling attempted to sustain trafficking from Peru to Colombia, delays, costs and risks prevented it from effectively substituting for the air bridge and restoring prices. Once again, operations began in Colombia two months before Peru and United States street prices began to rise three months after initiation of the force-down or shoot-down policy. Although the force-down or shoot-down policy remained in effect throughout the 1990s, Colombian growers were catching up with supply shortage and, by mid-1996, United States street prices had returned to their floor;

(e) Raids on the Colombian laboratory complex at Miraflores, Colombia, in December 1996 and January 1997. Since this laboratory complex probably processed a third of all the cocaine from Colombia and a large fraction of Peruvian production, which was smuggled into Colombia, both demand for and the price of Peruvian cocaine base abruptly dropped. Without warning, this shock rippled through the distribution chain to cause a surge in United States street prices four months later, illustrating that downstream traffickers are unable to capitalize on the knowledge of a future shortage. The nine-month moving average trend line for the price index in figure VIII truncates the spike in price, but the data points reveal that price actually rose to more than double the floor value. In late 1997, as dispersed replacement laboratories came into production, cocaine base prices returned to their former levels;

(f) Go-fast-boat operations. In 1998, coordinated operations began in the eastern Caribbean to intercept go-fast boats leaving Colombia, headed for the Yucatán peninsula. Disabling rifle fire was approved for use by the United States Coast Guard to destroy the motors of the boats; thus, smugglers could no longer ignore United States Coast Guard helicopters when they were intercepted. Simultaneously, police intercepted go-fast boats arriving at the Yucatán peninsula. This operation shut down the principal trafficking route at that time. Two months after operations began, prices rose on United States streets;

(g) Operation Millennium was a follow-up operation to those shutting down the lanes through the western Caribbean used by go-fast boats. With Drug Enforcement Administration support, the Colombian National Police penetrated the major smuggling organization operating in the go-fast lanes as it attempted to shift from the western Caribbean to the eastern Pacific and arrested its experienced leadership. Shortages ensued, resulting from a lack of trusted smuggler connections, and this caused a prolonged price increase on United States streets, indicated in figure VIII by a vertical grey band.*

During the 1990s, there was a complete correspondence between major source-zone or transit-zone events and United States price index increases above the price floor: every major operation was followed by a price increase and long periods without major source-zone or transit-zone operations had no price increases above the floor. The autoregressive integrated moving average time series analysis used in the research verified the strong statistical correlation between source-zone and transit-zone operations and movements in the United States price index [11]. Five separate classes of Peruvian air-interception events were independently fitted to the price index.** For each source-zone air-interception class, a time lag before the onset of an increase in the United States price index over a six-month window was searched for. A five-month lag fitted all five classes best, an event with odds of happening at random of only $(1/6)^4$ = 8:10,000. Also, the fitted magnitudes representing impacts on United States prices were consistent with the relative magnitudes and consequences of the interception classes. This verified the non-linear deterrence response to multiple interceptions under severe consequences. Finally, the fitted time lags to United States price increases decreased as the cocaine moved towards the United States: five-month lags for air interceptions in Peru, a four-month lag for the Colombian laboratory complex destruction and a two-month lag for the transit-zone go-fast-boat interceptions. Since the early Colombian air interceptions could not be verified in detail, they were excluded from the analysis.

*Larry Lyons of the Drug Enforcement Administration explained this event.

**In fact, seven independent classes of source-zone air-interception events were fitted to the United States price index, but only five were statistically significant. The classes were: three separate months with four or more interceptions, two lists of months with two or three interceptions and two lists of months with a single aircraft interception. The latter two classes distinguished between those before and after the implementation of the force-down or shoot-down policy. Neither of the classes with only one interception per month was statistically significant, while the coefficients representing impacts on prices for the other classes agreed with the relative magnitudes and consequences of the interceptions.

Subsequent to the go-fast-boat operations, there have been so many disruptive events in Colombia that autoregressive integrated moving average (ARIMA) analysis cannot sort them out without more precise data on their source-zone impacts. Nonetheless, the United States price rises represented by the grey bands in figure VIII are certainly a result of these source-zone operations.

Price-quantity relationship during market excursions

The price-quantity relationship derived from STRIDE data for three distinct operational periods with very different street price levels was compared, proving that the slope remains constant although the prices shift up or down for all traffickers in the distribution chain. During the initial epidemic period, from 1983 to mid-1985, prices hesitated at $294 per gram before continuing their decline. Between mid-1997 and late 1998, the gram price dropped to the floor value of $134. During the period 2000-2002, following the go-fast-boat interceptions, the price rose once again to a sustained level of $196 per pure gram. Although these three price levels were so different that the data clusters about each trend line did not overlap the other price excursions, the trend line slopes were the same within random expectations.* This implies that the price mark-ups at the wholesale and dealer levels are independent from the general price of cocaine.

Other time series indicators of damage to the cocaine business

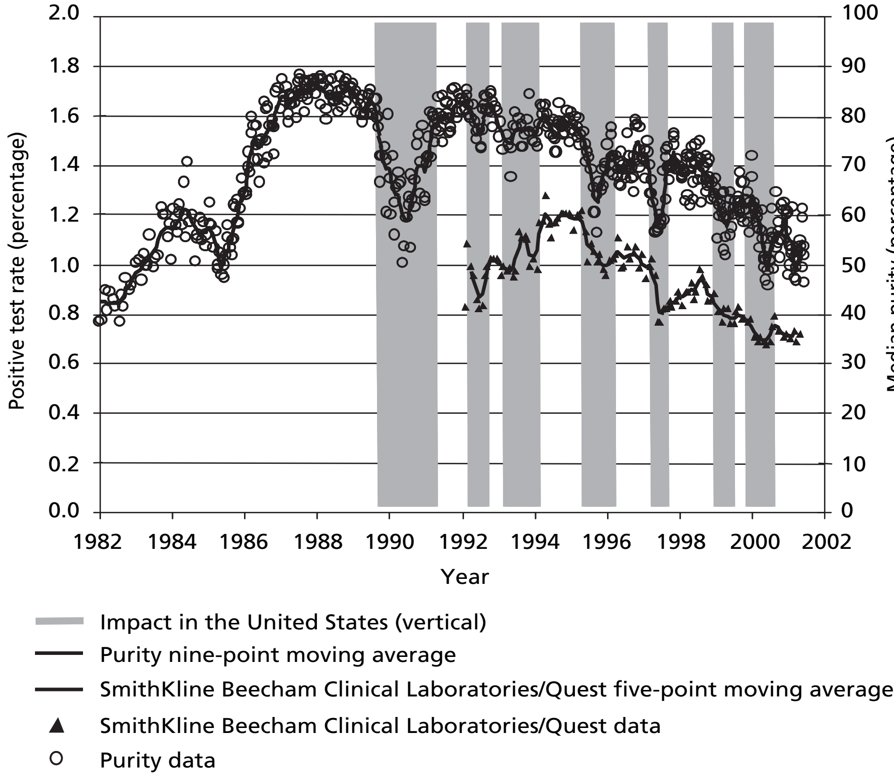
Previous work identified four separate indicators of reduced cocaine availability on United States streets: *(a)* the positive test rate for cocaine in the testing programme conducted by the SmithKline Beecham Clinical Laboratories (SBCL) (now known as Quest Diagnostics); *(b)* the positive test rate for cocaine in the Drug Usage Forecasting (DUF) (now called Arrestee Drug Abuse Monitoring (ADAM)) of the Department of Justice; *(c)* the number of cocaine treatments delivered by treatment centres participating in the Treatment Episode Data Set (TEDS) data collection programme; and *(d)* the number of hospital casualty department admissions linked to cocaine at participating hospitals in the Drug Abuse Warning Network (DAWN) data collection programme [4, 14, 22-25]. These other indicators expand the understanding of the impact of cocaine shortages in the United States and provide further empirical corroboration that the STRIDE-derived price index represents real price movements.

The SBCL/Quest programme typically conducts from 250,000 to more than 800,000 drug tests per month, covering a broad spectrum of the United States workplace. This accounts for the relatively small scatter about a five-month moving average, although the series exhibits pronounced drops during each of the operational periods (see figure IX). It should be noted that the systematic rise

*The null hypothesis that all three slopes are the same cannot be rejected at the 17 per cent level, while the differences in price are significant at the 4:100,000 level.

in positive test rate before 1994 was largely a result of SBCL establishing their customer base rather than an overall increase in cocaine abuse. Conversely, the decline by steps following the force-down or shoot-down operations and again following the go-fast-boat operations were real reductions in usage that did not recover as they slowly did following the Colombian laboratory raids.

Figure IX. Median cocaine purity and positive test rate

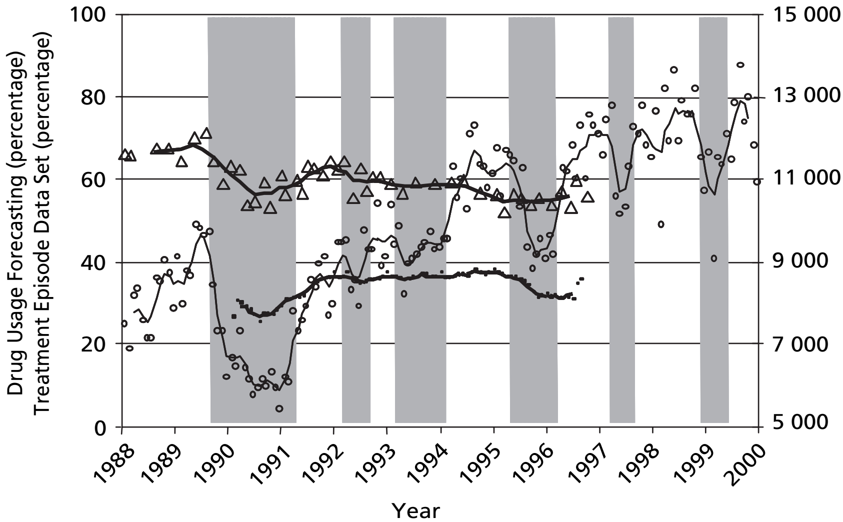


Sources: System To Retrieve Information from Drug Evidence (STRIDE) database and SmithKline Beecham Clinical Laboratories/Quest positive test rate time series.

The other time series in figure IX is median purity, also derived from STRIDE data. Each significant dip during operations makes a large contribution to price index rise because purity is in its denominator. During a shortage, traffickers clearly maintain sales volumes and revenues by cutting the cocaine quality. Early in the cocaine epidemic, when there was high demand and short supply, purity was around 40 per cent, which might be the lowest marketable purity for United States customers. More recently, the downward trend in purity shown in figure IX reversed around 11 September 2001, when most Coast Guard units were transferred to counter-terrorism roles. However, there are other plausible explanations for this reversal; thus, the root cause cannot be determined at this time.

The DUF/ADAM programme conducts semi-random drug testing on several hundred arrestees in 23 major cities every calendar quarter. For the purposes of the research, the percentage of arrestees testing positive is taken as an indicator of cocaine use. Figure X depicts the clear dip during the war on drugs and a small dip during Operation Support Justice III. Methodological problems with this indicator were corrected at the end of this series and may be responsible for the ambivalent behaviour from mid-1994 to mid-1996. Since arrestees abuse cocaine at a high rate and the chemical detected in the tests persists in the blood, even significant drops in availability and frequency of consumption do not appear strongly in this indicator.

Figure X. Data from the Drug Usage Forecasting, Treatment Episode Data Set and Drug Abuse Warning Network, 1988-2000



Impact in the United States (vertical)

△ Drug Usage Forecasting percentage of treatments

- Treatment Episode Data Set percentage of admissions

○ Drug Abuse Warning Network hospital casualty department mentions

The TEDS time series represents the percentage of treatments linked to cocaine as a primary or secondary drug of abuse. During the periods of the war on drugs and the force-down or shoot-down policy, TEDS dropped several percentage points and the tight clustering of data about the moving average suggests that the discernable dips at the beginning of operations Support Justice III and IV are significant.

Although DUF and TEDS move little in response to United States price index increases, DAWN exhibits large, statistically significant downward excursions for each operational period. Since these data include only those users between the ages of 12 and 35, an ageing user population is probably not the cause of the systematic rise. Rather, the shift towards crack cocaine or the occasional purity dip followed a month or so later by recovery, which catches cocaine abusers by surprise, probably combine to explain the overdoses and variation in emergency events.

Implications of the models for control of the supply of illicit narcotics

Given the broad regularities revealed by the analysis and modelling efforts, the question arises of how this information can be used to enhance supply-control operations to stem the traffic of illicit narcotics. Closely related topics are usage and trafficking models as a basis for insights, employing the models to recognize the strategic vulnerabilities of narcotics distribution systems, the limits to drug-business adaptability, measures of effectiveness for counter-drug efforts and the evolution of a common analytical framework for synthesizing disparate counter-drug efforts.

Empirically validated models of drug abuse and trafficking as a basis for insight

Three models are summarized below: the user function of drug abuse, the generic trafficker's business model and the price-quantity scaling of discounts up from farm gate to major drug trafficker organizations and from there down the distribution hierarchy to street users.

Cocaine users vary widely in their dose frequency: the great majority take only a few doses a year, while very heavy users average over 10 doses daily. This pattern follows a mathematical regularity, with the number of users diminishing as an inverse power of the amount they consume. Consequently, only 5 per cent of users are responsible for 80 per cent of all cocaine consumption. And at the opposite extreme, two thirds of all users consume only 2 per cent of the cocaine. This pattern illustrates why it is important to know whether heavy users get quantity discounts from dealers or are dealers themselves, whether heavy users must also pay elevated prices in times of shortage and experience the same purity changes and whether heavy user consumption tracks the more visible rise and fall of cocaine abuse among casual users.

Trafficker business models at every level of distribution share several common traits as a result of a lack of trust, interception risks from law enforcement and intense competition. These common traits permit a useful conceptual simplification that should enhance strategic thinking.

Lack of trust limits organizational size to the span of familiarity of at most a few central individuals, which is about 50-80 people, relegates high-risk tasks such as smuggling to small cohesive teams of 5-10 members and limits the number of customers to about 32 for cocaine and an estimated 17 for heroin and 50 for marijuana. Business associations and recruitment tend to rely on trust mechanisms such as common ethnicity and long-term familiarity and on coercive controls, such as holding family members hostage. Aside from an epidemic expansion during which there is little competitive pressure, lack of trust prevents vertical or horizontal integration and promotes the outsourcing of specialized tasks such as laboratory processing, smuggling and money-laundering.

Interception risk from law enforcement is perceived by smugglers as a 5-10 per cent chance of arrest and, for cocaine, the seizure rate is about 12 per cent for all levels of the distribution chain.

Competition forces price mark-ups towards a common value clustering around 2.5 because encroachment reduces excessive profits, while the need to make criminal profits prevents accepting lower mark-ups. Other narcotics competing with cocaine are probably also pushed towards the same mark-up. While a 2.5 mark-up might seem large, especially when one can turn over capital four or five times a year, interception losses reduce it to 2.2 and there are major expenses such as payments to corrupt officials, risk premiums on employee wages and inflated prices for all illicit materials and services. Evidence from the source zone suggests that traffickers require a 2.5 mark-up, while other aspects of the business model can be adjusted. For example, as supply caught up with demand, two intermediate trafficking levels between farmers and major smugglers were squeezed out of business. Also, collection and smuggling organizations today accept a 2.5 mark-up, although they consolidate hundreds of farmer lots into half-ton loads, which is in sharp contrast with distribution traffickers who limit themselves to approximately 32 customers.

Other drugs deviate in some respects from the cocaine distribution model. Most opium is produced in short growing seasons, so opium, morphine and heroin must be stockpiled somewhere along the distribution chain. Also, unprocessed marijuana and synthetic drugs can be produced close to users. However, heroin distribution does resemble cocaine in that it has a long distribution chain and local law enforcement interception of marijuana or synthetic production can, to a degree, force an increase in the length of these drugs' distribution chains, thereby also increasing their street price.

Because trafficker organizations operate autonomously at each step along the distribution chain, prices compound at successive steps from farm gate to user. This multiplicative increase generates the price-quantity scaling relationship seen for all major illicit drugs of mass distribution in the United States. If price shocks occur at early steps in the cocaine distribution chain, subsequent traffickers have little choice but to pass on the price increase, raised by their standard mark-up. The net result is that shocks are passed along from source zone to consumer zone as a uniform percentage increase, amplifying, for example, a 30 per cent rise in cocaine base price per pure gram ($1.00 to $1.30) at source zone laboratories to a comparable percentage price rise ($150 to $195) for street users. Generally, street users do not see a real price increase directly, but rather experience it as a drop in purity. With loss of satisfaction, many users abandon cocaine, as is revealed by independent indicators such as workplace drug-testing and hospital casualty department visits resulting from cocaine abuse.

Strategic vulnerabilities of the drug business

Each nation's counter-cocaine efforts contribute to the 2.5 mark-up per step all along the distribution chain exhibited in the price-quantity relationship. These efforts reduce supply, increase cocaine prices to users and reduce usage; hence they have tactical utility. But there are also strategic vulnerabilities of the entire distribution chain.

The price-quantity relationship implies large transaction quantities for major drug trafficker organizations and very low unit prices throughout the source zone. Concentration at the top also implies that there are relatively few trafficker organizations conducting the bulk of the business, estimated by the research at between 7 and 28. At the source, coca growers have the least resources of all participants in the cocaine business and their crops are clearly visible from above and cannot be easily relocated. All of these features represent inherent vulnerabilities for the cocaine distribution system as a whole.

Traffickers recognize these vulnerabilities and either locate activities outside areas of police control or pay high premiums to corrupt officials. Cocaine profits also enable insurgent and paramilitary organizations to challenge Andean governments for control of remote areas, thereby creating a nexus of terrorism, lawlessness, weapons-smuggling, human rights violations and environmental damage. Reducing this mayhem would benefit many other areas in addition to those of drug control, including civil order, humanitarian relief, environmental preservation, economic development and political stability.

If traffickers' net revenues are a measure of their ability to avoid law enforcement, then low source-zone prices argue in favour of the international community investing in strategic counter-drug operations there. To understand this, let us consider the fraction of total trafficker revenues accessible to those in the source zone. Farmers, in aggregate, receive only 1 per cent of total cocaine revenues from all levels of trafficking and major drug trafficker organizations receive only 2.4 per cent of total revenues.* Even if major drug trafficker organizations received one half of the transit-zone trafficker's profits, they would get only 5 per cent of total revenues. By comparison, the aggregate of all transit-zone traffickers receive more than 5 per cent of total revenues and are much more diversified than the major drug trafficker organizations. Therefore, additional counter-cocaine dollars spent to deter source-zone farmers or major drug trafficker organizations should cause more damage to the entire distribution system than comparable investments at other steps. Even without causing strategic collapse, increasing the source-zone trafficker costs relative to their low prices should drive up street prices by a comparable percentage, thereby pricing many users out of the market.

*Since 12 per cent of the flow is lost at each step, the actual rate of return on sales is $2.5 \times 0.88 = 2.20$ for each trafficking step. Ignoring source-zone and transit-zone consumption of 8.6 per cent of total flow, the total revenues from all trafficking is the street price minus at-farm-gate costs: $2.2^5 - 0.5 = 51$, where the 0.5 is the farmers' costs. Farmer revenues are 0.5 and major drug trafficker organization revenues are $2.2 - 1.0 = 1.2$. Similarly, transit-zone traffickers' revenues are $2.2^2 - 2.2 = 2.6$.

The early history of the cocaine business in Peru illustrates the economic vulnerability of growers and those collecting and selling to major drug trafficker organizations. From 1980-1990, production caught up with demand, purity rose from 40-80 per cent on United States streets and real prices fell by a factor of six in both Peru and the United States. This price crunch was not shared equally by all traffickers along the distribution chain; rather, two Peruvian middlemen between the farmer and a major smuggler were squeezed out. This means that *(a)* those at the bottom of the distribution chain are the most vulnerable and *(b)* the mark-up of 2.5 in price is sustained by all surviving traffickers even as others are squeezed out of business. The vulnerability of those at the bottom is underscored by the market reaction during Operation Support Justice IV in Peru: full production continued unabated, but the extra cost of increased smuggler pilot fees came out of growers' profits. With these findings, the reactions to an effective crop eradication programme causing large persistent supply shortages can be predicted. According to the price-quantity relationship, shortages cause a serious drop in the street purity of cocaine, a significant rise in prices and, eventually, users defect to other drugs or quit altogether. With fewer users, there is less need for trafficking, yet the business model implies major drug trafficker organization transaction quantities remain unchanged at the top of the price-quantity relationship. Hence some major drug trafficker organizations would be squeezed out and law enforcement would have fewer major drug trafficker organizations to pursue.

Strategic supply-control efforts focused on the source zone could inflict much more than incremental damage on the cocaine business. Severing the air bridge between Peru and Colombia caused a collapse across the entire Peruvian cocaine business by reducing cultivation to only 18 per cent above pre-epidemic levels. Similarly, the forced eradication in the Chapare region of Bolivia reduced cultivation there to 12 per cent of its peak production. These effective supply-control efforts illustrate that a focus on source-zone vulnerabilities can cause a collapse and sustain much diminished trafficking through a combination of eradication and deterrence of major traffickers.

Supply-control strategies

Some of the options for directing supply-control efforts to exploit these source-zone vulnerabilities are given below:

(a) The coca eradication campaign should be continued and strengthened because, among other things, it costs law enforcement only half as much to spray a hectare as it costs growers to replace the crop, and coca growers have the least resources of anyone in the cocaine business. Since the coca crop is visible and impractical to disperse,* intense eradication could rapidly undermine

*Although coca grows throughout large portions of the Amazon basin, only 10 per cent or less of the land that could support coca is cultivated and 50 per cent of all production comes from only 9 per cent of this 10 per cent. Infrastructure constraints force crop concentration and, in unpublished work, the present researchers estimated that production costs would increase as the inverse of the square root of the density; for example, reducing the density to one quarter of its present level would double costs.

the major revenue source supporting lawlessness and ongoing environmental damage in Colombia [5]. If successful, there should be preparations for humanitarian assistance to some indigenous growers, as well as relocation assistance to the excess population in growing areas, who were attracted there by the boomtown profits of the past;

(b) Coca cultivation should be forced towards the interior of Colombia to increase the costs of transporting essential supplies and chemicals and of maintaining the illicit economy in a remote area and to create opportunities to intercept the transport of illegal fuel and chemicals for coca processing. As a result of previous successful blockades of transportation, such as the Peruvian airbridge-denial operations and effective manual eradication in Bolivia, 88 per cent of the coca cultivation in excess of 1970 levels is now grown in Colombia;

(c) The leverage offered by the psychology of deterrence should be exercised against a limited number of major drug trafficker organizations [5, 7, 9]. Arresting the leaders and principal members of only 9 per cent of the major drug trafficker organizations (one to three organizations) each year and extraditing them to the United States should have a deterrent effect equivalent to the threat of lethal force and cause 80 per cent to quit;

(d) Raiding cocaine hydrochloride laboratories, intercepting cocaine or essential chemical movements within Colombia or intercepting boats leaving Colombia would require a very much greater effort to deter others and have a strategic effect and might not be practical [7]. However, technological breakthroughs in the ability to detect laboratories, material movements, or go-fast boats could facilitate a strategic breakthrough. Nonetheless, these actions are valuable tactical efforts in supply-control;

(e) Three additional actions would enhance deterrence effects: (i) increasing the level of distrust within and among trafficker organizations by employing informants and other intelligence-gathering methods to exploit traffickers' greatest fear; (ii) working to build trustworthy governments, justice systems, police and military organizations in source-zone and transit-zone countries to encourage cooperation from all those wishing to escape from fear and coercion, while ensuring certain justice for arrested traffickers; (iii) being persistent, so that traffickers no longer believe they can wait out the counter-drug commitment to achieving and sustaining a strategic effect;

(f) Surge operations should be employed to enhance the impact of limited forces (see the following section).

Limits to drug business adaptability

Evidence has been presented above arguing that most traffickers' business models are trapped by the price-quantity relationship. Also, lack of trust causes trafficker organizations to remain small, while their leaders limit the information available to most subordinates to thwart their ambitions of taking over. These features limit the flexibility and adaptability of individual trafficker organizations. The adaptability of the cocaine business as a whole is that of the anthill: when a few are lost, others take their place. Furthermore, it takes time to replace an organization after most of its members have been arrested and this creates a law enforcement opportunity to conduct surge operations.

A surge strategy consists of focusing overwhelming pressure on one area until those traffickers are put out of business or forced into a risky move to a new method or location. It is probably impractical for trafficker groups to simply close down for a few months in response to focused law enforcement pressure because they must continually turn over capital rapidly, pay suppliers, avoid competition, pay corrupt officials for protection and so forth. For example, after the effective interception of western Caribbean go-fast-boat lanes, the major smuggling organization operating those lanes immediately attempted to relocate to the Pacific coast, but the police exploited the forced change and were able to arrest the principals. In this case, the smuggling connections took more than a year to reconstitute, providing law enforcement an opportunity to focus elsewhere. Another surge operation, Operation Frontier Shield, was directed against trafficking into Puerto Rico. It quickly deterred much of the go-fast-boat smuggling and levels have remained lower to this day [9]. In preparation for a surge, law enforcement needs to monitor trafficker operations to identify the typical delays between successive trafficking events or any seasonal pattern or characteristic, thus enabling the interception surge to cause the most damage and to sustain the pressure until the traffickers are forced to react.

Measures of effectiveness for counter-drug efforts

The degree of effectiveness of counter-drug efforts should be gauged by the degree of damage inflicted on traffickers' businesses either locally or throughout the entire distribution chain. The principal quantification of risk from the traffickers' perspective is the price mark-up per pure gram necessary to stay in business. For example, Peruvian cocaine-base prices plunged below break-even levels after the implementation of the force-down or shoot-down interception policy for flights over the air bridge to Colombia. Such a drop is clear evidence of damage to the cocaine business and four years later it was found that the excess coca cultivation beyond historical levels in Peru had dropped by 80 per cent. Often, the real price mark-up is expressed as a significant drop in purity because traffickers cover their shortages by adulterating the higher grade product. Street purity may be the best estimate of shortage because it varies more than price during supply shortages. Countries therefore need to collect and report on the purity and transaction quantity associated with drug prices at each trafficking level.

A common indicator of supply-control effectiveness is the seizure rate. Unfortunately, long-term growth in seizures by themselves typically indicate more trafficking rather than less narcotic being delivered [8]. In the short term, the seizure rate is difficult to measure accurately because traffickers hide the true volume of traffic, which is in the denominator of the ratio. Nonetheless, approximate estimates might be sufficient to check consistency with observed price mark-ups, since a 10-15 per cent seizure rate should be associated with a 2.5 mark-up. Significant and rapid changes in seizure rates should induce corresponding increases in price mark-ups or cause observable changes to trafficking practices. Very high seizure rates and associated arrests are likely to deter traffickers from continuing, hence seizures can begin to fall although trafficking is being damaged. This happened after the implementation of the force-down or shoot-down policy: traffic plunged as did coca prices, but a one-month relaxation of interception pressure enabled trafficking and prices to recover sharply and interception was quickly reinstated [5]. Since seizures are a poor indicator by themselves, a better measure would be arrest rates, along with the consequences of arrest, because arrests are much more likely to deter traffickers [5, 7]. Below a deterrence threshold, traffickers ignore risks, but above this threshold deterrence can cause four to eight times as many to quit as are arrested [5, 9].

Other measures might provide early warning of traffickers gaining advantage over law enforcement: for example, the willingness of major traffickers in Colombia and Mexico after 11 September 2001 to cooperatively assemble large 10-metric-ton-plus loads for smugglers indicated growing horizontal integration resulting from reduced law enforcement pressure. Similarly, the greater number of customers per marijuana dealer compared to cocaine, or cocaine compared to heroin, indicates relatively greater law enforcement effort for the latter drugs.

Some indicators directly measure the defection of users from cocaine abuse. The SBCL/Quest positive test rate data did this on a monthly basis and quickly revealed significant changes among the mostly casual users. The DAWN data clearly reveal the impact of effective supply-control operations; however, the systematic rise in hospital emergencies while abuse rates have been declining calls for a clear interpretation. Unfortunately, these measures at the street-use level lag four to five months behind the initiation of operations in the source zone. Their principal value is to verify the strategic impact and calibrate more local measures.

A common analytical framework for counter-drug strategies

These models and associated methods offer a set of working hypotheses for characterizing the illicit narcotics business worldwide. These working hypotheses could assist researchers, the intelligence community, law enforcement and other counter-drug operational planners and policy makers who need a common analytical framework to measure, compare and express their specialized knowledge and interests. Such a framework facilitates cooperative counter-drug efforts at different steps along the distribution chain, as well as among different agencies and countries, who often have divergent goals and employ measures of performance that cannot be readily compared or related to one another.

To be effective, a common framework would have to be capable of addressing the variety of situations in drug trafficking and be empirically verifiable. The models and methods developed in the course of the present research provide the beginnings of such a framework. Three models form the backbone of the framework: the generic model of trafficker organizations resulting from the simplifying constraints, the price-quantity relationship and associated multiplicative market price dynamics and the user function and associated validating measures of abuse. A complementary set of methods deals with the wide variation in trafficker and user behaviour and the fragmented data representing what is known about that behaviour. For example, taking medians or other percentiles of many sub-samples yields useful parameters for characterizing the long-tailed distributions of cocaine consumption or price normalized by purity. These parameters are always well defined, handle pathological cases (such as infinite price for zero purity), remain stable with the addition of more data and are well behaved when analysed by the familiar statistical analysis methods. Since price mark-ups are ratios, as are seizure rates and the breakdown of purchased quantities into those sold, such data are plotted on logarithmic scales. When a logarithmic variable is plotted against another, one might discover a scaling relationship such as that between normalized price and transaction quantity. While such plots suppress much detail and compress wide variations, they clarify underlying regularities that motivate and constrain the narcotics trade. Variations in these regularities can then take on significance and the structures they reveal point to vulnerabilities and consequences for traffickers. These structures can be empirically validated and relate the efforts of most counter-drug efforts all along the distribution path from farm gate to street. Finally, these patterns and structures can be related to the trafficker's perspective, especially considerations of risk, competition and lack of trust.

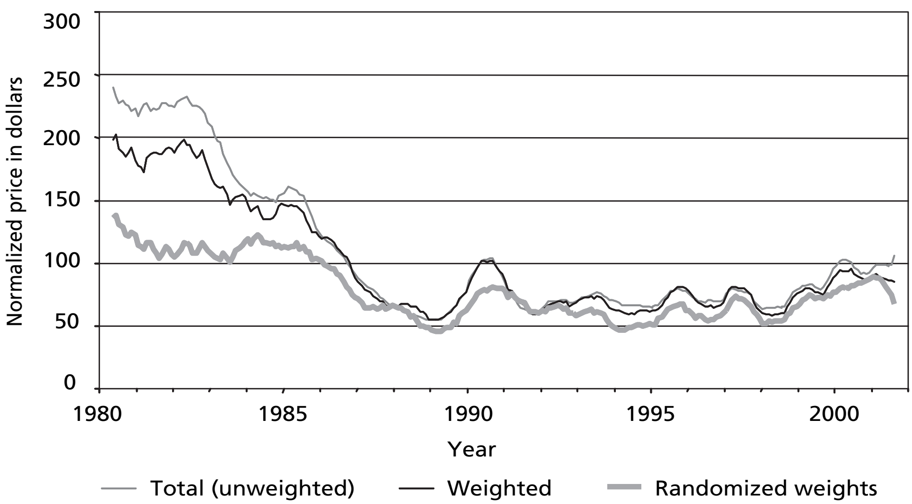
While specific information on trafficker activity tends to be very sensitive and often not shared by the originating agency or country, this framework implies that even consolidated listings or samplings of events described by a few key characteristics can be useful. Although the analytical methods and models used here do not substitute for detailed operational intelligence, they do provide a verifiable and dynamic map of the worldwide cocaine trade and possibly a modelling approach for other illicit narcotics of mass distribution. Such a map is useful in recognizing the strategic vulnerabilities of narcotics trafficking and should encourage broad-based public and political support for joint efforts, effective planning, sustained support and widely accepted measures of performance.

Annex A. Robustness of the normalized price index

As noted above, the STRIDE data collection process reflects local priorities and there is no effort to orchestrate a rigorously balanced statistical survey of all United States markets. Here the robustness of the STRIDE-derived normalized price index for cocaine is explored for purposes of identifying major data trends. The potential impact of applying standard statistical survey techniques in pursuit of precisely estimating, over time, the true national average price of street-level cocaine transactions is addressed.

As a simple illustrative exercise, figures A.I and A.II are considered. The pairs of thin-lined curves are identical in the two figures. Derived from STRIDE, the upper (lighter) curve depicts monthly median normalized cocaine prices, with each median computed across the entire United States. The lower (somewhat darker) curve, similarly based on STRIDE, also presents a monthly measure of United States normalized cocaine prices, calculated as a weighted average of the monthly medians associated with each state and Washington, D.C., with the weights matching the corresponding actual sampling frequencies that appear in STRIDE. Clearly, there is no practical difference between the pairs of curves.

Figure A.I. Alternative cocaine price index construction methodologies: random factors x prices and randomized weights



More rigorous methodologies, however, would argue that the Drug Enforcement Administration sampling is not representative of actual market purchase frequencies and seek to replace, at great time and expense, the implicit Drug Enforcement Administration weights by some measure of total cocaine consumption levels. Figure A.I addresses what the potential impact of such an approach would be, by considering the extreme case of arbitrarily random weights for each reporting jurisdiction in each month (that is, a random number between 0 and 1, properly normalized by the sum of all of the random numbers). One typical realization of this process is given by the thick (lowest) curve in figure A.I. It should be noted that there is some consistent deviation in the normalized price values, with the thick (lowest) curve generally taking on smaller numerical values. This ranking results because the randomly chosen weights tend to de-emphasize those states with relatively more Drug Enforcement Administration sampling, precisely the ones that generally have higher normalized prices. However, as far as fundamental trends and specific excursions in normalized price are concerned, the thick (lowest) curve captures these extremely well. If the focus is indeed on trends and features, the simple problem considered here indicates that it may not be cost-effective to attempt to obtain and incorporate excessively precise consumption figures into the analysis.

Figure A.II. Alternative cocaine price index construction methodologies: random factors x prices and "proportional" randomized weights

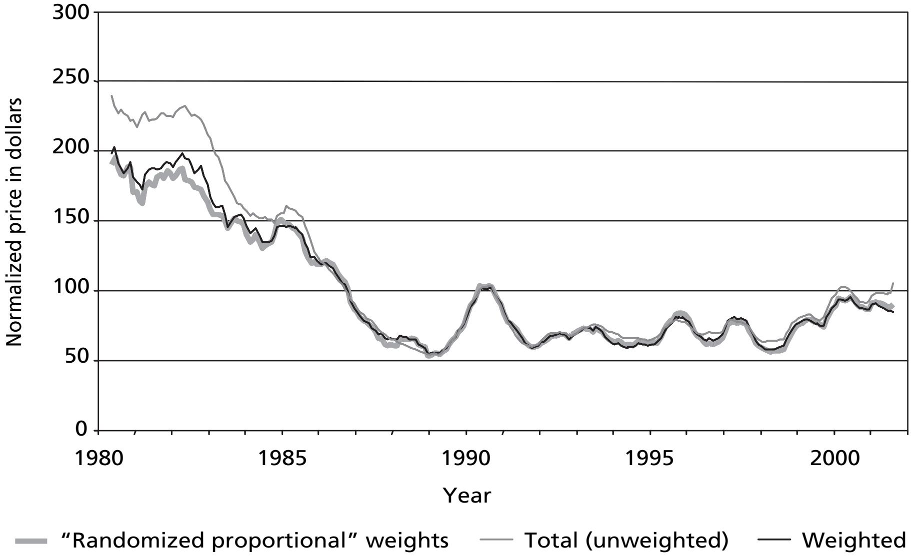


Figure A.II extends the simulation study by assuming that the Drug Enforcement Administration sampling frequencies are probably not too unrealistic, that is, they would naturally focus more on higher-consumption areas. Rather than using completely random weights, each reporting jurisdiction's weight is now computed by beginning with the Drug Enforcement Administration sampling count, multiplying by a random number between 0 and 1 and then normalizing all of these products accordingly. The particularly thick (lowest) curve shown in figure A.II is a typical realization of this process. It should be observed that it essentially matches the other curves, a result that further strengthens the conclusion of the preceding paragraph.

Finally, it was noted that an additional randomization had been applied to the determination of the thick (lowest) curves in figures A.I and A.II. Every STRIDE price was multiplied by a random factor, taking on values between 0.5 and 1.5. Thus, these figures also address the issue of how much effort should be devoted to collecting precisely representative cocaine prices. The robustness of the curves is consistent with the notion that substantive cocaine price and purity trends can be detected even when the available data do not conform perfectly with stratified sampling theory.

Annex B. Detecting trends: sample mean vs. sample median

For extremely heavy-tailed distributions lacking theoretical moments, typical of STRIDE normalized price data, the sample mean is extremely unstable (for example, it does not converge with increasing sample sizes) and is not an insightful measure of central tendency. Comparing individual sample means in an attempt to detect underlying population trends is a completely meaningless exercise. The wild variability of the sample mean probably disguises trends, even when they are profound in actual magnitude and duration. However, the sample median is an insightful distributional characterization and a powerful tool for detecting trends.

These points are illustrated by the following simple simulation study. Let us assume that price data are sampled daily over a five-week period, with the underlying theoretical distribution of prices adhering to Pareto distributions with systematically changing medians: constant within any given day, but increasing by 10 per cent each day, that is, essentially doubling every week. Suppose that 33 data values are collected daily, generating a weekly sample of 231 points, along with an associated sample mean and sample median. The entire five-week sampling process is iterated repeatedly to obtain the following characterizations: *(a)* the probability that the weekly sample means or medians portray the proper ordering, that is, Week 1 < Week 2 < Week 3 < Week 4 < Week 5 and *(b)* the probability that the weekly sample means or medians have Week 1 < Week 5.

The sample means portray the proper five-week ordering with a probability of only 0.03, while the sample median figure is 0.98. Similar results hold for the second measure of performance. The sample means indicate Week 1 < Week 5 with a probability of 0.84, that is, one-sixth of the time the sample means reverse the ordering, even though the actual difference is a factor of $2^4 = 16$. The sample median, on the other hand, reflects the proper ordering essentially all of the time.

Annex C. Consumption and seizures by trafficking zone

Where appropriate, the amounts seized by country are aggregated into regions or zones, as indicated in table C.1. The columns of table C.1 parallel the steps for estimating consumption. Population from an atlas scaled by United Nations prevalence yields total users in thousands. Assuming that all users consume the same amount, total consumption can be expressed. Combined with seizures, this should balance total production. Given total production, it is possible to calculate the consumption per user that balances the flows.

For total production, the United States estimate of Colombian production also reported by the United Nations is adopted [6].* The rationale for this is to approximately compensate for two opposing adjustments: scaling downward for counting immature crops in the Colombian surveys and for inefficiencies plus traditional use in Peru and Bolivia versus scaling up for market purity, which is assumed to be about 80 per cent.** Since these upward and downward adjustments are of similar size, they approximately cancel each other out.

Cocaine seizure totals for Canada, Western Europe and the United States need to be adjusted because seizures of lower purity street-grade cocaine remove less from the flow than seizures of wholesale kilogram bricks. For Europe, it is assumed that one half of all seizures, two steps of four, are of street-grade cocaine and total seizures are adjusted downward from 44 to 39 metric tons. Similarly, for the United States, it is

*The joint Colombian and United Nations coca survey effort yielded a much larger estimate of production: 680 metric tons versus 520 for the United States coca survey. Without going into the details of the differences between the two surveys, choosing the lower estimate is justified because a large portion, between 30 and 40 per cent, of the Colombian crop must be immature as a result of the recent rapid expansion and replacement of eradication losses.

**Here, the United Nations estimate is taken rather than the STRIDE value for 2000 of 70 per cent.

assumed that two thirds of seizures are street-grade cocaine, which adjusts 134 metric tons down to 112. For source-zone and transit-zone countries, the amounts seized at the consumer level are considered trivial and are therefore ignored. This is justified because, in the source-zone and transit-zone countries and regions, seizures exceed consumption by factors of two to four. However, in consumer-zone countries, consumption exceeds seizures by large percentages.

Table C.1. Seizures and estimated consumption for countries, regions and zones

Country or region	*Population (Thousands)*	*Prevalence (Percentage)*	*Users (Thousands)*	*Consumed (Metric tons)*	*Seized (Metric tons)*	*Consumed (Metric tons*	*Seized (Metric tons)*
Source zones							
Colombia	42 300	1.2	508	13.0	63.9	13.0	63.9
Peru	25 700	1.0	257	6.6	11.3	6.6	11.3
Bolivia	8 300	0.7	58	1.5	7.7	1.5	7.7
Transit zones							
Central America	36 200	1.0	358	9.2	16.7	29.3	63.7
Mexico	98 900	0.5	495	12.7	34.6		
Venezuela	24 200	1.2	290	7.5	12.4		
Ecuador	12 600	1.5	189	4.9	10.6	10.9	22.7
Caribbean	25 620	0.9	236	6.1	12.1		
Consumer zones							
Brazil	170 100	0.8	1 361	34.9	7.6	65.0	12.1
Chile	15 200	2.1	319	8.2	2.9		
Argentina	37 000	2.3	851	21.8	1.6		
Africa	380 600	0.3	1 314	33.7	0.5	33.7	0.5
Western Europe	314 000	0.7	2 200	56.5	44.1	56.5	44.1
Adjusted seizures							39
Canada	30 200	0.7	211	5.4	1.7	216.3	134.0
United States	273 800	3.0	8 214	210.9	132.3		
Adjusted seizures							112
Total	1 494 720	1.1	16 861	432.8	360.0	432.8	332.2

Total seizures	332
Total consumption	433
Total production	765
Consumption per user per year	26 grams

References

1. United States of America, Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, *National Household Survey on Drug Abuse* (Washington, D.C., 1996).
2. Rockwell International Special Investigations, *Measuring Deterrence: Approach and Methodology*, contract No. TC-89-037, October 1989.
3. Mary Lane and others, *Measuring the Deterrent Effect of Enforcement Operations on Drug Smuggling, 1991-1999* (Washington, D.C., Office of National Drug Control Policy, 2001).
4. B. Crane, A. Rivolo and G. Comfort, "An empirical examination of counter-drug interdiction program effectiveness", paper No. P-3219 (Alexandria, Virginia, Institute for Defense Analyses, 1997).
5. R. Anthony, B. Crane and S. Hanson, "Deterrence effects and Peru's force-down/shoot-down policy: lessons learned for counter-cocaine interdiction operations", paper No. P-3472 (Alexandria, Virginia, Institute for Defense Analyses, 2000).
6. United Nations, *Global Illicit Drug Trends, 2001* (United Nations publication, Sales No. E.01.XI.11).
7. R. Anthony, L. Lyons and A. Fries, "Evaluation of the strategic value of interdicting cocaine laboratories and arresting major traffickers", paper No. P-3859 (Alexandria, Virginia, Institute for Defense Analyses, 2004).
8. United Nations Office for Drug Control and Crime Prevention, *World Drug Report 2000* (United Nations publication, Sales No. GV.E.00.0.10).
9. R. Anthony, "A calibrated model of the psychology of deterrence for narcotics traffickers", *Bulletin on Narcotics* (United Nations publication), vol. LVI, No. 1 (2004).
10. S. Soneji and A. Fries, "An intervention analysis of cocaine prices in the U.S.", document No. D-2384 (Alexandria, Virginia, Institute for Defense Analyses, 1999).
11. S. Soneji and others, "Time series intervention analyses on U.S. cocaine prices", *Proceedings of the Fifth Annual U.S. Army Conference on Applied Statistics*, B. B. Bodt, ed. (United States Military Academy West Point, New York, United States of America and Army Research Laboratory (ARL-SR-110), Aberdeen Proving Ground, Maryland, United States of America, 2001).
12. A. Fries, "Academic criticisms of the DEA's STRIDE database", document No. D-2768 (Alexandria, Virginia, Institute for Defense Analyses, 2002).
13. J. Landon and P. Wilson, "Analysis of STRIDE data on cocaine transactions, 1980-2001", document No. TR 2002/2 (Washington, D.C., Institute for Reliability and Risk Analysis, George Washington University, 2002).
14. A. Fries, "Empirical corroboration of cocaine price and purity trends", *Proceedings of the Expert Group on Technical Challenges to the Drug Community* (Vienna, United Nations Office on Drugs and Crime and United States Office of National Drug Control Policy, 2003).
15. R. Anthony, "Farmgate-to-street model of narcotics trafficking", *Proceedings of the Expert Group on Technical Challenges to the Drug Community* (Vienna, United Nations Office on Drugs and Crime and United States Office of National Drug Control Policy, 2003).

16. Arturo Briceno, *An Econometric Model of the Peruvian Coca Supply* (Lima, 1992).
17. United States of America, Department of State, Bureau for International Narcotics and Law Enforcement Affairs, *International Narcotics Control Strategy Reports* (Washington, D.C., 1985-2000).
18. S. A. Cuanto, *Coca's Impact in Peruvian Economy, Peru 1980-1992: Methodological Appendix* (Lima, 1993).
19. World Bank, Environment and Agriculture Division, *Agricultural Policies for Economic Efficiency* (Peru, 1992).
20. J. Nunez and R. Reategui, *La Economia Cocalera en el Alto Huallaga: Impacto Economico*, degree paper (Lima, Universidad del Pacífico, 1990).
21. United Nations, Office for Drug Control and Crime Prevention, *Situation Analysis on the Fall in Coca Prices in Peru* (1995).
22. SmithKline Beecham Clinical Laboratory, positive test internal monthly data for cocaine, opiates and amphetamines, October 1996 and updates.
23. United States of America, Department of Justice, National Institute of Justice, *Drug Use Forecast*, 1989-1997.
24. United States of America, Department of Health and Human Services, Substance Abuse and Mental Health and Mental Health Services Administration, *Treatment Episode Data Set*, *1992*-1998.
25. United States of America, Department of Health and Human Services, Substance Abuse and Mental Health and Mental Health Services Administration, *Drug Abuse Warning Network*, *1989*-2001.

A calibrated model of the psychology of deterrence*

R. ANTHONY

*Research Staff Member, Institute for Defense Analyses,
Alexandria, Virginia, United States of America*

ABSTRACT

Can law enforcement cause 80 per cent of narcotics traffickers to give up their favoured methods or cease smuggling altogether, even though seizure rates under the most favourable conditions seldom exceed 30 per cent and arrest rates are lower still? Several counter-cocaine operations have achieved these goals. Principal among them is the force-down or shoot-down policy, in which the Government of the United States of America assisted the Peruvian Air Force with detection and monitoring so that smugglers could be prevented from flying cocaine base from Peru to Colombia. With the threat of lethal force, an 8 to 12 per cent interception rate held down trafficking to less than 15 per cent of former levels, causing the collapse of the Peruvian cocaine trade. Less severe consequences worked at higher interception rates in the transit zone to the United States. The present article presents an analysis of an extensive set of interviews with incarcerated drug smugglers that derives a mathematical function representing their willingness to smuggle. This function is the basis for a model of the psychology of deterrence, which was verified and calibrated using detailed data from counter-cocaine operations. Three major findings are that (a) *traffickers ignore the risk of interception up to a point,* (b) *interception rates beyond a certain threshold cause trafficking activity to collapse rapidly and* (c) *some traffickers are undeterrable, even when faced with certain interception.*

Keywords: deterrence; cocaine; interception; trafficking.

Introduction

Cocaine smuggling is a high-risk activity. Even though some smugglers are attracted by risk or desperate for quick profits, one would expect that, as the chances of being caught, imprisoned or killed increased, many would stop smuggling. The present article focuses on the following two questions:

(a) What is the level of risk necessary to deter most smugglers?

*The views expressed in the present article are those of the author and do not represent an official position of the Institute for Defense Analyses or of its sponsors.

(b) If smugglers can be deterred, can their behaviour be predicted through use of a mathematical model?

These issues can be addressed directly, by examining interception operations at different levels of risk, indirectly, by examining the literature on risk-taking, or subjectively, by interviewing captured smugglers. As data are available for only a few operations, they are insufficient to derive a functional form without other justification to link the data sets. Nevertheless, these known operations can be compared with a model, once one has been developed. Although the literature provides some clues as to the form of a deterrence model, it is not sufficient to build a useful mathematical model of the behaviour of a large group of smugglers. While interviews with imprisoned smugglers seem liable to bias, prisoners serving long prison terms apparently want to share their knowledge and brag about their exploits, if they can do so without self-incrimination or adding to their term of imprisonment. Although interview data represent only the opinions of smugglers, their opinions are the basis for their decisions to smuggle and thus the ultimate basis for deterrence. Fortunately, interview data are available with sufficient detail to build a mathematical model of deterrence.

Throughout the present article, the following military definition of deterrence is used: "the prevention from action by fear of consequences—deterrence is a state of mind brought about by the existence of a credible threat of unacceptable counteraction" [1]. This definition addresses the following essential elements: the perpetrator's psychology (state of mind) that arises from the probability of consequences (a credible threat) and the consequences themselves (unacceptable counteraction).

Interview data

In 1989, the United States Customs Service sponsored a research team from Rockwell International to interview former drug smugglers in federal prisons concerning the conditions under which they would be willing to continue various illicit activities [2]. The interview team contacted the United States Bureau of Prisons and selected inmates whose offence code identified them as currently serving sentences for violations of pertinent sections of the Comprehensive Drug Abuse Prevention and Control Act of 1970. The research team selected a sample of inmates from nine federal prisons in five states and one state prison in Texas. To avoid selection biases, no distinction was made on the basis of weight or type of narcotic, arresting agency or location of arrest, length of sentence, age, sex or demographic profile. Prior to the interviews, inmates were told that their answers would be treated in confidence and that any information given by them would not be traceable to them. Almost half of the sample of inmates with drug convictions agreed to participate, yielding a final sample size of 112. Of those, 109 gave useful responses and five of those did not answer some of the key questions. This is nevertheless a high response rate for so long an interview process and so sensitive a subject.

The responding inmates were quite diverse demographically and in terms of experience. They were between 20 and 50 years old; half were citizens of the United States, while the others consisted of equal numbers of Colombians, Mexicans and other, various nationalities. There were about equal numbers of secondary school dropouts, those who had completed secondary education and university graduates, as well as a few holding doctoral degrees. A total of 50 per cent had smuggled marijuana, 40 per cent had smuggled cocaine and 10 per cent heroin. Some had smuggled more than one drug or drugs other than cocaine, heroin or marijuana.* They had smuggled drugs between one and 10 times and an average of approximately six times.

Interviewers were selected on the basis of their investigative expertise, bilingual capability, law enforcement background and drug control experience in order to ensure that frivolous answers were kept to a minimum. Interviewers commented that the variety of inmate respondents and the commonality of their responses indicated that there were no identifiable biases. Some of the survey respondents commented that, before being imprisoned, they had underestimated the chances and consequences of being caught. Some inmates went so far as to suggest that the authorities should conduct a campaign of informing currently active smugglers about the penalties for drug smuggling. This led the interviewers to conclude that the answers given from the perspective of the associate were more likely than the answers given from the point of view of the inmate himself or herself to be representative of the majority of active smugglers.

Three principal questions (Q1, Q2 and Q3), all of a common form, yielded the data that became the framework for a mathematical model of the willingness to smuggle. Those questions and the allowed responses to each of them were as follows:

> I would not smuggle drugs into the United States if my chances of getting caught (Q1); caught and convicted (Q2); caught, convicted and imprisoned (Q3); were 1 in 10 times (R1); 1 in 5 times (R2); 2 in 5 times (R3); 4 in 5 times (R4); or every time (R5).

Inmates were asked to answer the questions from two separate points of view. The first, under the heading of "self", required responses from inmates regarding their own actions, perspectives and future smuggling intentions. The second, under the heading of "associate", required similar responses, but from the perspective of a former associate or friend in the smuggling business. In each of questions Q1, Q2 and Q3, inmates were asked to choose, first for "self" and again for "associate", from one of five probabilities of interception: R1, R2, R3, R4 or R5. Table 1 summarizes the responses provided by the inmates. It should be noted that the responses in each of these probability-of-interception categories represented those who would be deterred by the selected level of risk, but not

*The data obtained were aggregated and it is consequently not certain that the cocaine smugglers' responses were distributed similarly to those of the other smugglers. However, the eventual consistency and uniformity of the results justify combining the data.

by the next lower level of risk. Thus, someone not willing to smuggle against a probability of being caught of "2 in 5 times" might be willing to do so if the probability were only slightly greater than the previous category of "1 in 5 times".

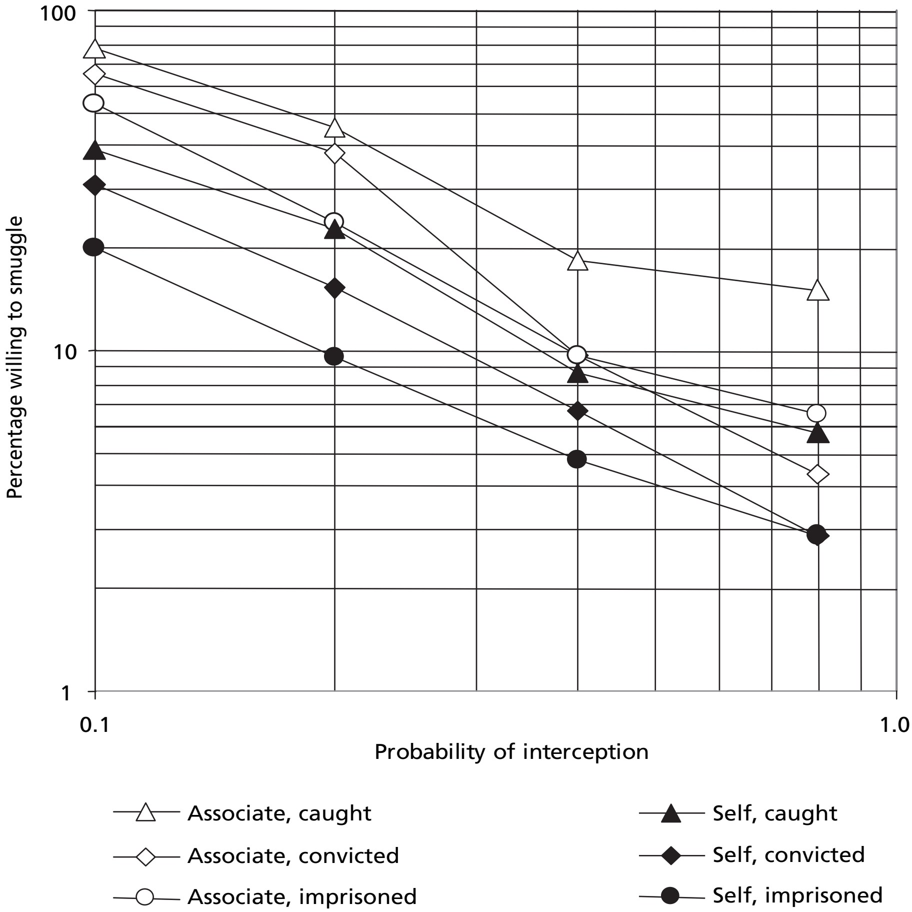
Table 1. Additional number of inmates not willing to smuggle as risk increases from one probability to the next higher probability

	Consequences					
	Caught, convicted and imprisoned		*Caught and convicted*		*Caught*	
Probability of interception	*Self*	*Associate*	*Self*	*Associate*	*Self*	*Associate*
1 in 10	83	43	72	32	63	21
1 in 5	11	27	16	25	17	29
2 in 5	5	13	9	26	15	25
4 in 5	2	3	4	5	3	3
Every time	3	6	3	4	6	14
Total respondents	104	92	104	92	104	92

For the remainder of the present article, the data in table 1 should be considered to constitute six separate subsets of data for the combinations of the two types of response (self versus associate) and the three degrees of severity of punishment (imprisoned, convicted, caught). Figure I displays all six data subsets as cumulative trends on a log-log plot. Each trend declines in an approximately linear fashion, and the extension of each trend would intersect the probability of interception of 1.0 (anticipation of certain apprehension) at a finite value. This means that some inmates would be willing to smuggle even if they knew beforehand that they would be caught, consistent with comments volunteered by some of the survey respondents. It should be noted that the proportions for each data subset that correspond to a "probability of interception" equal to 0.1 follow a regular progression in "percentage willing to smuggle" for all six cases. From the top down, the ordering is "associate, caught", "associate, convicted", "associate, imprisoned", "self, caught", "self, convicted" and "self, imprisoned". This progression makes sense because the inmates were more willing to smuggle against less severe penalties and because imprisoned smugglers judge the consequences more severely than they imagine their former associates would have while free. At a probability of 0.4 and on to 0.8, however, responses to the "associate, convicted" question abruptly deviate from this logical order. It is assumed that some element of the interview process corrupted these few responses; therefore, this case was ultimately dropped from the data subsets employed to characterize the willingness function mathematically.*

*To restore the regular progression in figure I and table 1 would require only six responses to be moved from "2 in 5 times" to "every time". Moving three responses would bring the number in that bin for "associate, convicted" within one standard error deviation from the estimated value.

Figure I. Proportion of inmates willing to smuggle for each probability of interception for each of the six data subsets



Mathematical representation of the willingness function

Three simple functional formulations of the willingness function were considered. In all three, W represents the willingness to smuggle and P_I represents the probability of interception. First, the Pareto function, $W(P_I) = (1 - P_I)^{\alpha - 1}$, matches the general qualitative appearance of military deterrence plots; there are so many examples of this behaviour that it is a standard technique for representing data. Second, the attrition filter function, $W(P_I) = \exp(-\alpha P_I)$, represents a multi-stage process of absorption through a series of filters; for example, the willingness of criminals to repeat offences once released from prison [3]. It is plausible that traffickers experience their activities as going from risk situation to risk situation and therefore perceive their chances of success in the same way. Third, the risk perception function, $W(P_I) = (P_0/P_I)^{\alpha}$, with $W(P_I) = 1.0$ for

$P_I \leq P_0$, arises from the psychophysics of perception processes. For example, subjects asked to match various light intensities to sound intensities or to a numerical scale produced a power-law relationship between the two physical intensities [4]. Matching two scales works well as it does not require subjects to adapt a common scale; it allows them to select their own numbers and a scale is derived from the ratios of the numbers chosen. One might expect that smugglers would match their perceptions of the penalties of being caught and suffering various consequences against the probabilities of being caught in a similar power-law relationship.

When the interview data were compared with each of the three candidate functions, neither the Pareto nor the attrition filter function was even approximately correct; however, the risk perception function appeared to represent the data quite well. Because the response data were grouped into probability of interception intervals (bins), whereas the willingness function was a continuous cumulative distribution, differences in the cumulative willingness function between the upper and lower bin boundaries were taken to estimate the fraction of inmates who responded within the range of each bin. The fitting procedure minimizes the sum of squares of deviations between observed counts and model predictions. The last two bins, 0.4 to 0.8 and 0.8 to 1.0,* were combined so that all sample sizes were greater than four. As shown in table 2, it was determined that one common exponent parameter, α, and a different P_0 for each data subset gave excellent agreement (excluding the troublesome "associate, convicted" combination discussed previously) [3].

Table 2. Parameters of estimated willingness function

	Self			*Associate*		
Parameter	*Imprisoned*	*Convicted*	*Caught*	*Imprisoned*	*Convicted*	*Caught*
P_0	0.021	0.032	0.041	0.054	0.068	0.078
Standard error	0.004	0.005	0.005	0.005	0.004	0.004
Common exponent	–1.029					
Standard error	0.068					

Interpretation of model parameters

The willingness function was interpreted to imply that smugglers entirely ignore some small probability, P_0, of being apprehended. This small probability is a threshold degree of risk of apprehension that must be exceeded before smugglers are deterred. Once this threshold has been exceeded, however, only

*The data in the 0.8 to 1.0 interval represent both the intended responses for 0.8 to 1.0 and those for greater than 1.0; in other words, those willing to smuggle knowing they would be intercepted.

somewhat larger probabilities of apprehension deter relatively large numbers of would-be smugglers. This threshold is the point at which deterrence sets in and is a strongly non-linear feature of the willingness function and, hence, of deterrence in general.

The willingness function also implies that a non-zero fraction, $P_0^{1.029}$, of the smuggler population would be willing to smuggle even if they knew that they would fail, that is, when $P_I = 1.0$. Inmates volunteered comments that validated this surprising result. Some said that the advance bonuses they received provided for their families better than the wages they could have earned in the period of time they were in prison.

Additional support for the willingness function

Other risk-taking contexts were examined to determine whether the form of the willingness function held generally. Only two contexts with comparable data were found: early automobile use and United States Coast Guard enforcement of restrictions on fisheries (allowed areas for fishing and fishing practices).

Although driving an automobile was a legal activity in 1900, it was essentially an "extreme sport", with a 1:200 chance of dying for every eight hours of driving time; however, only 1 in 6,600 Americans drove at that time [5]. By 1910, approximately 50 times more people were driving and the fatality rate per hour had dropped to one fiftieth of its former value. This is an inverse power relationship identical to the drug smugglers' willingness function established by this research.

The United States Coast Guard provided data on the frequency of boarding to inspect fishing vessels, on the number of violations of fisheries' restrictions and on the number of fishing vessels in the area during inspections [6]. For two fishing districts with ample data, it was possible to compute the decline in the proportion of fishing crews violating restrictions and the proportion of fishing boats being boarded. The relationship between these two proportions follows a simple inverse power comparable to the earlier drug smuggling findings. These data also corroborated the unexpected feature of the willingness function, that there is a residual proportion that is not deterred even by very high probabilities of being intercepted. For instance, typically less than 10 per cent ever violate fishing restrictions, but 1 per cent continue to violate even when the probability of being apprehended is as high as 80 per cent.

Deterrence of cocaine traffickers

Several operations mounted against cocaine traffickers were able to achieve a significant level of deterrence. The data from those operations provide information about real trafficker behaviour and an opportunity to validate the form of the willingness function and calibrate its parameters.

Trafficker flights over the Caribbean into the United States

Before 1989, counter-drug pilots could only intercept about 5 per cent of the trafficker flights bringing drugs into the United States from South America, and at most another 15 per cent were deterred [7]. Between 1991 and 1996, however, the capability to detect air traffickers improved significantly after the operational deployment of two relocatable over-the-horizon radars. With the radars' wide area of coverage, detection efficiency increased to the point that about three quarters of all trafficker flights were observable. Consequently, interception rates increased over this period by between 20 and 35 per cent and attempts to fly cocaine across the Caribbean declined from about 350-400 per year to about 80-100 per year. Combined with the interceptions, this meant that successful drug trafficking flights were at less than 20 per cent of former levels.

Operation "Frontier Shield" against go-fast boat drug smuggling into Puerto Rico

The 90-day "pulse phase"* of Operation "Frontier Shield", mounted by the United States Coast Guard, and the companion Operation "Gateway", mounted by the United States Customs Service, began on 1 October 1996 and were designed to intercept non-commercial trafficking into Puerto Rico and the eastern Caribbean [7]. The United States Customs Service imposed severe penalties for non-compliance with customs inspections, consisting of the seizure of boats or aircraft, thereby significantly increasing the costs of illegal entry and smuggling into Puerto Rico. During the pulse phase, there were 1,251 targets of interest, 648 boardings were carried out by the Coast Guard, 7 vessels were seized, 19 arrests were made and 6 metric tons of cocaine were seized [8]. Overall, seizure rates for all forms of trafficking increased from 10-15 per cent before the operation to nearly 30 per cent during the pulse phase, which is consistent with having doubled the interception forces. During this period, trafficking in the eastern Caribbean dropped sharply, from 38 per cent of total traffic down to 23 per cent, and it continued to drop thereafter. The willingness to smuggle is somewhat less than double these percentages because initially nearly 60 per cent of all drug traffic traversed the eastern Caribbean. Thus, the combination of deterrence and interceptions did not quite make successful trafficking drop below 20 per cent.

Air-bridge-denial operations

The strongest evidence for deterrence is the extensive operational data collected during several operations mounted in the 1990s against trafficker flights

---

*Pulse operations begin with a concentration of counter-drug effort to deter traffickers. The belief is that the effects of this disruption of organized trafficking will last long after the concentration of effort has subsided to a more typical level of effort.

transporting cocaine base from Peru to Colombia. Calibration of the full deterrence model requires independent estimates of P_I and W for each period of interest.* The Narcotics Affairs Section of the United States Embassy in Peru obtained reports on almost all flights over the air bridge and verified the reports of air interceptions. The ratio of interceptions to flights is an estimate of P_I. For the value of W, the following data sources were used: *(a)* estimates of the amount of cocaine being smuggled on each observed flight, supplied by the Tactical Analysis Team of the United States Embassy, and *(b)* the estimate of all cocaine produced based on satellite crop surveys of total Peruvian coca production, carried out by the United States Crime and Narcotics Center. Then an estimate of W is the ratio of all cocaine flown to all available to be flown. During 1993-1994, a period of stable counter-drug air interception effort, the estimates made by the Tactical Analysis Team accounted for nearly 90 per cent of estimated Peruvian production. Also during this 16-month period, air trafficking surged on three occasions to more than double its minimum, yet the observed amount flown during the first and second halves of this period balanced within 1.5 per cent [3]. This argues for the consistency and near completeness of the flight observations because analysts could not have adjusted monthly reports to satisfy end-of-period balances. Therefore, the drop from this unimpeded trafficking period to the stable periods following the implementation of the force-down or shoot-down policy period provides a quantitative estimate of the impact of deterrence.

Three Peruvian air-bridge-denial operations and the intermediate periods provide calibration data for the willingness of cocaine smugglers to continue their flights despite facing lethal consequences [3]. Operation Support Justice III** began in September 1991 in Colombia and in November 1991 in Peru and ended on 29 April 1992. This early operation was quite effective in reducing traffic because Peruvian Air Force pilots often fired upon aircraft being used by traffickers. However, lacking adequate procedural constraints, the Peruvian Air Force mistakenly fired on a C-130 aircraft operated by the Government of the United States, killing one crew member. The Government of the United States ended its detection and monitoring support for the operation shortly afterwards.

Operation Support Justice IV began in November 1992 in Colombia and in January 1993 in Peru and ended on 1 May 1994. The Government of the United States resumed detection and monitoring support, with tight restrictions on engagements to prevent fatalities. Although coca prices dropped significantly, traffic in illicit drugs continued unabated, with almost all of the available

---

*Monthly reports of interceptions were grouped into statistically significant operational periods (each period had four or more interceptions) and all sources of detection bias were examined to estimate the plausible range of uncertainty. As these adjustments and uncertainty ranges did not significantly affect the result, they have not been explained in the present article [3].

**Operations Support Justice I and II were short training exercises. Although Operation Support Justice II was a notable success, with one interception made and 42 aircraft seized, and resulted in a small dip in coca base prices, it lasted less than two months and cannot be called a major operation.

cocaine base transported. The operation ended when the Government of the United States withdrew its support because Colombia wanted to reinstate lethal engagements.

The intermediate periods before and between the various Support Justice operations did not involve lethal interception, except for the period between the end of Operation Support Justice IV and the implementation of the force-down or shoot-down policy. During this latter period with poor detection efficiency, only a few interceptions were made. The force-down or shoot-down policy was implemented after both the United States Congress and President approved support for lethal engagements under prescribed operational conditions. The policy was implemented in January 1995 in Colombia, but only effectively implemented in Peru in March 1995. The operation was an immediate success: illicit traffic plummeted for the following seven months and coca base prices dropped well below the point at which growers could break even. Operations continued throughout 1999, with only a brief pause in December 1995. By 1999, coca cultivation in Peru had shrunk to 34 per cent of its former level. Although it seems remarkable that the pilots of the aircraft carrying the illicit drugs did not surrender to Peruvian interceptors and accept their chances in court or jail when confronting almost certain death in the air, it has been shown that people are willing to accept 1,000 times as much risk when they are in control as when someone else is controlling the risk [5].

Other events in Colombia and Peru

Raids of Colombian cocaine hydrochloride laboratories were carried out by the Colombian National Police in December 1996 and January 1997. The raids destroyed a large complex of cocaine laboratories and diminished demand for cocaine base from Peru for several months, compounding the effect of the force-down or shoot-down policy.

While there were many noteworthy secondary events in the 1990s, none of them can be credibly argued to be the primary cause of significant damage to the coca market in Peru [3]. In September 1992, five months after Operation Support Justice III ended, the Government of Peru captured Abimael Guzman, the leader of the Sendero Luminoso (Shining Path) guerrillas. Thereafter, the influence of the Sendero Luminoso plummeted and government control within the growing areas increased, but the level of trafficking remained unchanged. In mid-1995, two months after the force-down or shoot-down policy was implemented in Peru, three principal leaders of the Cali Cartel were arrested. However, the precipitous drop in cocaine prices in Peru had already occurred in conjunction with the implementation of the force-down or shoot-down policy. Formal time series analysis has demonstrated that the air interception operations alone explain price fluctuations in the United States during the 1990s [9]. The timing of other secondary events was such that they could not be a plausible cause of trafficking decline or price and purity effects in the United States [3].

Calibration of the willingness function with real operational data

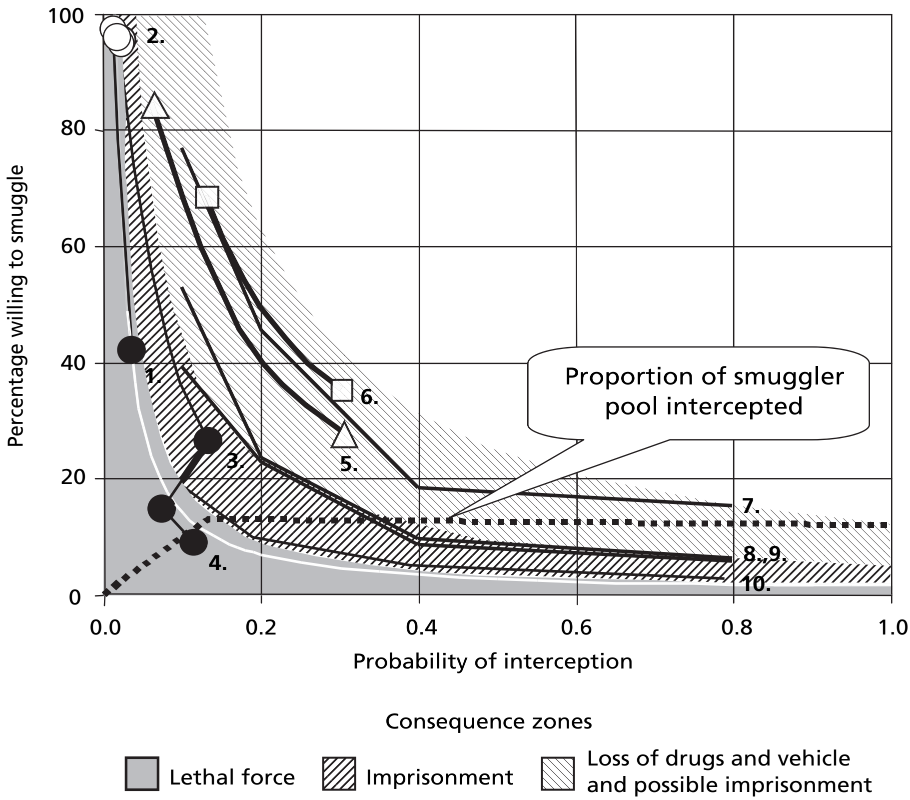
Figure II shows the willingness function overlaid with real counter-cocaine operational data as a means of calibrating the deterrence thresholds for several sets of conditions. The willingness functions for three different values of P_0 define the boundary contours for three zones of deterrence consequences, representing risk exposure to distinctly different consequences for smugglers: *(a)* lethal force, *(b)* imprisonment and *(c)* loss of drugs and vehicle and possible imprisonment.

The "lethal-force zone" is bounded on the right by a threshold parameter of 2 per cent. The estimate of the average deterrence threshold for lethal interception across all of the operational periods with potential lethal force is 1.2 per cent, which defines the willingness function contour shown within the lethal zone in figure II. Just below a deterrence threshold of 1.2 per cent, traffickers appear to ignore even the risk of lethal consequences. Above that threshold, from 1.2 to 2 per cent, increasing numbers of traffickers would stop smuggling. The "imprisonment zone" encompasses deterrence threshold values from 2 to 5 per cent, corresponding to the interval between "self, imprisoned" and "associate, imprisoned" for the smugglers interviewed in prison. Somewhere in this range of thresholds defining the zone, almost all traffickers reach their individual threshold of deterrence when facing a threat of imprisonment. Finally, the "loss of drugs and vehicle and possible imprisonment zone" encompasses deterrence thresholds from 5 to 13 per cent. This zone begins where the previous one ended and reaches nearly double the deterrence threshold for "associate, caught". This upper boundary is a practical choice, representing a limit to deterrence arising from a combination of consequences in which interception leads to capture in only some instances. For operations with these mixed outcomes and as the interception rate increases, the willingness of smugglers to continue should fall within this zone.

While the loss of vehicle or drugs alone becomes a deterrent somewhere in the remaining 13-100 per cent range of "thresholds", economic losses probably outweigh the psychological concerns as interception rates reach 30 per cent or more. Figure II does not show the boundaries of the willingness function in this range because no data are available for them and psychological deterrence provides little advantage as a force multiplier.

At the far left with circular data points, figure II shows the complex sequence of air-bridge-denial efforts by plotting the interception rates and the willingness of traffickers to smuggle for each successive operation. Black circles designate operational periods involving the threat of lethal force, while white circles represent rules of engagement involving non-lethal interception. Typically, for the non-lethal interceptions, the Peruvian Air Force strafed the smugglers' aircraft on the ground after the pilots had escaped, leaving the cocaine on board.

Figure II. Willingness function overlaid with operational data



Notes: The dates indicate the period of data consolidation for the points plotted in figure II with the following clarifications: two additional periods without lethal interception cover the periods before and after Operation Support Justice IV, which itself is split into two periods, the second beginning in October 1993. The force-down or shoot-down policy continued from its initiation to the end of the period shown, including a transition period of December 1995 to November 1996.

Air-bridge-denial operations in Peru:

1. Operation Support Justice III: November 1991 to April 1992;
2. Operation Support Justice IV: January 1993 to May 1994;
3. Force-down or shoot-down policy: March 1995 to November 1995;
4. Raids on Colombian laboratory complex: December 1996 to June 1997.

Other interception operations against cocaine smuggling:

5. Trafficker flights across the Caribbean: 1991 to 1996;
6. Operation "Frontier Shield": October 1996 to December 1996.

Inmate willingness to smuggle (cumulative) data subsets:

7. Associate, caught;
8. Associate, imprisoned;
9. Self, caught;
10. Self, imprisoned.

Interception procedures changed dramatically when the force-down or shoot-down policy was enforced. During the first nine months of implementation of the policy, only 26 per cent were willing to risk a 14 per cent chance of being intercepted. Since these percentages are averages taken over several months of an exponential decline in trafficking, the force-down or shoot-down data point in figure II does not fall precisely on the willingness function for interception with lethal consequences. Soon, smuggler pilots adjusted to the force-down or shoot-down policy because trafficking stabilized at a level well below 20 per cent during the transition period (plotted in figure II but not labelled or named). During this operational period and one including the raids on the Colombian cocaine hydrochloride laboratories, smugglers were constantly under threat of lethal consequences and the points in figure II marking these periods do match the willingness function. Figure II also shows Operation "Frontier Shield" and other operations to block trafficker flights across the Caribbean. Trafficking declined in response to the increased probability of interception for both of these operations according to the appropriate willingness function profile, that corresponding to the consequence of losing vehicles, cocaine and possibly being imprisoned if apprehended. During Operation "Frontier Shield", traffickers learned to avoid eastern Caribbean routes and have never returned in such numbers; it is believed that deterrence effects linger in the smugglers' memories.

Figure II also shows, by a dotted line, the proportion of the total potential smuggler pool that is intercepted for each level of interception probability corresponding to the deterrence model profile that defines the upper boundary of the zone with the least severe consequences. Initially, the number of interceptions rises in direct proportion to the probability of interception. But once the deterrence threshold has been exceeded, the decline in those willing to continue smuggling exactly compensates for the increased probability of interception to yield a constant number actually intercepted. This remarkable observation depends upon the fact that α, the fitted exponent in the willingness function, is statistically indistinguishable from -1.0. Why this is so cries out for explanation, but, unfortunately, leads to too complex a discussion for the present article. However, it clearly suggests that the willingness function is a general property of the psychology of extreme risk-takers.

Other aspects of deterrence

During the interviews, the drug smugglers provided answers that offered insights into other aspects of deterrence [3]. Inmates were asked how many loads a smuggler or the owner of the drugs who did not smuggle would be willing to lose before stopping or changing methods. Unlike the willingness function, the mathematical representation of their responses was an exponential decay, which implies that each subsequent loss causes the same proportion of the remaining traffickers to quit. Thus, the inmates anticipated that for each additional load lost, with associates smuggling, 27 per cent of the remaining smugglers would drop out and, with associates as owners, 24 per cent would drop out. Smugglers were also asked about their inclination to change location or method of operation as a function of perceived consequence. Their responses followed the same pattern as the willingness function, with the exception that the new thresholds were approximately half those of the former deterrence thresholds.

Finally, three important findings emerged from the analysis of drug-smuggler responses to questions on whether they would accept a greater risk of interception if offered higher fees. First, smugglers strongly discount the value of fees as the risk of interception increases. The fees necessary to induce everyone who might be willing to smuggle to do so rise in proportion to the square of the increased risk, such that, when the risk doubles, the fee required to persuade someone to smuggle increases fourfold. Second, the fees demanded by smugglers escalate as the consequences become more severe. Pilots flying cocaine from Peru to Colombia received about $20,000 per flight if they expected to be unopposed, $60,000 if they expected they might be intercepted but did not face a lethal threat and $200,000 if facing a lethal threat, but with little chance of being intercepted. If they estimated their chances of being intercepted with lethal consequences to be greater than a few per cent, few were willing to pilot aircraft transporting cocaine even when offered $2,000,000 [3,10]. Third, reconciling the inmate interview responses when they were asked, without reference to fees, if they would be willing to smuggle with their responses when they were asked about their willingness to take more risk in return for higher fees revealed the following: when simply asked about their willingness to smuggle, inmates assumed that they would be more generously compensated when they accepted a greater risk. This also suggests inmates can assess willingness to take risk in a single subjective response, independent of explicit considerations of compensation. The finding that subjects making decisions on taking risks for gains consider first whether the level of risk is acceptable and then examine the reward is just now emerging from psychology experiments carried out on college students [11].

Why deterrence matters

Although it has proved impractical to intercept or arrest more than approximately 30 per cent of those involved in the clandestine business of drug trafficking, law enforcement and other organizations attempting to control the supply of narcotics can nevertheless cause most traffickers to cease trafficking through the leverage offered by the psychology of deterrence. In the present article, several examples are given of deterrence causing a drop of 80 per cent or more in the number of successful smuggling attempts with interception rates of 30 per cent or less. If counter-cocaine activities did not cause the traffickers to cease trafficking in all cases, at least they forced them to resort to less effective and more expensive methods of trafficking.

The central contribution of the research described in the present article is the development of a mathematical expression for the psychology of deterrence, called the willingness function. This derived function is supported by a remarkably diverse set of interviews with imprisoned smugglers and calibrated by data on several counter-cocaine operations, the most detailed and complex of which were the series of operations that shut down the air routes used by traffickers to fly cocaine base from Peru to Colombia. Other data from early automobile use and the enforcement of restrictions on commercial fishing also support the mathematical formulation. In addition, the very simplicity of the willingness function suggests that it might apply to extreme risk-taking in general.*

However, the promise of effective deterrence itself poses a dilemma. The resources and effort necessary to deter traffickers can only be justified if those directing counter-drug operations have the vision and confidence that deterrence is possible. This is one purpose of the present article: to describe the evidence justifying confidence in the ability of law enforcement organizations to deter traffickers from their criminal trade. Another dilemma arises from the typical methods used for measuring counter-drug performance, which is by number of traffickers arrested and by quantity of drugs seized. Because successful deterrence causes the production and movements of drugs to abate drastically, arrests and drug seizures also decline as traffickers give up their criminal enterprise. Therefore, new measures based on the economic and organizational damage to trafficker infrastructures are needed, such as dislocations in price per pure gram, disruptions to organizations revealed from arrested traffickers and real supply shortages on the streets. Such information would provide an early warning of changes, as well as more effectively sustain the support for successful deterrence.

References

1. Joint Publication 1–02, *Department of Defense Dictionary of Military and Associated Terms* (Department of Defense, United States of America, 2000).
2. Rockwell International Special Investigations, *Measuring Deterrence: Approach and Methodology*, Contract TC-89-037 (October 1989).
3. R. Anthony, B. Crane and S. Hanson, *Deterrence Effects and Peru's Force-Down/Shoot-Down Policy: Lessons Learned for Counter-Cocaine Interdiction Operations*, Paper P-3472 (Institute for Defense Analyses, Alexandria, Virginia, United States, 2000).
4. S. S. Stevens, *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects* (New York, John Wiley and Sons, 1975).
5. C. Starr, "Benefit-Cost Studies in Sociotechnical Systems", in *Perspectives on Benefit-Risk Decision Making* (National Academy of Engineering, Washington, D.C., 1972).

*The present article does not address directly experiments on the psychology of taking business risks and related modelling approaches, such as those based on expected utility theory [11]. Other, as yet unpublished, research carried out at the Institute for Defense Analyses re-analyses these documented results and obtains regression fits of comparable or better quality using the model developed in this article. This unpublished research also proves mathematically that the general mathematical form of expected utility theory models cannot fit the willingness function that is supported by the data presented in this article.

6. B. Crane, S. Warner and M. Kuchma, *Fisheries Law Enforcement: Assessment of Deterrence*, Document D-2381 (Institute for Defense Analyses, Alexandria, Virginia, United States, 1999).
7. B. Crane, *Deterrence Effects of Operation Frontier Shield*, Paper P-3460 (Institute for Defense Analyses, Alexandria, Virginia, United States, 1999).
8. R. W. Passarelli and L. J. A. Felder, *High-Level Assessment of Operation Frontier Lance*, CRM 98-104 (Center for Naval Analyses, United States, 1998).
9. S. Soneji, R. Anthony, A. Fries and B. Crane, "Time Series Intervention Analyses on U.S. Cocaine Prices", in *Proceedings of the Fifth Annual U.S. Army Conference on Applied Statistics (1999)*, ARL-SR-110 (Aberdeen Proving Ground, Maryland, United States, 2001), pp. 67-76.
10. United States Agency for International Development, *Peru's Cocaine Economy: Estimates of the Economic Impact of Peru's Cocaine Industry in 1992* (Lima, 1993).
11. J. Sokolowska and A. Pohorille, "Models of risk and choice: challenge or danger", *Acta Psychologica*, vol. 104, No. 3 (Elsevier Science, June 2000), pp. 339-369.

A reduction in the availability of heroin in Australia

RICHARD P. MATTICK

Professor of Drug and Alcohol Studies, School of Public Health and Community Medicine, University of New South Wales, and Director of the National Drug and Alcohol Research Centre, Sydney, Australia

LIBBY TOPP

Senior Researcher, National Drug and Alcohol Research Centre, Sydney, Australia

LOUISA DEGENHARDT

Senior Lecturer, School of Public Health and Community Medicine, University of New South Wales, Sydney, Australia

ABSTRACT

After December 2000, Australia's illicit drug markets experienced an unexpected, dramatic reduction in the supply of heroin. The reduction in supply was sustained throughout 2001 in jurisdictions in which heroin had been easily available, decreasing in price and relatively pure. The heroin shortage provided an unprecedented opportunity to examine the impact of a marked reduction in the availability of the preferred drug of the majority of participants in a contemporary injecting drug market. In the present article, the authors examine the existing knowledge regarding the heroin shortage. A historical overview of Australia's heroin markets is provided, and the heroin shortage is characterized in terms of changes in the price, purity and availability of the drug. The short-term changes associated with the heroin shortage are examined, and data are presented relating to: (a) *self-reported patterns of heroin use among injecting drug users;* (b) *health outcomes, including treatment episodes and ambulance attendances at suspected opioid overdoses; and* (c) *drug-related criminal activity recorded by the Attorney-General's Department of New South Wales. The hypotheses regarding the causes of the heroin shortage are presented, and an ongoing in-depth study, funded by the Australian law enforcement sector and designed to examine the causes, effects and implications of the shortage, is described. Subsequently, the global implications of a localized and potentially short-term change in heroin availability are considered, particularly in the light of indications of a shift to other drug use among some primary heroin users. Technical challenges to understanding the causes of the heroin shortage are identified.*

Keywords: heroin, Australia, supply reduction, heroin shortage.

Introduction

Starting in late December 2000, the illicit drug markets in Australia experienced an unexpected and dramatic reduction in the supply of heroin. The reduction in supply was sustained throughout 2001 in all jurisdictions in which heroin had, for some years previously, been freely available, decreasing in price and relatively pure [1]. The reduced availability of heroin was remarkable in terms of its magnitude and prolonged nature; it provided an opportunity to examine the impact of a marked reduction in the availability of the preferred drug of the majority of participants in a contemporary illicit drug market.

The present article provides an overview of the existing knowledge regarding the heroin shortage. A brief historical background of heroin use in Australia is presented, and the heroin shortage is characterized in terms of its effects on the price, purity and availability of the drug. The short-term effects of the shortage are examined using data relating to: *(a)* changes in self-reported patterns of heroin use among injecting drug users (IDU); *(b)* health outcomes, including treatment episodes and ambulance attendances at suspected opioid overdoses; and *(c)* changes in criminal incidents for drug possession or use, as well as drug dealing or trafficking, recorded by the Bureau of Crime Statistics and Research of the Attorney-General's Department of New South Wales. The hypotheses advanced thus far regarding the causes of the heroin shortage are critically examined; they involve: *(a)* conditions in source countries; *(b)* strategic changes among heroin traffickers; and *(c)* increased success of interdiction efforts. An ongoing study designed to investigate fully the causes, effects and implications of the reduced availability is described. The global implications of a localized and potentially short-term change in heroin availability are then considered, particularly in the light of indications of a shift to other drug use among some primary heroin users. Technical challenges to understanding the causes of the heroin shortage are identified.

Monitoring trends in illicit drug markets

In recent years, there has been increased recognition of the importance of drug information systems. Both the United Nations (see the Declaration on the Guiding Principles of Drug Demand Reduction (General Assembly resolution S-20/3, annex)) and the European Monitoring Centre for Drugs and Drug Addiction [2] have highlighted the importance of making regular assessments of areas such as drug use among the general population, problematic drug use, treatment-seeking behaviour and drug-related morbidity and mortality. A number of drug information systems cover these and other areas in Australia, including national prevalence surveys, drug treatment surveys and statistics, illicit drug user surveys (including surveys of more specific population subgroups, such as youth, injecting drug users or incarcerated drug users), statistics on drug-related harm (for example, morbidity and mortality, law enforcement data on criminal activity), key informant interviews and police detainee surveys and urinalyses.

Together, the above-mentioned indicators contribute to the development of a more complete picture of illicit drug markets [3, 4]; however, because none of the indicators were established prior to the 1990s, this depth of knowledge extends back less than a decade. Thus, of the three periods in which heroin markets have undergone rapid expansion, the parameters of only the most recent have been systematically documented. The regular collection during this time of comparable assessments of key market indicators has provided a framework within which to detect trends over time in market conditions [5-8].

Australia's heroin markets in the late 1990s

The National Drug Strategy of Australia was established in 1985; governments at all levels—federal, state and territory—committed themselves to three years of supporting the Strategy [9]. From the inception of the Strategy, the importance of adopting a comprehensive, integrated approach to dealing with the harmful use of licit and illicit drugs was recognized [10]. The overall objective of the Strategy was, and remains, to minimize the harmful effects of drugs in Australian society [11, 12].

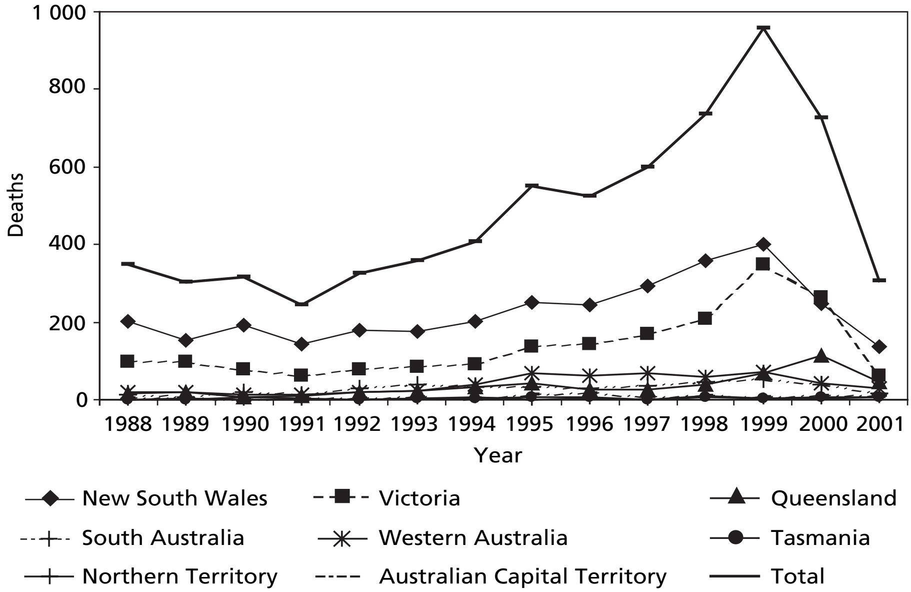
Under the current National Drug Strategic Framework 1998/99-2002/03 (which has since been extended to 2003/04) [13], the objective of the Strategy is to be realized through the implementation of programmes and policies directed at three broad areas: supply reduction; demand reduction, including abstinence-oriented interventions; and reduction of the adverse health and social consequences of drug use. The Strategy was designed not only to strengthen partnerships between the federal and state governments, but also to forge links between the health and law enforcement sectors by achieving an appropriate balance between supply reduction and demand reduction activities [14].

The increase in the number of opioid overdose deaths during the middle and late 1990s [15, 16] (see figure I) sparked concerns that illicit heroin markets were undergoing a third period of rapid growth. An important difference between the latest growth period and the earlier ones, however, is that the most recent expansion occurred within the context of vastly improved drug monitoring and information systems [3, 4]. Those drug monitoring and information systems provided a solid empirical basis from which to begin to derive an understanding of the changes that occurred.

During the period 1996-2000, the annual collection of comparable data from each of the eight jurisdictions of Australia made possible the first detailed description of illicit drug markets throughout the country. Those data indicated that, throughout the late 1990s, heroin was the drug injected most often in Australia [17]. Six jurisdictions contained viable heroin markets,* as defined by the consistent availability of the drug to regular market participants and its

*In two smaller, satellite markets in which heroin has not been readily available, the illicit use of other opioid preparations has dominated, namely methadone in Tasmania [18] and morphine in the Northern Territory [19].

Figure I. Opioid overdose deaths among persons aged 15-44 years in Australia, by jurisdiction, 1988-2001



Source: Australian Bureau of Statistics.

widespread use among them [20]. Reports regarding the preferred drugs and patterns of use of IDU suggested that, in particular, heroin dominated illicit drug markets in the south-east part of mainland Australia. That area included the two most populous states, New South Wales and Victoria. On the basis of the proportion of total opioid overdose deaths that occur within each jurisdiction (see figure I), it has been estimated that New South Wales and Victoria account for one half and one quarter, respectively, of the heroin markets in Australia [21]. Although hidden populations cannot be precisely defined, the capital cities of New South Wales (Sydney) and Victoria (Melbourne) are assumed to contain the majority of illicit drug market participants in those states. Law enforcement intelligence indicated that Sydney remained the centre of heroin importation and trafficking in Australia during the period 2000-2001 [23].

The data collected also suggested that heroin markets underwent rapid expansion during the late 1990s. In the six jurisdictions containing viable heroin markets, annual surveys of sentinel populations of IDU and key informants who worked in the field repeatedly found that the availability of heroin had remained consistently high [20]. Furthermore, the price of heroin either remained stable or decreased every year: in Sydney, it decreased by almost one half, from a median of 400 Australian dollars* per gram in 1996 to A$ 220 in 2000 [6]. The

*All prices cited in the present article refer to Australian dollars, the value of which has fluctuated recently between approximately 0.56 and 0.60 United States dollars.

average purity of heroin seized in Australia increased in the same period, from 44 per cent in the period 1996-1997 to 58 per cent in the period 1999-2000, reaching a peak of 65 per cent in the period 1998-1999 [22]. Such a pattern of increased availability, decreased price and increasing quality is consistent with an expanding market [24]. The illicit drug monitoring systems implemented in Australia in the mid-1990s were able to document the parameters of such growth for the first time.

Evidence of a reduction in the availability of heroin in 2001

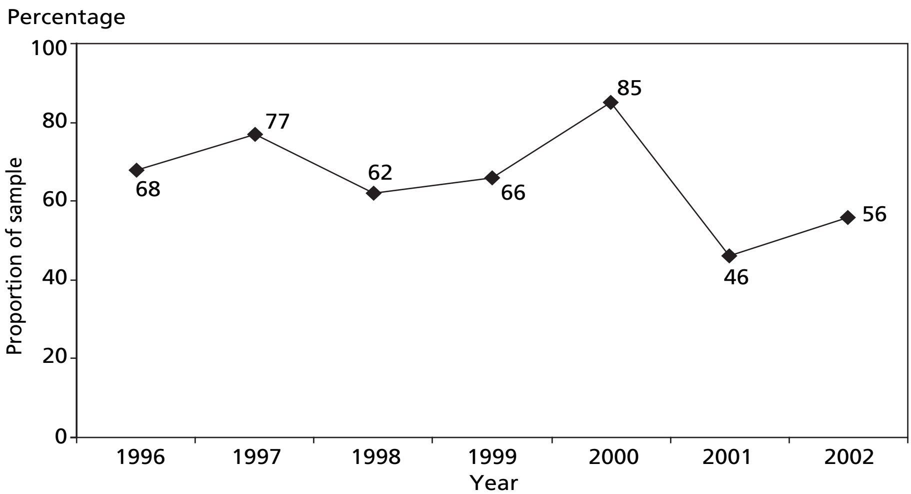
Starting at the end of 2000, however, Australia's heroin markets experienced an unexpected and dramatic reduction in the availability of heroin. Unsolicited reports of a heroin shortage came from staff of drug treatment agencies and needle and syringe exchange programmes, as well as researchers in the field. In response to those reports, timely surveys of IDU were conducted in Sydney and Melbourne to examine their veracity [25-28]. In those studies, the great majority of IDU reported a reduction in the availability of heroin, as indexed by significantly increased "search time" (the time taken to successfully obtain drugs), and marked increases in the price of heroin. Most IDU had also perceived a reduction in the purity of heroin, although figures obtained from analysed seizures of heroin are clearly a measure of purity preferable to the subjective impressions of users.

The findings were validated not long afterwards by the Illicit Drug Reporting System (IDRS), Australia's strategic early warning system that has monitored the price, purity, availability and patterns of use of illicit drugs since 1996 [29]. IDRS annually collects comparable and detailed information from a sentinel population of IDU regarding their history and patterns of drug use. The IDRS data make possible a rigourous examination of changes in patterns of drug use because a comparable sample of a sentinel population of IDU, recruited using the same methods and from the same geographical regions over time, are interviewed each year and asked questions regarding their own recent behaviour. Those data thus do not rely on retrospective recall to the extent of the studies cited above.

Consistent with recommendations regarding best practice in the monitoring of drug trends [30], IDRS also triangulates a number of data sources against the quantitative IDU survey to ensure the validity of its findings [5-8, 31]. Comparable quantitative data that demonstrate changes between 2000 and 2001 in the behaviour of IDU are presented below, but it should be noted that all of the trends thus demonstrated were validated through qualitative reports of key informants or experts who, through their work, have regular contact with illicit drug users [32]. Further, in order to be as concise and cohesive as possible, data relating only to New South Wales are presented in the remainder of this article. Data on price and availability suggest that other jurisdictions, particularly Victoria, experienced a more dramatic reduction in heroin availability than New South Wales, a fact that is consistent with law enforcement intelligence indicating that Sydney remained the centre of heroin importation and trafficking in Australia throughout the period 2000-2001 [23]. However, it is reasonable to focus on New South Wales data given that it is estimated that New South Wales accounts for approximately one half of the heroin users in Australia [33].

IDU recruited for IDRS described the availability of heroin* and recent changes in availability.** Comparison of those results over time (figure II) shows that, in New South Wales (as in all jurisdictions with a viable heroin market), the availability of heroin was dramatically reduced in 2001 compared with 2000 [32]. From 2000 to 2001, there was a marked decline (from 85 per cent to 46 per cent) in the proportion of IDU describing heroin as being "very easy" to obtain and a concomitant increase (from 1 per cent to 16 per cent) in the proportion of IDU describing heroin as being "difficult" or "very difficult" to obtain. Likewise, in 2001, a far greater proportion of the IDU sample reported that heroin had recently been more difficult to obtain (37 per cent in 2001, compared with 7 per cent in 2000) (see figure III).

Figure II. Proportion of injecting drug users in New South Wales who described heroin as being "very easy" to obtain, 1996-2002



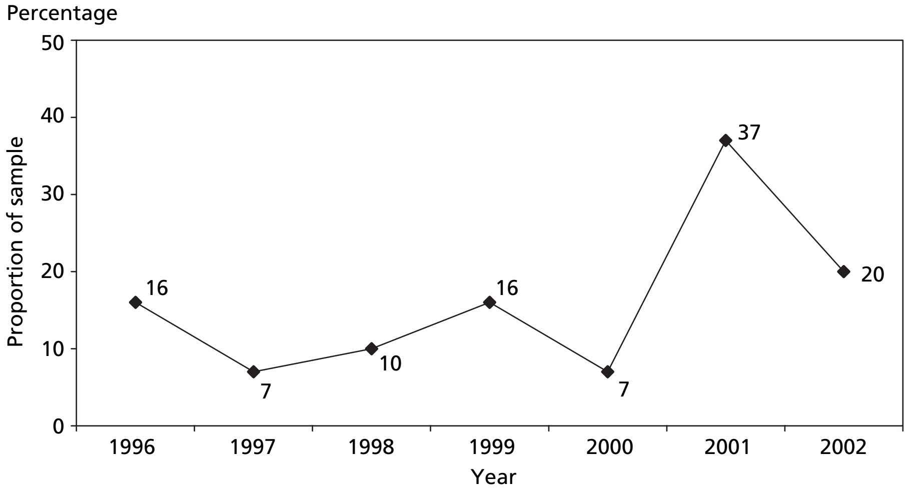
Source: Illicit Drug Reporting System, interviews with injecting drug users.

Assuming that demand is relatively stable, decreased availability should increase prices [34]. Consistent with reports by IDU of a heroin shortage in 2001 were the first increases in price recorded since 1996. Figure IV shows the median price paid between 1996 and 2002 for the last purchase of a gram of heroin among IDU in New South Wales. It also shows that, following a period of stable or decreased prices, the price of a gram of street heroin rose from A$ 220 in 2000 to A$ 320 in 2001.

*The question asked was "How easy is it to get heroin at the moment?"; the possible response options were "Very easy", "Easy", "Difficult" and "Very difficult".

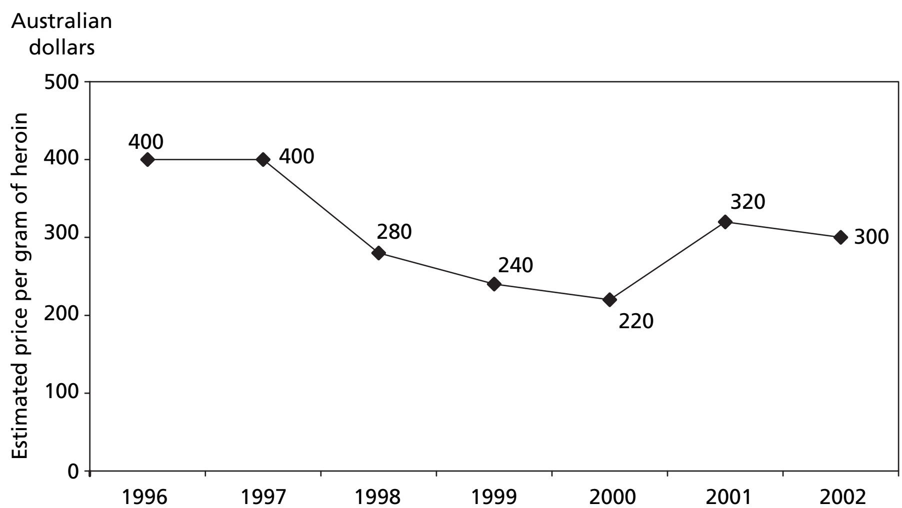
**The question asked was "Has the availability of heroin changed in the last six months?"; the possible response options were "Easier", "Stable", "More difficult" and "Fluctuates".

Figure III. Proportion of injecting drug users who reported that heroin had recently been more difficult to obtain, 1996-2002



Source: Illicit Drug Reporting System, interviews with injecting drug users.

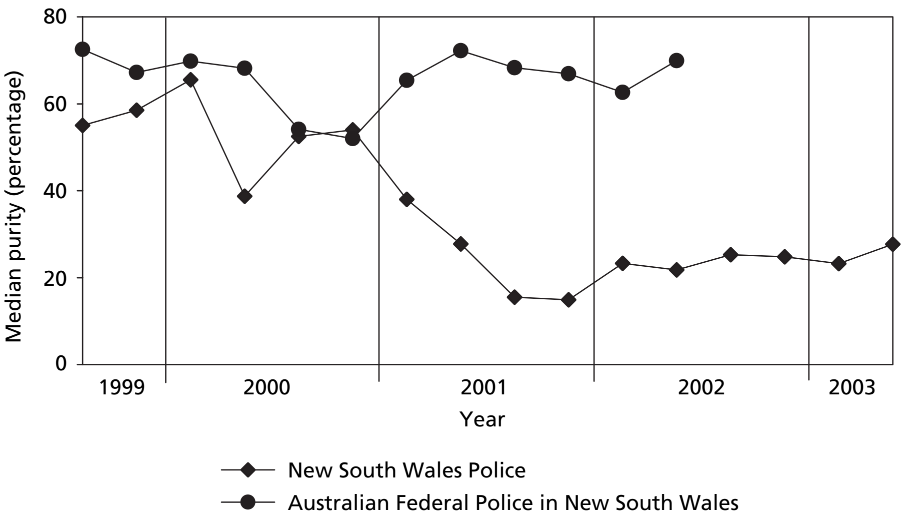
Figure IV. Estimates by injecting drug users of the price of heroin in New South Wales, 1996-2002



Source: Illicit Drug Reporting System, interviews with injecting drug users.

As would be expected during a period of reduced drug availability, the purity of heroin at the street level declined at around the time of the heroin shortage. Figure V shows the median purity of heroin seized in New South Wales by the New South Wales Police and the Australian Federal Police. Seizures made by the New South Wales Police result from activities ranging from street-level arrests of suspected users to hauls from targeted operations, whereas seizures made by the Australian Federal Police are more likely to be the result of higher-level intercepts, often at the border. The median purity of heroin seizures made in New South Wales and analysed by the Australian Federal Police has remained relatively stable over time at approximately 70 per cent (see figure V), although during the third and fourth quarters of 2000 it decreased, to approximately 50 per cent [35]. The purity of seizures analysed by the New South Wales Police, however, appears to be more variable, declining from 65 per cent in the second quarter of 2000 to 38 per cent in the third quarter of that year and dropping as low as 28 per cent in the second quarter of 2001. The New South Wales Police data beyond that date have not yet been made available.

Figure V. Purity of heroin seizures analysed in New South Wales, by quarter, 1999-2003



Source: Australian Crime Commission.

Following the 2001 findings of IDRS, the evidence appeared clear: despite reservations about the veracity of initial anecdotal reports, in 2001 illicit drug markets in New South Wales had undergone a fundamental shift in which the availability of heroin, the most widely preferred injectable drug in the state, had suddenly, dramatically and inexplicably decreased.

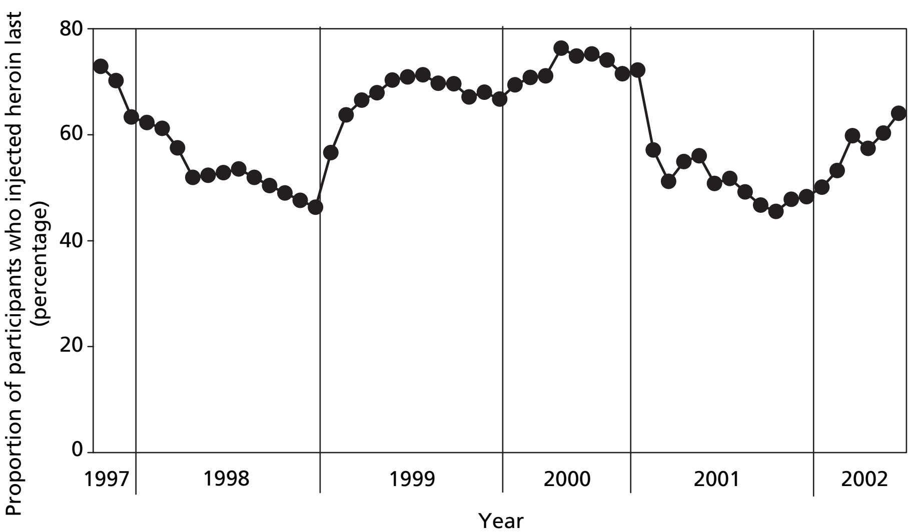
Short-term changes associated with the reduced availability of heroin

Several short-term changes associated with the reduced availability of heroin have been observed. The data presented below relate to some of those changes, discussed under three broad headings: *(a)* patterns of heroin use; *(b)* health-related outcomes; and *(c)* criminal activity.

Patterns of heroin use

As would be expected during a period of relative shortage of a drug, there were declines in a number of indicators of heroin use between 2000 and 2001. For example, among IDU recruited for the New South Wales IDRS in 2000 and 2001, the median number of days on which heroin had been used in the preceding six months among heroin users declined, from 180 days in 2000 (a figure indicating that more than one half of those who had used heroin in the six months preceding the interview had used on a daily basis during that period) to 158 days in 2001. The proportion of IDU who reported that heroin was the last drug they had injected prior to the interview decreased from 78 per cent in 2000 to 57 per cent in 2001. Similarly, the proportion of the sample who reported that heroin was the drug they had injected most often in the preceding month decreased from 79 per cent in 2000 to 58 per cent in 2001. The proportion of the sample who reported daily heroin use in the preceding six months declined from 49 per cent in 2000 to 41 per cent in 2001, and the proportion who reported that they had used heroin on the day preceding the interview fell from 78 per cent in 2000 to 62 per cent in 2001.

Figure VI. Reports of heroin being the last drug injected among persons attending a needle and syringe exchange programme in the inner city of Sidney, 1997-2002



Numerous data sources other than IDRS validate the findings of reduced prevalence and frequency of heroin use among IDU. Figure VI shows the findings of a needle and syringe exchange programme conducted in the inner city of Sydney, which routinely records the drugs last injected by participants. Following a steady increase between 1999 and late 2000 in the proportion of participants who last injected heroin, there was a marked decrease early in 2001, from 71 per cent of participants who last injected heroin in August 2000 to 51 per cent in March 2001. The consistency of a range of data clearly demonstrates that the reduced availability of heroin was associated with a reduction in the prevalence and frequency of heroin injecting among illicit drug users in New South Wales.

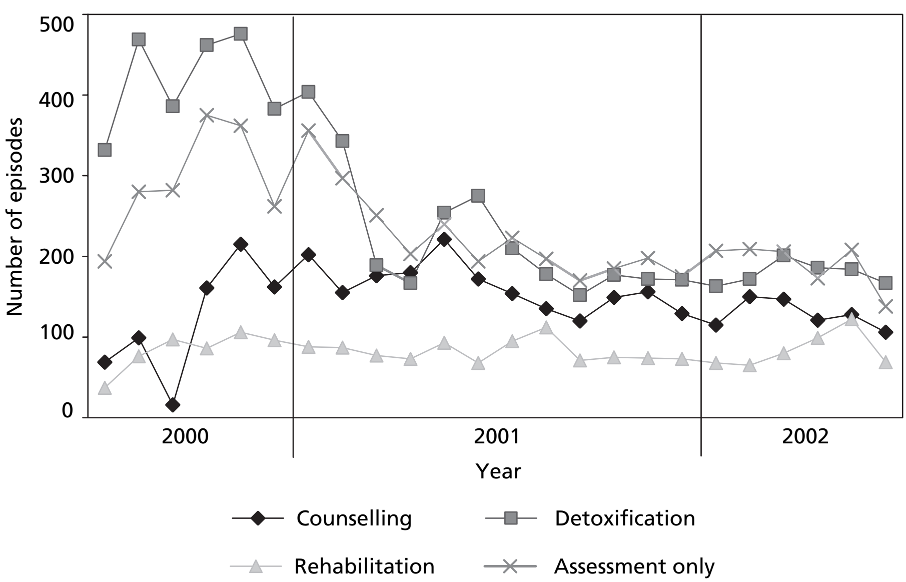
Health-related outcomes

This section presents data relating to two notable outcomes of problematic heroin use: treatment for dependence and overdose.

Treatment

Figure VII shows the number of treatment episodes in New South Wales between July 2000 and June 2002 in which opioids were identified as the main drug of concern. The number of episodes of treatment involving detoxification decreased noticeably during the first half of 2001, as did the number of assessment episodes. Overall, this has manifested itself as a reduction in the number of treatment episodes where opioids were the primary drug of concern, a reduction that was sustained until June 2002.

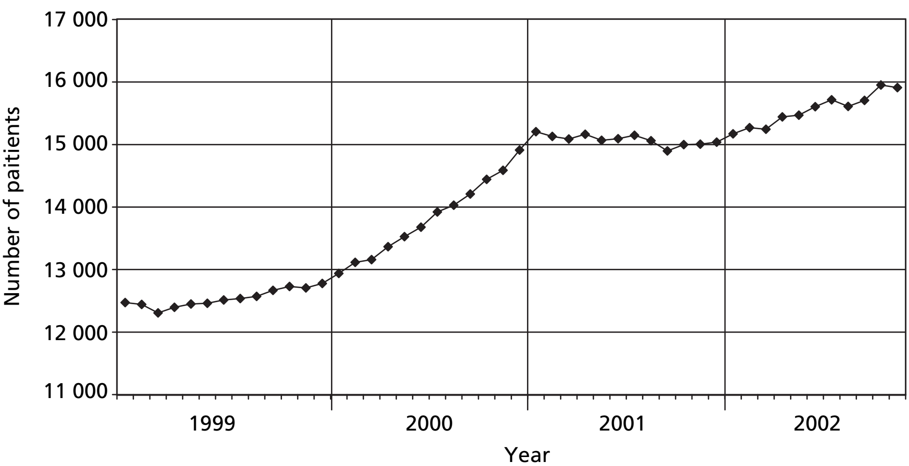
Figure VII. Opioid treatment episodes in New South Wales, by main treatment type, July 2000-June 2002



Source: Drug Programs Bureau, New South Wales Health, Australia, National Minimum Data Set.

These findings can be considered complementary to the patterns depicted in figure VIII. Between January 1999 and January 2001, consistent with an increase in the provision of treatment for opioid dependence in New South Wales, there were consistent increases in the number of persons registered for either methadone or buprenorphine maintenance treatment. Between January 2001 and January 2002, the number of persons in maintenance treatment in New South Wales remained relatively stable at around 15,000 persons. From January 2002, the number of opioid maintenance patients again began to increase; thus, in November 2002, there were nearly 16,000 persons enrolled in maintenance treatment in New South Wales. The retention of persons in treatment during the heroin shortage is notable. If the number of opioid-dependent persons has remained fairly stable since 2001 and a higher proportion are enrolled in maintenance treatment, it stands to reason that there would be a decline in demand for other kinds of treatment.

Figure VIII. Methadone and buprenorphine maintenance patients in New South Wales, January 1999-December 2002



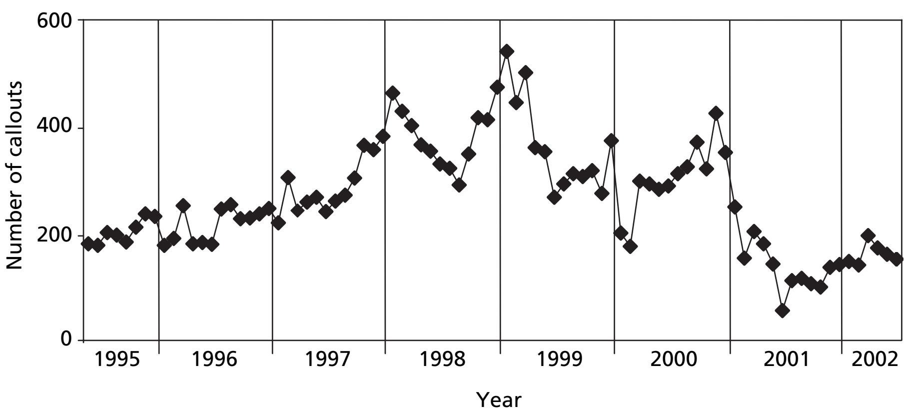
Source: Pharmaceutical Services Branch, NWS Health, Australia.

The findings relating to a decrease in the number of treatment episodes during a period of reduced drug availability may appear counter-intuitive; many suggested at the time of the heroin shortage that demand for treatment would increase. However, it may be the case that, as a result of decreased heroin purity and enforced reductions in prevalence and frequency of use, heroin users were simply not as physically dependent on the drug as they may have been during periods in which heroin was relatively abundant and, as a result, the need for treatment was not as great. As the number of people enrolled in maintenance treatment increased throughout 2002, data relating to other types of treatment episodes throughout the whole of 2002 are eagerly awaited; however, at least in the first half of the year, the number of such treatment episodes did not appear to increase.

Overdose

Figure IX shows the number of callouts by the Ambulance Service of New South Wales to suspected opioid overdoses between May 1995 and June 2002, an indicator with demonstrated sensitivity to heroin use and associated problems [36]. Starting in January 2001, there were dramatic and consistent decreases in the number of ambulance callouts per month. During June 2001, the number of callouts was lower than the number recorded in 1995. As would be expected if heroin had become less available, the number of people requiring medical assistance for opioid overdose decreased.

Figure IX. Ambulance callouts to suspected opioid overdoses in New South Wales, May 1995-June 2002



Source: Ambulance Service of New South Wales, Australia.

Summary of health-related outcomes

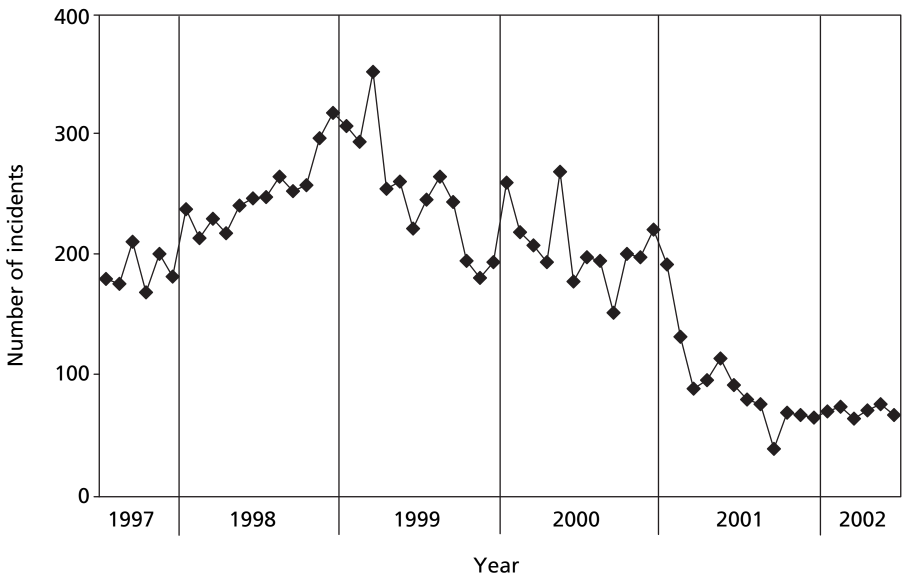
The two health-related outcomes discussed in this section suggest that there were significant health benefits to heroin users in at least some domains as a result of the heroin shortage. The demand for treatment for opioid dependence was not as great as during periods in which heroin was relatively abundant, which suggests that the severity of opioid dependence among heroin users decreased as a result of the shortage. Further, the number of persons experiencing opioid overdoses decreased. That is likely to be a consequence of decreased prevalence and frequency of heroin use and significant polydrug use, as well as decreased heroin purity [37-40].

Criminal activity

Heroin possession and use

The number of New South Wales Police incidents involving heroin possession or use that were recorded between 1997 and 2002 is shown in figure X. In early 2001, incidents involving heroin possession or use declined rapidly and stabilized at a lower level for the remainder of 2001 and at least the first part of 2002. That decrease appears to have been an extension of a steady reduction that was first recorded in the last half of 1999. There are, of course, difficulties in identifying changes in drug user behaviour based solely on police data, because such data can reflect changes in law enforcement resources and strategies, as well as changes in criminal activity. However, the data from 2001 onwards are consistent with the reports by IDU of changes in patterns of drug use, which are discussed in the preceding section.

Figure X. Incidents involving heroin possession or use recorded by New South Wales Police, 1997-2002



Source: Bureau of Crime Statistics and Research, Attorney-General's Department of New South Wales, Australia.

Summary of criminal activity

The two types of criminal activity discussed in the present section suggest that the levels of some types of criminal activity will decrease during a period of reduced drug availability; the number of incidents involving both heroin possession and use and heroin dealing and trafficking decreased during the first half of 2001, concomitant with the shortage. However, the number of incidents involving both classes of crime also appeared to decrease in the second half of 1999.

Current research into the causes, effects and implications of the reduced availability of heroin

An 18-month programme of research into the causes, effects and implications of the heroin shortage is currently being coordinated by the National Drug and Alcohol Research Centre in three Australian jurisdictions (New South Wales, Victoria and South Australia). The research, which is being funded by the National Drug Law Enforcement Research Fund of Australia, involves the following:

(a) Interviews are completed with two groups of IDU entering methadone or buprenorphine maintenance treatment: one group entering before the heroin shortage occurred, and one group entering during the period of the shortage. The IDU are compared for their current levels of functioning to determine whether there are any differences between the two groups in terms of current functioning and drug use. They are also asked about the heroin shortage, and their reasons for entering treatment when they did;

(b) Interviews are completed with key informants from both health and law enforcement agencies. The key informants have been selected on the basis of their knowledge and experience at a number of levels within the health and law enforcement bureaucracies;

(c) Analyses are completed of indicator data related to health and law enforcement consequences of drug use and of data related to seizures and law enforcement activity;

(d) An examination is carried out of protected information from law enforcement agencies at the state and federal levels regarding the nature of illicit drug markets, activity within different markets before, during and after the heroin shortage and information regarding operational law enforcement activity.

The findings of the project will be published in 2004 (see the web site of the National Drug and Alcohol Research Centre (www.ndarc.unsw.edu.au)). A number of events have been hypothesized as being the causes of the heroin shortage in Australia. Evidence to assess the support for those hypotheses is currently being systematically collected and evaluated as part of an ongoing project. More details are to be made available pending the results of research being conducted at the National Drug and Alcohol Research Centre.

Caveats

The heroin shortage in Australia has attracted extensive international attention; other countries may be somewhat envious of the reduced availability of a drug that is associated with a disproportionate amount of the total harm related to illicit drugs [21]. There is no doubt that there were benefits of the heroin shortage, particularly in terms of the health of heroin users (including a significant decrease in the rate of opioid overdose) and reductions in heroin-related criminal activity. Whatever factor or combination of factors the above-mentioned research demonstrates to have caused or contributed to the reduction in heroin availability, both the health and law enforcement sectors may justly take pride in their achievements, and the Australian Government's approach to drug issues through its National Drug Strategy, including the significant increase in funding provided through its National Illicit Drug Strategy, will again be justified.

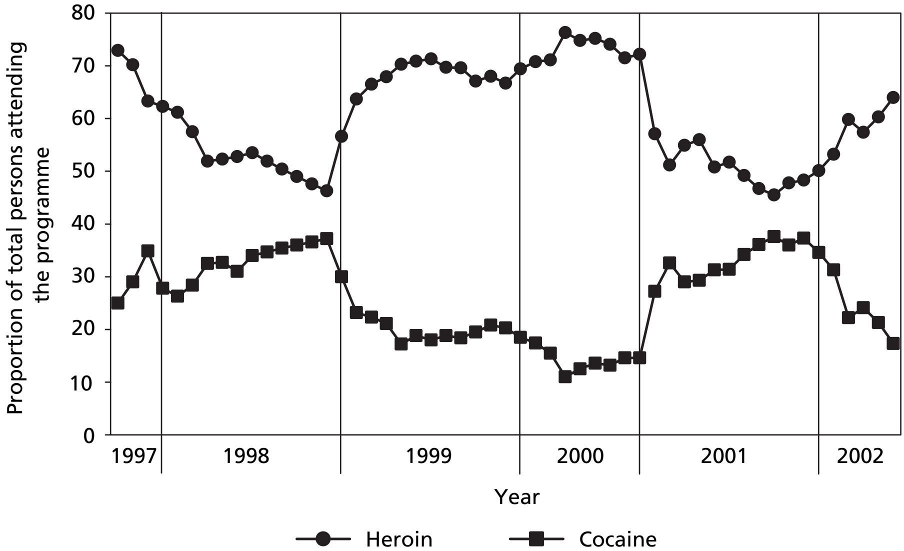
It should be kept in mind that in no other contemporary illicit drug market has the availability of heroin undergone such a sustained reduction. This is the first empirical evidence of the consequences of such a shortage. For this reason, the intuitive assumption that any reduction in drug use can be equated with a reduction in the overall level of drug-related adverse health and social consequences (the approach underpinning the National Drug Strategy) warrants careful contemplation. In this section, some perspective on the heroin shortage is provided by a consideration of three aspects to it: *(a)* evidence of a shift to patterns of other drug use among some primary heroin users; *(b)* the duration of the shortage; and *(c)* an estimation of the proportion of the total world heroin market accounted for by Australia.

Drug substitution among heroin users

Although there was a clear reduction in the prevalence and frequency of heroin use during 2001, there was some evidence of drug substitution among those who remained in the market. For example, data relating to cocaine use are considered. There was a substantial increase in the proportion of persons in the New South Wales IDRS sample reporting recent cocaine use in 2001 (84 per cent, compared with 63 per cent in 2000); and in the median number of days on which they had used it in the preceding six months (90 days in 2001, compared with 12 days in 2000). Similarly, there were marked increases in the proportion of IDU who reported that cocaine was the drug they had most recently injected (from 11 per cent in 2000 to 37 per cent in 2001) and the drug they had injected most often in the preceding month (from 9 per cent in 2000 to 34 per cent in 2001). Although some have found a shift from a depressant to a stimulant counter-intuitive, an association between heroin and cocaine use has been noted in other countries [41, 42]. The shift could perhaps have been predicted to some extent in Sydney as a result of the marked increased in 1998 in cocaine injection among primary heroin users that has been sustained since that time [5].

Data sources other than IDRS are consistent in suggesting that at least some IDU in Sydney shifted to cocaine injection in the face of reduced heroin availability. Figure XI shows the close relationship between use of the two drugs among persons attending a needle and syringe exchange programme in Sydney.

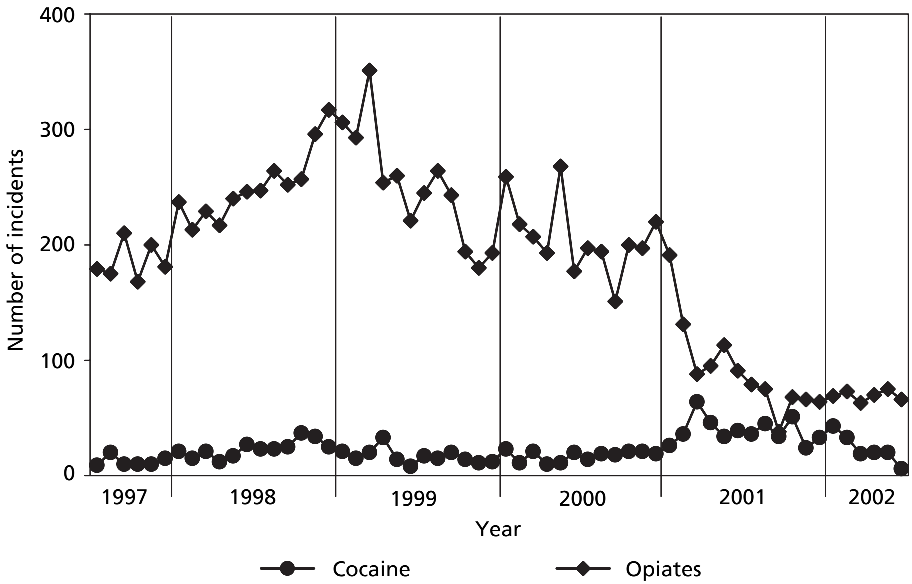
Figure XI. Reports of heroin or cocaine being the last drug injected among persons attending a needle and syringe exchange programme in the inner city of Sydney, 1997-2002



Similarly, there appears to be at least some relationship between the number of criminal incidents involving heroin and those involving cocaine (see figures XII and XIII), whereby the two types of crime appeared to mirror each other more closely from the beginning of 2001. In early 2001, incidents for heroin possession or use declined rapidly; in the same months, there was a concomitant increase in incidents related to cocaine possession or use (see figure XII). A similar pattern is evident for dealing or trafficking in cocaine: in early 2001, the number of incidents related to dealing or trafficking in cocaine increased; at the same time, the number of incidents related to dealing or trafficking in heroin decreased (see figure XIII).

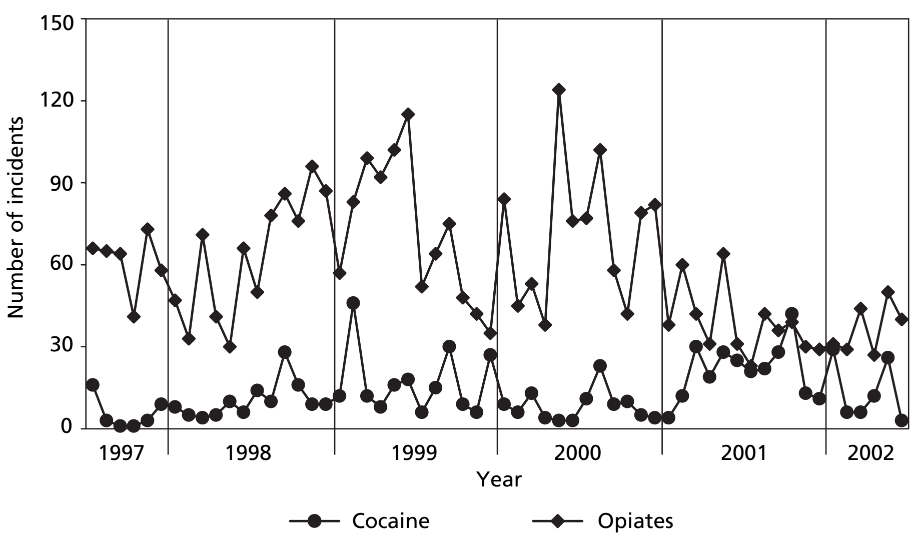
With respect to the substitution of only one drug for heroin among primary heroin users, the above-mentioned data, despite being limited to the findings of a single jurisdiction, emphasize that it is simplistic to assume that, if heroin is made less available, then there will be a consequent reduction in net drug-related harm. The 2001 national IDRS (which takes into account the results from all eight Australian jurisdictions) documented shifts to other drugs among primary heroin users in all jurisdictions that had viable heroin markets prior to the shortage. The shifts included increases in the injection of illicit methamphetamine, pharmaceutical opioid preparations, such as morphine and pethidine, and non-injectable pharmaceuticals, such as methadone syrup and benzodiazepines [32]. When assessing the benefits of the heroin shortage, it is important to take into account such shifts, particularly in Australia, where the treatment and intervention system is primarily designed to accommodate opiate use, dependence and harm [1].

Figure XII. Incidents recorded by the New South Wales Police involving possession or use of cocaine and opiates, 1997-2002



Source: Bureau of Crime Statistics and Research, Attorney-General's Department of New South Wales, Australia.

Figure XIII. Incidents recorded by the New South Wales Police involving dealing or trafficking in cocaine and opiates, 1997-2002



Source: Bureau of Crime Statistics and Research, Attorney-General's Department of New South Wales, Australia.

Such findings emphasize an important point: the overall net harm resulting from illicit drug use may not have altered to the same extent as the reduction in heroin use. The substitution of other drugs for heroin may make the net harm that occurs similar over time. For example, one consequence of the heroin shortage may have been that cocaine use became a more normative behaviour among some heroin injectors. Both Australian [5, 43] and international [44-49] research has consistently documented higher levels of harm among cocaine injectors than among other IDU. Cocaine injection is associated with a number of undesirable outcomes, including more frequent injection, higher levels of injection-related health problems, more frequent sharing of injecting equipment, greater sexual risk-taking, poorer health, higher HIV seroprevalence, increased criminality and striking rates of psychiatric comorbidity [50]. Given the consistency of the findings of huge levels of harm accruing to cocaine injectors, it must be considered that any benefits obtained from the reduction in heroin availability may be partly offset by the use of other drugs, including cocaine.

Duration of the heroin shortage

The global implications of the heroin shortage must also be considered in terms of the duration of the shortage. There is no doubt that the shortage was a reliable and valid phenomenon and that it was accompanied by increased prices, decreased purity, changes in patterns of drug use and a range of changes in health and criminal activity outcomes. Although data indicate that the availability of heroin has not returned to its pre-2001 levels, it is equally clear that the peak period of the shortage was relatively brief.

Australia's share of global heroin consumption

When reflecting on the global implications of the heroin shortage in Australia, some perspective is provided by consideration of the proportion of total world heroin consumption accounted for by the Australian heroin market. It is difficult to develop reliable estimates of heroin consumption, and any attempt to do so must inevitably be considered with an appropriate degree of caution. However, it is also reasonable to argue that it is possible to derive estimates that are at least indicative. In Global Illicit Drug Trends 2001 [51], a report prepared by the former Office for Drug Control and Crime Prevention (now called the United Nations Office on Drugs and Crime) that provides seizure figures from around the world, it is stated that, in 1999, seizures of heroin made in Oceania (Australia and New Zealand), the great majority of which were made in Australia, accounted for 1.1 per cent of total world heroin (and morphine) seizures. The total weight of heroin seized throughout the world in 1999 was 36,200.48 kilograms. In addition, it is estimated in the report that, in the same year, the interdiction rate (the quantity seized divided by the quantity produced) for opiates was 15 per cent. Together, the two figures can be used as the basis for estimating that a total of approximately 241,336.5 kilograms of heroin were produced (100/15 x 36,200.48) in 1999. If the assumption is made that the total produced is equal to the total consumed and if the upper estimate of the former National Crime Authority (which has been integrated with two other agencies to form the Australian Crime Commission) [52] of the amount of heroin consumed per annum in Australia (8,000 kilograms) is accepted as valid, it can be calculated that Australia accounts for approximately 3.3 per cent of total world heroin consumption. Alternatively, if it is assumed that seizures reflect market activity, the estimate of the United Nations Office on Drugs and Crime [53] suggests that Australia accounts for approximately 1 per cent of the global heroin market.

When Australia's proportionate role in the world heroin market is considered in this fashion, the global implications of Australia's heroin shortage can be considered within the appropriate perspective. There is no doubt that, although Australia has high rates of heroin consumption relative to countries in western Europe and the Americas [51]; with a total population of fewer than 20 million people, Australia accounts for a tiny share of the total world heroin market, the absolute upper limit being 3 per cent and a more reasonable estimate being 1 per cent.

Supply reduction, demand reduction and value for money

The monetary cost of reducing consumption by policing and customs must be viewed with the cost of successful strategies implemented in Australia to reduce demand. In Australia, 35-45 per cent of all opioid-dependent persons are in treatment at any point in time, mainly in detoxification, in pharmacotherapy treatment with methadone, buprenorphine and naltrexone and in residential rehabilitation [54]. In other countries, it is estimated that up to 80 per cent of the heroin-dependent population is in treatment. There is reason to believe that the reduction in heroin use is significant in detoxification and rehabilitation and within opioid pharmacotherapy [55]. If one third of heroin consumption is obviated by virtue of such treatment, the role of demand reduction is both marked and sustainable.

Technical challenges

The technical challenges involved in monitoring illicit drugs to obtain a better understanding of the nature of fluctuations in their availability are implicit in the arguments presented above. The challenges include developing methods to systematically document, survey or study:

(a) Local production processes and yields of various illicit drugs in source countries;

(b) The importation processes;

(c) Seizures occurring at borders and at the street level;

(d) The organization of high-level drug markets and how people enter and work in those markets;

(e) The price, purity and availability of illicit drugs.

Whether countries, especially developing countries, can develop such technical capacities will depend on the investment made.

Conclusions

Australia has succeeded in reducing the supply of heroin within its territory. The reduction does not appear to be attributable to reduced production in Afghanistan or Myanmar. It may be attributable to policing action deterring the importation of heroin into Australia, possibly by making existing importers view other markets more favourably or by reducing the number of importers through compromising the operations of trafficking syndicates. The reasons for the shortage are still being studied, and the extent to which the change in heroin supply is reliable and replicable remains to be documented.

In the face of a marked reduction in the availability of the preferred drug of the majority of IDU in Australia, the implications of the potential for fundamental shifts in the use of different illicit drugs must be considered. There is a need to increase the accessibility of interventions designed for users of illicit drugs other than heroin and to provide training for frontline workers who have limited experience dealing with such users. In Australia, governments at all levels have recognized that stimulant use, dependence and related harm have been identified as priority areas for funding and further research. The extent of stimulant-related research, as well as training for health and law enforcement professionals, currently being conducted in Australia should ensure that the treatment and intervention system is better equipped to accommodate the differential harm associated with stimulants as opposed to opiates. Whether the supply of heroin will return to its pre-2001 levels (and the impact that that may have on patterns in the use of heroin, stimulants and other drugs) remains to be seen. The heroin shortage constituted a unique natural experiment that demonstrated the importance of continued routine documentation of changes in patterns of illicit drug use so that policy makers may respond appropriately to the demands of dynamic and evolving illicit drug markets.

References

1. L. Topp, C. Day and L. Degenhardt, "Changes in patterns of drug injection concurrent with a sustained reduction in the availability of heroin in Australia", *Drug and Alcohol Dependence*, vol. 70, No. 3 (2003), pp. 275-286.
2. European Monitoring Centre for Drugs and Drug Addiction, *Annual Report on the State of the Drugs Problem in the European Union 2000* (Luxembourg, Office for Official Publications of the European Communities, 2000).

3. C. Fry and L. Topp, "Illicit drug trends in Australia", *Global Workshop on Drug Information Systems: Activities, Methods and Future Opportunities*, Vienna, 3-5 December 2001 (2002), pp. 38-41.

4. F. Shand and others, "The monitoring of drug trends in Australia", *Drug and Alcohol Review*, vol. 22, No. 1 (2003), pp. 61-72.

5. S. Darke, S. Kaye and L. Topp, "Cocaine use in New South Wales, Australia, 1996-2000: 5 year monitoring of trends in price, purity, availability and use from the illicit drug reporting system", *Drug and Alcohol Dependence*, vol. 67, No. 1 (2002), pp. 81-88.

6. S. Darke and others, "Heroin use in New South Wales, Australia, 1996-2000: 5 year monitoring of trends in price, purity, availability and use from the Illicit Drug Reporting System (IDRS)", *Addiction*, vol. 97, No. 2 (2002), pp. 179-186.

7. S. Darke, L. Topp and J. Ross, "The injection of methadone and benzodiazepines among Sydney injecting drug users 1996-2000: 5-year monitoring of trends from the Illicit Drug Reporting System", *Drug and Alcohol Review*, vol. 21, No. 1 (2002), pp. 27-32.

8. L. Topp and others, "The emergence of potent forms of methamphetamine in Sydney, Australia: a case study of the IDRS as a strategic early warning system", *Drug and Alcohol Review*, vol. 21, No. 4 (2002), pp. 341-348.

9. N. Blewett, *National Campaign against Drug Abuse: Assumptions, Arguments and Aspirations*, NCADA Monograph No. 1 (Canberra, Australian Government Publishing Service, 1987).

10. *The National Campaign against Drug Abuse 1985-1988: Evaluation and Future Directions*, NCADA Monograph No. 12 (Canberra, Australian Government Publishing Service, 1989).

11. G. Bammer and others, "Harm minimization in a prohibition context: Australia", *Annals of the American Academy of Political and Social Science*, special issue, vol. 582, 2002, pp. 80-93.

12. J. Fitzgerald and T. Sewards, *Drug Policy: the Australian Approach*, ANCD Research Paper No. 5 (Canberra, Australian National Council on Drugs, 2002).

13. Australia, Ministerial Council on Drug Strategy, *National Drug Strategic Framework 1998-99 to 2002/03: Building Partnerships: a Strategy to Reduce the Harm Caused by Drugs in Our Community* (Canberra, 1998).

14. E. Single and T. Rohl, *The National Drug Strategy: Mapping the Future: an Evaluation of the National Drug Strategy 1993-1997* (Canberra, Australian Government Publishing Services, 1997).

15. W. Hall, L. Degenhardt and M. Lynskey, "Opioid overdose mortality in Australia, 1964-1997: birth-cohort trends", *Medical Journal of Australia*, vol. 171, No. 1 (1999), pp. 34-37.

16. W. Hall, M. Lynskey and L. Degenhardt, "Trends in opiate-related deaths in the United Kingdom and Australia, 1985-1995", *Drug and Alcohol Dependence*, vol. 57, No. 3 (2000), pp. 247-254.

17. M. MacDonald, M. Robotin, and L. Topp, "Drug use trends among injecting drug users (IDU): findings from the Australian Needle and Syringe Program (NSP) Survey, 1995-2000, *Drug Trends Bulletin* (Sydney), October 2001.

18. R. Bruno and S. McLean, *Tasmanian Drug Trends 2000: Findings from the Illicit Drug Reporting System (IDRS)*, NDARC Technical Report No. 109 (Sydney, National Drug and Alcohol Research Centre, 2001).

19. B. O'Reilly and P. Rysavy, *Northern Territory Drug Trends 2000: Findings of the Illicit Drug Reporting System*, NDARC Technical Report No. 104 (Sydney, National Drug and Alcohol Research Centre, 2001).

20. S. Darke, W. Hall and L. Topp, *The Illicit Drug Reporting System (IDRS) 1996-2000*, NDARC Technical Report No. 101 (Sydney, National Drug and Alcohol Research Centre, 2000).

21. W. Hall, M. Lynskey and L. Degenhardt, *Heroin Use in Australia: Its Impact on Public Health and Public Order*, NDARC Monograph No. 42 (Sydney, National Drug and Alcohol Research Centre, 1999).

22. L. Topp and others, *Australian Drug Trends 2000: Findings of the Illicit Drug Reporting System (IDRS)*, NDARC Monograph No. 47 (Sydney, National Drug and Alcohol Research Centre, 2001).

23. Australian Bureau of Criminal Intelligence, *Australian Illicit Drug Report 2000-01* (Canberra, 2002).

24. J. P. Caulkins and P. Reuter, "What price data tell us about illicit drug markets", *Journal of Drug Issues*, vol. 28, No. 3 (1998), pp. 593-612.

25. C. Day and others, "Decreased heroin availability in Sydney in early 2001", *Addiction*, vol. 98, No. 1 (2003), pp. 93-95.

26. P. Dietze and others, "The heroin drought in Melbourne: indicators, behaviour changes, impacts and return to supply", paper presented to the 13th International Conference on the Reduction of Drug-Related Harm, Ljubljana, 3-7 March 2002.

27. P. Miller, C. Fry and P. Dietze, *A Study of the Impact of the Heroin "Drought" in Melbourne: Results of the Drug Availability Monitoring Project (DAMP)* (Melbourne, Turning Point Alcohol and Drug Centre, 2001).

28. D. Weatherburn and others, "Supply control and harm reduction: lessons from the Australian heroin 'drought'", *Addiction*, vol. 98, No. 1 (2003), pp. 83-91.

29. J. Hando and others, "The development of an early warning system to detect trends in illicit drug use in Australia: the Illicit Drug Reporting System", *Addiction Research*, vol. 6, 1998, pp. 97-113.

30. P. Griffiths and others, "Drug information systems, early warning and new drug trends: can drug monitoring systems become more sensitive to emerging trends in drug consumption?", *Substance Use and Misuse*, vol. 35, Nos. 6-8 (2000), pp. 811-844.

31. L. Topp and R. McKetin, "Supporting evidence-based policy-making: a case study of the Illicit Drug Reporting System in Australia", *Bulletin on Narcotics* (United Nations publication), vol. LV, Nos. 1 and 2 (2003), pp. 23-30.

32. L. Topp and others, *Australian Drug Trends 2001: Findings of the Illicit Drug Reporting System (IDRS)*, NDARC Monograph No. 48 (Sydney, National Drug and Alcohol Research Centre, 2002).

33. W. Hall and others, "How many dependent heroin users are there in Australia?", *Medical Journal of Australia*, vol. 173, No. 10 (2000), pp. 528-531.

34. R. J. MacCoun and P. Reuter, *Drug War Heresies: Learning from Other Vices, Times, and Places* (Cambridge, Cambridge University Press, 2001).

35. Australian Crime Commission, *Australian Illicit Drug Report 2001-02* (Canberra, 2003).

36. L. Degenhardt, W. Hall and B.-A. Adelstein, "Ambulance calls to suspected overdoses: New South Wales patterns July 1997 to June 1999", *Australian and New Zealand Journal of Public Health*, vol. 25, No. 5 (2001), pp. 447-450.

37. P. O. Coffin and others, "Opiates, cocaine and alcohol combinations in accidental drug overdose deaths in New York City, 1990-98", *Addiction*, vol. 98, No. 6 (2003), pp. 739-747.

38. S. Darke and others, "Fluctuations in heroin purity and the incidence of fatal heroin overdose", *Drug and Alcohol Dependence*, vol. 54, No. 2 (1999), pp. 155-161.

39. S. Darke and D. Zador, "Fatal heroin 'overdose': a review", *Addiction*, vol. 91, No. 12 (1996), pp. 1765-1772.

40. M. Warner-Smith and others, "Heroin overdose: causes and consequences", *Addiction*, vol. 96, No. 8 (2001), pp. 1113-1125.

41. D. M. Hartel and others, "Patterns of heroin, cocaine and speedball injection among methadone maintenance patients: 1978-1988", *Addiction Research*, vol. 4, 1996, pp. 323-340.

42. J. A. Inciardi and L. D. Harrison, eds., *Heroin in the Age of Crack Cocaine* (Thousand Oaks, California, Sage Publications, 1998).

43. I. van Beek, R. Dwyer and A. Malcolm, "Cocaine injecting: the sharp end of drug-related harm!", *Drug and Alcohol Review*, vol. 20, No. 3 (2001), pp. 333-342.

44. D. A. Bux, R. J. Lamb and M. Y. Iguchi, "Cocaine use and HIV risk behaviour in methadone maintenance patients", *Drug and Alcohol Dependence*, vol. 37, No. 1 (1995), pp. 29-35.

45. R. E. Chaisson and others, "Cocaine use and HIV infection in intravenous drug users in San Francisco", *Journal of the American Medical Association*, vol. 261, No. 4 (1989), pp. 561-565.

46. A. J. Giannini and others, "Cocaine-associated violence and relationship to route of administration", *Journal of Substance Abuse Treatment*, vol. 10, No. 1 (1993), pp. 67-69.

47. R. Hudgins, J. McCusker and A. Stoddard, "Cocaine use and risky injection and sexual behaviours", *Drug and Alcohol Dependence*, vol. 37, No. 1 (1995), pp. 7-14.

48. J. J. Platt, *Cocaine Addiction: Theory, Research and Treatment* (Cambridge, Massachusetts, Harvard University Press, 1997).

49. E. E. Schoenbaum and others, "Risk factors for human immunodeficiency virus infection in intravenous drug users", *New England Journal of Medicine*, vol. 321, No. 13 (1989), pp. 874-879.

50. C. S. Tutton and J. W. Crayton, "Current pharmacotherapies for cocaine abuse: a review", *Journal of Addictive Diseases*, vol. 12, No. 2 (1993), pp. 109-127.

51. *Global Illicit Drug Trends 2001* (United Nations publication, Sales No. E.01.XI.11).

52. Australia, National Crime Authority, *National Crime Authority Commentary 2001* (Canberra, 2001).

53. United Nations International Drug Control Programme, *Global Impact of the Ban on Opium Production in Afghanistan* (Vienna, November 2001).

54. F. Shand and R. P. Mattick, *Clients of Treatment Service Agencies: May 2001 Census Findings*, National Drug Strategy Monograph No. 47 (Canberra, Australian Government Publishing Service, 2002).

55. R. P. Mattick and W. Hall, eds., *A Treatment Outline for Approaches to Opioid Dependence: Quality Assurance Project*, National Drug Strategy Monograph No. 21 (Canberra, Australian Government Publishing Service, 1993).

Understanding the dynamics of international heroin markets: making better use of price data to measure the impact of drug control strategies

J. McCOLM

Strategic Intelligence Analyst, Her Majesty's Customs and Excise, United Kingdom of Great Britain and Northern Ireland

ABSTRACT

The present article was prepared to support work undertaken in the United Kingdom of Great Britain and Northern Ireland towards the development of a better understanding of the dynamics of international heroin markets and also to support attempts to use analyses to measure the impact of international drug control strategies and interventions by law enforcement agencies in the United Kingdom. The author explores the types of data available to analysts in the United Kingdom engaged in drug control assessment work. He also explores whether price data alone should be used to improve the current understanding of how heroin markets function. The author also provides recommendations for improving the collection and structuring of existing data. The hypotheses of existing analyses are discussed, together with the weaknesses of the underlying data. The author concludes that, for data to be used in a meaningful way, a supply-side model of international heroin supply needs to be constructed. That would enable analysts to examine the data in their proper context and would allow the data to be interpreted and communicated to policy makers in a format that would facilitate the taking of action. The author provides examples of where price data can be used in a model to influence, and measure the impact of, drug control strategies.

Keywords: heroin; United Kingdom; Afghanistan; opium; heroin prices; illicit drug markets

Introduction

Analysts working for many departments and agencies of the Government of the United Kingdom of Great Britain and Northern Ireland have been trying to improve their understanding of international markets for illicit drugs and how they function. A number of projects have been initiated in recent years to inform policy makers. Under the auspices of the Concerted Inter-Agency Drug Action group,* two projects were commissioned: one to help fill intelligence gaps on the origin of heroin destined for the United Kingdom and the other to provide details of the routes used to smuggle heroin into the country. The two projects have provided intelligence analysts in the United Kingdom with an analytical basis for making assumptions and testing theories relating to the sources and routes used by those smuggling heroin into the country.

There are still major gaps in understanding the economic characteristics of international heroin markets. Those gaps relate to the market structure of heroin supply routes and to the influences that the amount and cost of opium produced in Afghanistan, as well as the myriad relationships and payments involved in the smuggling of heroin into the United Kingdom, have on the eventual street price of heroin in the United Kingdom. International price data are required to fill those gaps. Equally important are the analysis and interpretation of the data.

Since 1998, the Concerted Inter-Agency Drug Action group has adopted an "end-to-end" disruption policy targeting the most harmful drugs (called class A drugs in the United Kingdom),** both as they arrive in the United Kingdom and on the major routes leading from source countries to the United Kingdom. Data are available on the sources of heroin found in the United Kingdom, and on the trafficking routes used. However, there are currently no data that reveal the structure of the heroin trade; such data can be assessed to indicate potential vulnerabilities. Thus, there is a need for the structured collection and analysis of price data and transaction information along the routes leading from the illicit drug production areas to the illicit markets.

The systematic gathering of price data, together with the subsequent analysis of such data, is a key tool used to assess both the desirability of competing international projects designed to curtail the flow of heroin into Western Europe and the success or failure of existing projects in relation to their impact on the price and availability of heroin. In short, it provides a mechanism that shows what works.

To analyse such data, three key conditions need to be met:

(a) There must be sufficient price data from a broad range of sources for the analysis to be statistically meaningful;

(b) There must be a supply-side model for heroin flows that can show the impact of price increases further along in the supply chain;

(c) There must be a methodology similar to the demand-side model of National Economic Research Associates [1] used to assess the size of the illicit market in the United Kingdom, to interpret the data against and to predict the potential impact in the United Kingdom of supply-side shocks.

*The Concerted Inter-Agency Drug Action group is a committee comprising the heads of all agencies in the United Kingdom involved in reducing illicit drug supply. It is responsible for steering and coordinating drug law enforcement policies and activities.

**In the United Kingdom, heroin and cocaine are the two class A drugs at which the greatest amount of drug control activity is directed.

Current price data in the United Kingdom

In the United Kingdom, the overseas drug liaison officer network provides price data to analysts based in the country. However, there are a number of problems with the data. The key problems relate to quantifying the data provided. The data are from a myriad of different sources. Factors such as the quality of the heroin, the amount referred to and the relationship between the two parties involved in the transaction are often unclear. The data are accurate when they refer to a particular transaction. However, there is never enough data to provide a suitable sample size and therefore the danger remains that a wayward or uncharacteristically deviant figure could disrupt the findings. Although the data are collated and analysed, there is not enough depth in the data for them to be used to provide "hard" assessments of the function and nature of illicit drug transactions in international markets. (That is not to say that the data have no value, as the price intelligence collected by Her Majesty's Customs and Excise is used to highlight emerging trends or to illustrate the effect of a certain operation.) Another weakness of the price data collated in the United Kingdom by the National Criminal Intelligence Service is that they represent an average of all police constabularies and are thus not weighted to reflect the level of use in any price area.

Type of data needed

The Home Office of the United Kingdom has examined the type and nature of price data needed to provide meaningful intelligence in relation to changes in the availability of illicit drugs at the street level and on wholesale markets in the United Kingdom [2]. Currently, the main use of price data in the United Kingdom is to assign a value to the weight of drugs for court sentencing. Data of the National Criminal Intelligence Service show a range of street-level and wholesale prices reported by 52 police forces and the customs authorities and an average street-level and wholesale price in the United Kingdom. However, while this is an important source of information on drug prices in the United Kingdom, with an established data collection methodology, there are problems with the use of the data. The wide range of reporting authorities and the time delay between reports mean that the data are unsuitable in their current form for providing intelligence that can be used to highlight the impact of ongoing border or upstream interdiction operations in the United Kingdom. The calculation of an average drug price over a period of 3-6 months masks any increase in price brought about by successful interdiction because the resulting changes may only last weeks or may be geographically specific.

The purity question

Data in the United Kingdom

Price is not the only factor that needs to be taken into account in a meaningful analysis. Purity should be an integral part of any attempt to measure the impact of an operation on a specific market. In the figures used by the Office of National Drug Control Policy of the United States of America to size the heroin market in the United States, heroin is calculated at 100 per cent purity; there is no relationship involving the purity of the drug in the figures of the National Criminal Intelligence Service for the United Kingdom, where what is reported is a price for heroin of unknown or variable purity.

According to research and analysis conducted by Her Majesty's Customs and Excise [3] at most points on the supply chain, drug traffickers and dealers respond to scarcity in two ways:

(a) The "Mars bar" effect: making a deal whereby the drug is sold at an agreed price, but the buyer is not aware that the amount of drug sold is actually less than the amount agreed on;

(b) Diluting the drug with more cutting agents, while holding the price at a constant level so as not to deter buyers or let competitors into the market.

That was seen in the United Kingdom in late 2001 and early 2002, when supply constriction following the ban on opium poppy cultivation issued by the Taliban in 2001 resulted in a marked fall in the purity of imported heroin seized by Her Majesty's Customs and Excise, from an average of 64.9 per cent in the first quarter of 2001 to 36.7 per cent in the second quarter of 2002, according to data of the Forensic Science Service, an agency of the Home Office. The Forensic Science Service attributed this to poorer-quality raw products and also to an increased level of cutting [4]. Some samples had been cut as many as three times before their arrival in the United Kingdom. Thus, while data from Her Majesty's Customs and Excise [3] and the Forensic Science Service [4] showed a falling purity level, street-level data showed heroin prices in the United Kingdom remaining at a constant level. This demonstrates that an analysis of the heroin market in the United Kingdom based solely on prices, without any reference to purity, can produce misleading results. A more useful and responsive measure of the illicit drug market in the United Kingdom is that of price (at the wholesale and retail levels) adjusted for purity.

International data

The same problems apply in relation to international data, although there are benefits, such as the provision of data that have been verified and analysed by staff of the United Nations Office on Drugs and Crime both at its headquarters in Vienna and in the country in question. The problem with purity remains, although forensic data suggest that it may be less of a problem in relation to the international flow of heroin in relation to the flow of heroin into the United Kingdom and within the United Kingdom. Forensic analysis of heroin shows that cutting generally occurs at or close to the production stage and at various stages between the wholesale level and the street level. The Forensic Science Service analysed samples of heroin and found that approximately 40 per cent of the samples had been cut at the source and 60 per cent had been cut at the wholesale level.* Those finding are supported by intelligence. It is safe to assume that the pharmacology of heroin destined for the United Kingdom changes very little between its manufacture and its arrival on the European mainland.

Uses of price data

There are a number of problems in using data from the overseas drug liaison officer network of Her Majesty's Customs and Excise Firstly, there is no systematic process for the collection of the data. Secondly, there are no background data indicating the nature of the deal, the size of the deal, the relationship between the individuals concerned and credit agreement or payments in kind. Those two factors mitigate against the use of the data for analytical purposes. Authorities in the United Kingdom should make greater use of price data of the United Nations Office on Drugs and Crime. Those data are both up to date and verifiable and should be used as a baseline for any supply-side model or impact assessment upstream of the United Kingdom. While a number of small data sets would suffice to provide data for impact assessments, a comprehensive data set is required to build a supply-side price model. Such a model would have to be based on data from the United Nations Office on Drugs and Crime.

That does not mean that the data from the overseas drug liaison officer network are redundant. Those data assume more importance as they provide a credibility check on the data from the United Nations Office on Drugs and Crime. They can also act as an early warning system, showing price fluctuations caused by emerging over-supply or under-supply. What is needed is a systematic method of collecting the data, along with additional qualitative data relating to size of the deal, the relationship between the individuals involved and so on. In addition, the data could show that, although a specific operation has had no impact on the price level, it has had an impact in terms of the structure of the market, in terms of how deals are brokered, or it has had an adverse effect on a particular organization.

The price data set of the United Nations Office on Drugs and Crime is currently the only one of its kind. The data have been criticized by a number of authors, most notably Reuter and Greenfield, whose criticisms mainly consists of the following [5]:

(a) Figures for consumption are based on United States prices that inflate gross estimates;

(b) Consumption estimates are biased against consumption in "poorer" countries and in favour of consumption in "rich" countries;

(c) The analysis attributes too much value added to the international supply chain, when most value added occurs between the import stage and the street level.

---

*These figures are estimates based on the main cutting agent used: if paracetamol is used, according to intelligence, the cutting took place close to the source; caffeine is generally found to have been added closer to the consumption stage at the wholesale level.

The criticism of Reuter and Greenfield is valid, especially with regard to any estimate of the worldwide value of the heroin trade or any crude assessment of gross income ascribed to the illicit drug trade within a country. However, for developing a supply-side model what is required is the ratio of price increase across national boundaries or from between import into and export from a specific country. Those data are available from the United Nations Office on Drugs and Crime [6]. In fact, Reuter and Greenfield concurred that the data of the United Nations Office on Drugs and Crime could provide useful information on, among other things, the distribution of supply chain activities and value added across countries and the distribution of final consumption across countries [5].

Modelling

Currently the only model in the United Kingdom is the demand-side model of the National Economic Research Associates. That model measures the size of the market in the United Kingdom based on the extrapolation of prevalence data. It produces a figure reflecting how much is supplied to the market from that. To calculate how much heroin is destined for the United Kingdom, analysts of Her Majesty's Customs and Excise calculate a figure for unsuccessful supply, which encompasses customs and police seizures of heroin in the United Kingdom, seizures abroad of heroin known to be destined for the United Kingdom, and a percentage of seizures of over 5 kg of heroin effected in Western Europe.* The model is also used as the basis for attributing the flow of heroin, as defined by percentage and quantity, to different modes of transport. There are problems with the model; the main problem is that there is an inverse relationship between heroin purity and the size of the market. As purity falls, it is assumed that the size of the market increases and vice versa. This could be remedied if the consumption figure was worked out at 100 per cent purity. An additional problem with consumption models is that it is difficult to ascertain the exact nature of the drug being taken. That is particularly true in countries such as the Islamic Republic of Iran, where both opium and several different grades of heroin are consumed.

The model of the National Economic Research Associates is evidence that it is possible to construct models that illustrate the behaviour of illicit drug markets. It is also possible to construct supply-side models providing adequate data are available. Supply-side models are based on two determinants: *(a)* how much drug is available; and *(b)* how it arrives. To calculate how much drug is available, a figure is needed for the quantity of opium produced and its heroin equivalent. In relation to Afghan heroin, two figures are available: one is produced by the United States Department of State and the other by the United Nations

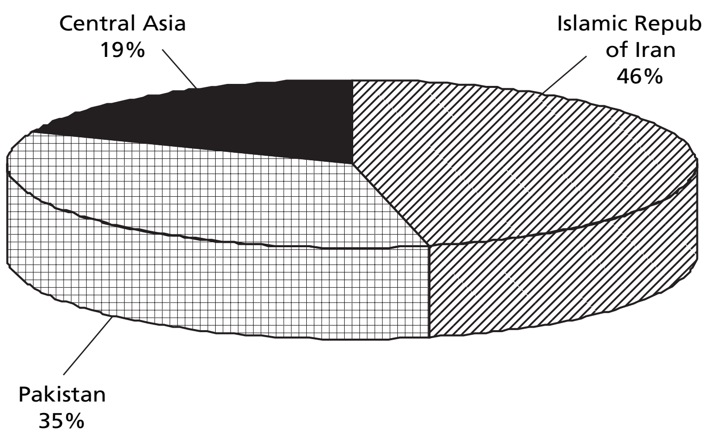
*The methodology regards Western Europe as a "single" wholesale drug market, as a consignment of 100 kg of heroin seized on the German-Polish border would not necessarily have been intended for Germany but in most cases would have been split and dispatched to Belgium, Germany and the United Kingdom. As such, 10 per cent of gross seizures of heroin in Western Europe reported to the Customs Cooperation Council (also called the World Customs Organization) and the International Criminal Police Organization (Interpol) are added.

Office on Drugs and Crime. Differing methodologies are used by the two agencies, which explains their different figures. In addition, conversion rates are based on the assumption that 10 kg of opium are required to manufacture 1 kg of heroin, despite the fact that analysis of price data and scientific research suggest that the ratio is closer to 7:1 or 6:1. However, all that is needed is a starting point that shows gross potential heroin supply.

In the opinion of the author, little is served by constructing a complex supply-side model to show the flow into and out of each country in the supply chain from Afghanistan to the United Kingdom. It is equally difficult to determine how much goes to specific countries or regions based on prevalence of use. There are data to support both, but the evidence is currently not nearly strong enough. To construct overarching models of this nature would be at best speculative.

It is possible to use models to inform policy makers and persons making decisions on international project funding, quantifying the potential risk. In 2001, the amount of heroin and morphine seized totalled 12,669 kg in the Islamic Republic of Iran, 9,492 kg in Pakistan and 5,214 kg in Central Asia as a whole. Figure I shows the seizure patterns for heroin and morphine in the Islamic Republic of Iran and Pakistan (countries bordering Afghanistan) and in Central Asia, expressed as a percentage to indicate percentage flows [6].

Figure I. Total amount of heroin and morphine seized in the Islamic Republic of Iran and Pakistan and in Central Asia, 2001



Source: United Nations Office on Drugs and Crime.

A static reading of this model indicates that, of the total amount of heroin and morphine that leaves Afghanistan, almost one half does so via the Islamic Republic of Iran, and one third does so via Pakistan. However, if the model is run over time, then any sustained rise or fall in the percentage flow via a particular country will need to be examined more closely. A similar chart produced five years earlier would have shown a much smaller percentage for Central Asia. Therefore, the increase in flow through Central Asia can be attributed to either increased flow or improved law enforcement or a combination of both. A model such as this can also be used as a baseline prior to the launching of an international project. If one of the aims of the project is to curtail the amount of heroin entering Pakistan then a better yardstick would be the amount of heroin seized in Pakistan expressed as a percentage of the total amount seized in its neighbouring countries. That figure has the advantage of not being influenced by the production level in Afghanistan or by abstracts of the domestic seizure total in Pakistan.

Models such as this are simple; they do not explain why certain routes are chosen or show where profits are made. The key to explaining this is market incentives expressed in terms of price differentials. If opium prices are low then that would force farmers to consider growing alternative crops to earn their livelihood. Falling international demand for heroin also diminishes the incentives for trafficking in opium and heroin. However, if consumer demand remains high, the market will be willing to pay high prices for opium to Afghan traders.

Traders in Afghanistan recognized that opium was one of the few commodities that could be produced in excess of local demand, and for which international demand was strong. Thus, they made a connection between Afghan opium production and international demand for heroin. By the late 1990s, the opium trade within Afghanistan did not have a risk premium, reflecting the fact that the authorities were not trying to prevent the trade. Traffickers correctly assumed that opium profits would increase rapidly once the borders of Afghanistan were crossed. Prices doubled once a consignment moved from southern Afghanistan to Pakistan, and the price of opium on the wholesale market in Tehran was six times as high as the price on the border of Pakistan [6]. It is generally assumed that the increase in the price of a consignment that has crossed the border is attributable to costs to cover such things as transport and bribes and to compensate for risk, as well as an element of speculation. What is interesting from an analytical perspective is that the increase in the price of opium as it moves across different borders is not constant.

According to Her Majesty's Customs and Excise pricing intelligence, over the past few years, wholesale heroin prices in the United Kingdom have been considerably higher than in the Netherlands, whereas there is little difference between the prices in Belgium and the Netherlands. This illustrates that there is a cost imposed on traffickers transporting heroin to the United Kingdom that is not imposed on those moving the drugs to Belgium. The additional cost is attributable to structural and transport costs imposed by moving concealed goods from a continent to an island, an element of profit and a cost to compensate for the risk posed by law enforcement interdiction. The same costs are likely to occur at all stages along transport routes leading from source countries. If it can be assumed that structural costs are constant, then increasing the costs imposed by law enforcement is likely to increase the cost of crossing a border. However, measurement of this is difficult; price data are not sufficient to illustrate any sustained impact of operations. Perhaps a better way of assessing the longer-term impact of drug control policies is to look at changes in price ratios across borders. If a project designed to improve law enforcement effectiveness on the Afghan-Tajik border resulted in the ratio of prices between Afghanistan and Tajikistan increasing from 2:1 to 4:1, then it could be deemed a success. However, if the ratio remained the same, then it would be clear that little impact other than an increase in seizures had been made.

Although not an exhaustive list of indicators, a model or group of models that illustrate flow and price differentials and are capable of showing changes over time are vital to providing a quantitative basis for both funding and measuring overseas assistance. Models can show the parts of the supply chain where costs and profits are the highest and, by implication, where traffickers are most vulnerable. Models are also better indicators of the success of operations or projects, as they measure outcome rather than output.

Usefulness to policy makers in the United Kingdom

A model of this nature can be used in the United Kingdom to plan in advance both where drug control interventions should be made and the type of interventions required to have the maximum impact. That is particularly important given the commitment of the United Kingdom Government to evidence-based policy-making and the need to pursue a drug control strategy based on what works and on value for money. Currently, funding is given in the United Kingdom for drug supply activities in areas identified as having top priority in terms of the threat posed to the country. However, there is little assessment of the type of interventions that would be the most effective in terms of meeting the government goals of reducing the availability of illicit drugs or at least in terms of increasing prices or reducing purity. An analysis based on the relative price increases between countries and within countries along the supply route leading from Afghanistan to Western Europe would be the key to this.

There are strong theoretical arguments favouring supply-side policies, which restrict the flow of narcotic drugs to consumer markets. From the prohibitionist paradigm, any restriction is per se a good thing. However, support for supply-side policies can also be seen from the harm reduction paradigm. Caulkins states that law enforcement can play a role in either micro or macro harm reduction, but the options for contributing to macro harm reduction are clearly greater [7]. International law enforcement actions need to show results, not just in terms of the amount of drugs seized, but also in terms of the impact in relation to constraining supply. For a policy such as "end-to-end" disruption to succeed, it would have to be shown that it had been successful in reducing drug-related harm. Indicators such as the size of the domestic market and street-level prices are currently being used in the United Kingdom to do that. Caulkins believes that these are valid indicators; he states that there is mounting evidence that consumption is in fact quite responsive to price [7]. In their study of the Australian heroin drought in the crop year 2000/01 Weatherburn and others indicate the relative price elasticity of demand for heroin [8]. That evidence from Australia suggests that the domestic impact of supply restriction can be measured. To support effective policy-making in the United Kingdom, a supply-side model for heroin flows needs to be constructed. Such a model would be able to indicate which policies or operations would have the greatest impact on consumption in the United Kingdom and could be used to predict how future supply restriction may have an impact in the country and further downstream.

There are other indicators of harm that are not based on harm to individuals as represented by gross consumption of heroin in the United Kingdom. Law enforcement in the United Kingdom has focused on the harm caused by drug use in specific communities in the country. More recently, this approach has focused on the harm caused by crack cocaine in many inner-city black communities in the country and on the rising levels of "black-on-black" violence. Law enforcement action has targeted body carriers of cocaine (couriers carrying cocaine on their person), both as they leave Jamaica and as they arrive at airports in the United Kingdom. This is a case where a qualitative judgement in relation to the levels of harm has been made over a quantitative one: cocaine dispatched from Jamaica may not represent the bulk of the cocaine arriving in the United Kingdom, but it has the highest number of negative externalities: associations with crime involving guns; murder of participants; predominance of crack markets; and general lawlessness. Caulkins, however, puts forward the alternative view that a ton of heroin smuggled in by body carriers is less harmful than the same amount smuggled in large shipments. He states that large shipments are worth enough to create incentives for employing violence and that any organization that moves such large quantities must be powerful, whereas a smaller dealer who arranges delivery by body carriers may have fewer resources. Ten small organizations, each importing by body carriers, may have less capacity to corrupt or threaten state institutions than one large organization that imports as much as the 10 small organizations do collectively [7]. These two views illustrate the type of qualitative analysis that needs to be carried out when assessing the nature of domestic harm.

Understanding how illicit markets function

An understanding of how the illicit market functions is imperative to planning operations and projects that have the greatest impact on *(a)* reducing the supply of narcotic drugs to the United Kingdom, *(b)* reducing the overall flow of narcotic drugs and *(c)* changes in purity and price in the illicit market in the United Kingdom. Much of the academic debate on this subject has focused on whether markets are "additive" or "multiplicative". An "additive" price structure would mean that an increase of $1,000 in the price of a kilogram of heroin as it crosses a national border would result in a similar price increase being passed on all the way down the supply chain to the user; this theory assumes that all actors—farmers, traders, traffickers and transport groups—are charged fixed prices for their goods and services. A "multiplicative" price structure means that a price increase of 10 per cent paid to cross a border would result in a series of 10 per cent price increases all the way along the supply chain; in other words, each price increase is multiplied until the drug reaches the user [6].

The above-mentioned assumptions about the nature of drugs markets have significant policy implications. If the price structure of an illicit drug market is additive, then interventions at the source, such as eradication, are close to meaningless, as any increased cost to the trafficker will not be passed on to the user. In some cases, they may be counterproductive, driving up the price for opium and stimulating the market to demand more opium production in the years that follow. If, however, the price structure is multiplicative, then interventions both at the source and in transit countries will pass on price increases (or, in some cases, reductions in purity) to the user; interventions will therefore serve to meet government targets for supply reduction. If the supply chain is totally multiplicative, then the impact of increased opium prices caused by eradication will be passed on to the user.

However, illicit markets are never as straightforward as economic theory suggests. In relation to heroin, a casual reading of price data and intelligence reveals that there are parts of the supply chain that are additive and parts that are multiplicative. Reuter points out that drug crop cultivation is a classic case of an additive price structure [9]. Such cultivation is dominated by a large number of farmers who only have a fixed number of outlets for their produce. The farmers have no market power to push up the prices that they charge. As such, the cost that an opium trader pays to the farmer for the opium is an additive cost. Intelligence indicates that, for British customers, Turkish traffickers have two different prices for a kilogram of heroin: one for delivery to the Netherlands, to be collected by the customer; and the other for delivery to the customer in the United Kingdom. As that price increase is close to the price difference in the illicit market, it is likely that an increase in that cost could be passed on; however, it will probably be reflected in an increased wholesale price and then the cost will be absorbed in the supply chain before the heroin reaches the street-level user.

Example: functioning of the illicit market in southern Afghanistan

With the Government of the United Kingdom committed to a policy of "end-to-end" disruption, it is vital that funding for overseas projects and operations targets the parts of the supply chain that are multiplicative, in order to have the maximum disruptive impact. For example, figures of the United Nations Office on Drugs and Crime show that prices for opium/heroin double between southern Afghanistan and Quetta, Pakistan [10]. Prices increase by only 30 per cent between Quetta and the port of Karachi. The first increase may be either multiplicative or additive, based on the level of risk. The increase between Quetta and Karachi would appear to be additive, reflecting fixed storage and transport costs in Pakistan. However, the price structure of the supply chain between southern Afghanistan and the Islamic Republic of Iran has multiplicative characteristics. Data from the United Nations Office on Drugs and Crime suggest that the price of opium increases sixfold once it crosses the Iranian border; in some cases, it has increased by a factor of 10 [10]. This is a reflection of the higher fixed transport cost, the increased risk of interdiction, the increased likelihood of death if apprehended and increased security costs.

Economic logic suggests that policy should focus on law enforcement on the Afghan-Iranian border, which would be more effective than similar assistance on the Afghan-Pakistan border. However, if the policy objective is to provide drug control assistance to Pakistan, based on qualitative intelligence indicating that Pakistan may be the source of up to one fifth of the heroin entering the United Kingdom,* then it would be important to examine the price structure in Pakistan.

Prices of heroin from Pakistan show that there is a very large increase, most likely of a multiplicative nature, between the Pakistan export price and the price of importing heroin from Pakistan into the United Kingdom. (The wholesale price of heroin in Pakistan is about $3,500 per kilogram while in the United Kingdom heroin imported from Pakistan can fetch around $24,000 per kilogram [11].) Economic logic suggests that, if British law enforcement activity is to have an impact on the supply of heroin from Pakistan to the United Kingdom, the area worth focusing assistance on is the part of the supply chain between Pakistan and the United Kingdom rather than the Afghan-Pakistan border.

Turkish drug traffickers: controlling prices on illicit markets

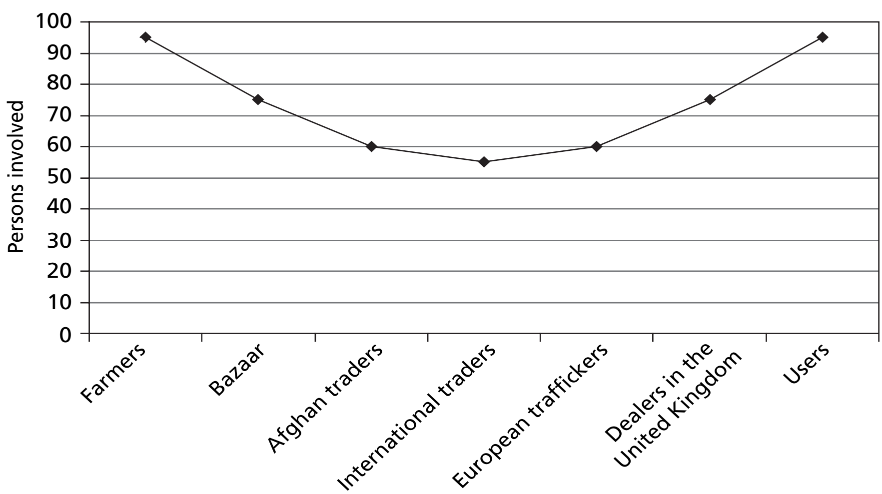
Understanding the pricing structure of the heroin supply is the key to finding out where overseas assistance should be directed. However, there is also a need to understand the nature of relationships between traffickers and whether those involved in the heroin trade act in accordance with or counter to economic logic. The main impact in the United Kingdom of the 2001 ban on opium poppy cultivation in Afghanistan was a reduction in the purity of imported heroin. Intelligence indicated that, although there were large increases in the price of morphine base and heroin on arrival in Turkey, it was only followed by some small increases in the wholesale price between Turkey and Western Europe. Turkish traffickers appeared to prefer to hold prices constant. Analysis has shown a correlation between falling purity levels in Turkey and the United Kingdom [11]; however, there may be other factors at play. It could have occurred for a number of reasons; for example, perhaps traffickers did not want to lose customers to rivals; that the close, trust-based relationship between traffickers and buyers meant that traffickers had to hold the price that had been offered. Analysis presented to the Concerted Inter-Agency Drug Action group in

*In July 2002, The Forensic Science Service reported that analysis of samples of seized heroin revealed that 27 per cent of the samples had originated in Pakistan. Pakistan is one of the principal countries of origin of heroin seized in the United Kingdom; however, in the United Kingdom, heroin from Pakistan is usually found among air passengers and freight and the total weight of such seizures is considerably less than that of seizures of heroin arriving via Turkey and mainland Europe.

2002 suggested that increases in the price of opiates in Afghanistan and the Islamic Republic of Iran were not passed on to users in the United Kingdom because of the key role played in international markets by Turkish drug traffickers. That theory suggests that any price increase upstream was absorbed by the key Turkish trafficking groups and not passed on to others downstream.

Turkish drug traffickers are positioned between Afghan farmers and traders and European wholesalers and users. They have established upstream and downstream links and have considerable control over the processing of opium and morphine base into heroin. Whether the illicit goods physically move through Turkey or not, Turkish traffickers hold a dominant market position, dealing with many suppliers and buyers who are dependent on them. It is estimated that approximately 80 per cent of heroin destined for Western Europe passes through Turkey [6]. That means that traffickers based in Turkey are in a position to set prices. In terms of building a price-based model to show market behaviour, it can be assumed that the price of heroin in Turkey is subject to "cartel" or oligopolistic behaviour. Turkish traffickers are able to control the illicit market because they are at the mid-point of international markets and, consequently, are in the best position to exert oligopolistic pressure to increase prices, there being a small number of actors involved at this stage in the supply chain.

Figure II. Number of persons involved in the heroin trade



According to the Concerted Inter-Agency Drug Action group, a conservative estimate of the price differential between heroin entering and leaving Turkey would be 400 per cent. That is heroin is sold to European buyers at four times the price paid for it (or for morphine, taking into account the conversion costs). This illustrates that the maximum profit with regard to international heroin trafficking from South-West Asia to Western Europe is made in Turkey. This is not to be confused with the maximum profit from trafficking or dealing in heroin, which the majority of research shows is made in the consumer country. The limited intelligence that Her Majesty's Customs and Excise has received indicates that, as the heroin consignments are moved from Turkey across Europe, costs return to being additive based on a fixed sum charged by transport groups to cross various national borders, which, in turn, is based on the level of law enforcement capability and the sum of the bribes to be paid.

With respect to the cost increases following the 2001 ban on opium poppy cultivation in Afghanistan, Turkish traffickers chose to respond to what they deemed short-term increases in their raw material costs by absorbing the increased cost instead of passing it on to their buyers in Europe. That approach runs counter to the multiplicative thesis. However, it is economically rational behaviour, particularly if the cost increases were considered to be temporary. The high profits made from the trade in Turkey enabled the traffickers to absorb the cost increases. Perhaps it was decided that it would be better to keep existing customers happy by delivering an agreed amount of heroin at agreed prices (albeit with a reduced purity level) than to face the possibility of raising the prices and then perhaps losing customers to alternative sources of supply. It is worth pointing out that other strategies, such as passing on the cost increase, would have to be pursued if the upstream cost increases were to continue beyond the short term.

Conclusion

It can be assumed that drug trafficking is a profit-driven business that functions in accordance with conventional economic rationale. If law enforcement is to be successful in its goal of reducing illicit drug use, then it needs to concentrate its efforts on finding out what works. A better understanding of how illicit drug markets function is vital. The author has highlighted existing sources of data and ways to interpret them in order to provide better insight into the operation of illicit drug markets. Analysts in the United Kingdom should not function in a "black box"; much analysis in this area is currently being undertaken by the United Nations Office on Drugs and Crime and by analysts and academics in other countries. There are considerable theoretical and empirical data that can be analysed to provide the basis for a model of the heroin supply in the United Kingdom.

Only by developing models of supply can law enforcement in the United Kingdom assess the type of drug control interventions that will have the maximum effect. Such models could be used to direct government resources to regions and countries where they are likely to have the greatest impact in terms of increasing the costs to drug traffickers and reducing their profit margins. The models can also be used by analysts to assess the impact over time of the interventions, particularly in terms of reducing the use, reducing the flow or increasing the prices of illicit drugs, as well as increasing the costs to drug traffickers. Quantification of the impact of upstream intervention has long been a problem for the Concerted Inter-Agency Drug Action group. The suggestions made in the present article offer a way forward, so that "outcome", as well as "output", can be measured.

References

1. Edward Bramley-Harker, *Sizing the UK Market for Illicit Drugs*, RDS Occasional Paper No. 74.
2. Government of the United Kingdom of Great Britain and Northern Ireland, Home Office, "Towards a price target for drugs", unpublished paper, 2003.
3. Jim McColm, "Towards a future heroin model: measuring the impact of supply-side interventions", unpublished paper, October 2002.
4. Government of the United Kingdom of Great Britain and Northern Ireland, Home Office, Forensic Science Service, "Heroin Impurity Profiling Report", unpublished paper, July 2002.
5. Peter Reuter and Victoria Greenfield, "Measuring global drug markets: how good are the numbers and why should we care about them?", *World Economics,* vol. 2, No. 4 (October-December 2001), pp. 159-173.
6. *The Opium Economy in Afghanistan*, (United Nations publication, Sales No. E.03.XI.13).
7. Jonathan P. Caulkins, "Law enforcement's role in a harm reduction regime", *Crime and Justice Bulletin* (Sydney), No. 64, January 2002.
8. Don Weatherburn and others, "The Australian heroin drought and its implications for drug policy", *Crime and Justice Bulletin* (Sydney), No. 59, October 2001.
9. Peter Reuter, "Comments on 'Sizing the UK market for illicit drugs'", unpublished manuscript.
10. "Greed, warlords and the opium trade", *The Opium Economy in Afghanistan* (United Nations publication, Sales No. E.03.XI.13).
11. *Global Illicit Drug Trends 2002* (United Nations publication, Sales No. E.02.XI.9), p. 33.

Price-setting behaviour in the heroin market

THOMAS PIETSCHMANN
Research Officer, Research and Analysis Section,
United Nations Office on Drugs and Crime, Vienna, Austria

ABSTRACT

The present article analyses price-setting behaviour in drug markets, starting with an analysis of the forces driving drug markets ("expected profit") and those restraining their expansion ("risks"), distribution patterns, the involvement of organized crime and the importance of market concentrations, before addressing in more detail the question of price-setting behaviour along the trafficking chain. Examples are drawn primarily from the opium/heroin markets, as they have experienced the most significant price changes in recent years, thus making possible the testing of two different hypotheses: additive versus multiplicative price-setting. On the basis of empirical analysis, the article argues that neither the additive nor the multiplicative model are fair reflections of reality. Following the ban of opium poppy cultivation by Afghanistan in 2001, the empirically determined rise in the price of heroin in the consumer markets of Western Europe (taking changes of purity into account) was less than the rise predicted by the multiplicative price model but significantly higher than that foreseen by the additive price model.

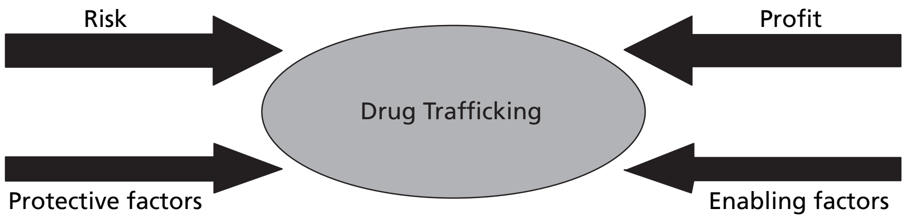
Keywords: drug markets, heroin markets, drug prices, heroin prices, purity, price-setting behaviour, risk, profit, intervention, distribution pattern, organized crime, opium, heroin, opium poppy ban, Afghanistan, Iran (Islamic Republic of), Pakistan, Tajikistan, Western Europe, France, Germany, Italy, United Kingdom of Great Britain and Northern Ireland, additive model, multiplicative model.

Overview

The present article is a slightly modified and updated version of a paper prepared previously by the United Nations Office on Drugs and Crime, entitled "Factors to be taken into account in modelling drug trafficking operations", which was presented at the United States Library of Congress in January 2003. It starts with some general considerations on the operations of illicit drug markets, such as profits and risk, distribution patterns, the involvement of organized crime and the importance of market concentrations, before addressing in more detail—based on empirical observations—the question of price-setting behaviour along the trafficking chain. Though the analysis is not restricted to any specific drug, examples are drawn primarily from the opium/heroin markets as they have experienced the most significant price changes in recent years, thus enabling the testing of two different hypotheses (the additive versus the multiplicative price-setting model).*

Drug trafficking, in conceptual terms, can be analysed by two key factors: profits and risk, as well as a number of enabling and protective factors (see figure I). Protective factors include what is usually described under the heading of "social capital" (community cohesion, usually fostered by local traditions, local culture, religion, local employment opportunities and so on). Enabling factors are, inter alia, ethnic diaspora, marginalization of significant proportions of the population, unemployment, uneven income distribution (relative poverty), lack of government control and civil strife. In addition, location plays a key role (transit countries).

Figure I. Conceptual representation of drug trafficking



The key motivation for participation in drug trafficking is usually profit, that is, the expectation of high returns in a relatively short period of time for relatively little work. The higher the profit expectations, the more drug traffickers, as well as people in general, will try to participate in such business operations. A specific question, which will be addressed below in more detail, is whether the return on investment, or profit per unit of substance trafficked, is the key driving force.

The main limiting factor for drug trafficking is risk. The higher the risk, the less, *ceteris paribus*, drug trafficking will take place. This will be true as long as the risk is not offset by growing profit expectations. The "risk function" differs widely among individuals and groups in society. For a few individuals risk can even be the driving force for their activities, though for the majority of people it is the factor that limits their potential involvement in the drug business. In general, males appear to be less risk-averse than females. The readiness to take risks is also stronger among youths and young adults than among the older

*The "additive" and "multiplicative" price transmission models were first introduced—to the knowledge of UNODC—by Caulkins [1] and discussed further by him [2, 3]. Similarly to the findings of this article, Caulkins sees these models as ends of a spectrum with the likely situation being some blend of the two extremes.

population. Risk is also a class phenomenon: people who have good career prospects due to their social origin and their educational background tend to be—for obvious reasons—more risk-averse than those whose economic prospects are, in any case, minimal to non-existent. For similar reasons, risk perceptions differ among people from different countries, often reflecting the levels of development.

In any case, much research is still needed on what governs the risk function and on the impact of various interventions on risk perceptions. Preliminary research in the late 1990s by the Institute for Defense Analyses of the United States of America suggested—on the basis of interviews with apprehended drug traffickers—that the impact of seizures on drug trafficking was relatively small, unless really huge amounts were taken out of the market. Otherwise, seizures were seen as merely the "cost of doing business". If the goal of deterrence was to reduce the operations of potential drug traffickers by 80 per cent or more, about 80 per cent of drugs in circulation had to be seized, or 20 per cent of the drug traffickers had to be arrested and imprisoned (i.e. more than 1 in 5), or drug traffickers had to believe that there was a chance of more than 1 in 25 (4 per cent) to risk their lives in drug trafficking operations. (This referred, inter alia, to pilots transporting coca base from Peru to Colombia, fearing to be shot down by the authorities.) An average trafficker was found to consider giving up trafficking only after having lost about four loads in a row.

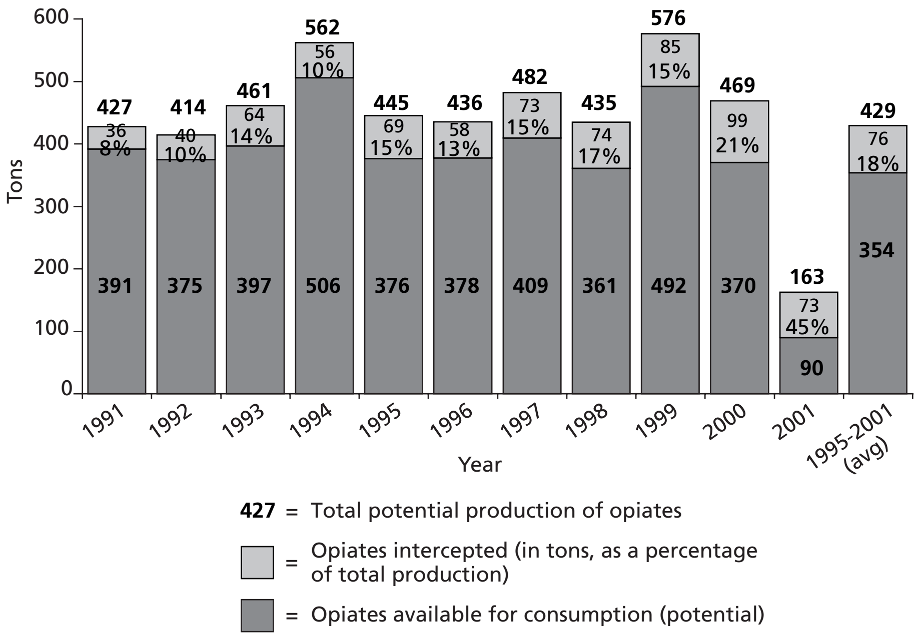
The level of "perceived risk" tends to be more important than "actual risk". Thus, mass media, for instance, seem to play an important role in shaping the risk function. The deterrence thresholds, which prompt the "average" trafficker to respond to perceived risks (e.g. by temporarily ceasing operations or looking for alternative routes) start, according to this model, if more than 30 per cent of the drugs are being seized, if a chance of 2-4 per cent exists to be personally arrested and imprisoned or if there is a risk of near to or less than 1 per cent to die in trafficking operations [4].

As long as risk and profit expectations are in equilibrium, drug trafficking will remain stable. The profit/risk equilibrium, however, can be broken. An increase in law enforcement activities, for instance, will raise risk. Increased levels of risk will, *ceteris paribus*, reduce drug trafficking, notably once the thresholds levels—outlined above—have been crossed.

As regards empirical information to determine whether such threshold levels are reached, in the case of cocaine, production estimates of the United Nations Office on Drugs and Crime [5] (or the United States Government) [6] and global seizure data (United Nations Office on Drugs and Crime) [7] suggest that 40 per cent of the potentially available cocaine was seized over the period 1995-2001. Over the last two decades the interception rates for cocaine increased from 37 per cent of the cocaine produced over the period 1986-1995 to 45 per cent over the period 2001-2002. This looks quite impressive, but unfortunately does not mean that the necessary threshold levels for cocaine traffickers to give up their activities have been reached. Such calculated interception rates have an inherent upward bias as they are not adjusted for changes in purity along the trafficking chain. As drugs move closer to the end consumer they tend to be diluted. The actual psychoactive content of cocaine usually declines from levels around 90 per cent in the producer country to some 60 per cent in the main consumer countries. Thus, the actual interception rates were at around 30 per cent over the period 2001-2002 [8]. Calculations of a global interception rate are based on aggregate seizures covering all levels (producer countries, transit countries and seizures in the consumer countries). Each level of trafficking is characterized by the existence of additional layers (e.g. importer, large-scale wholesaler, medium-market distributors and street-sellers). All of this means that at any individual layer along the trafficking chain the actual interception rate is significantly below a threshold level of 30 per cent.

If the interception rates at the global level for opiates are calculated, it can be seen that on average 18 per cent of the opiates produced were seized over the period 1995-2001 [9], up from 11 per cent over the period 1990-1995 (see figure II). In 2002 the interception rate amounted to 19 per cent. Based on the considerations discussed above, such seizure levels may be significant, but they still fall short of being a deterrent to trafficking. In the case of other drugs, interception rates are even lower.

Figure II. Global illicit supply of opiates, 1991-2001
(in tons of heroin equivalent)

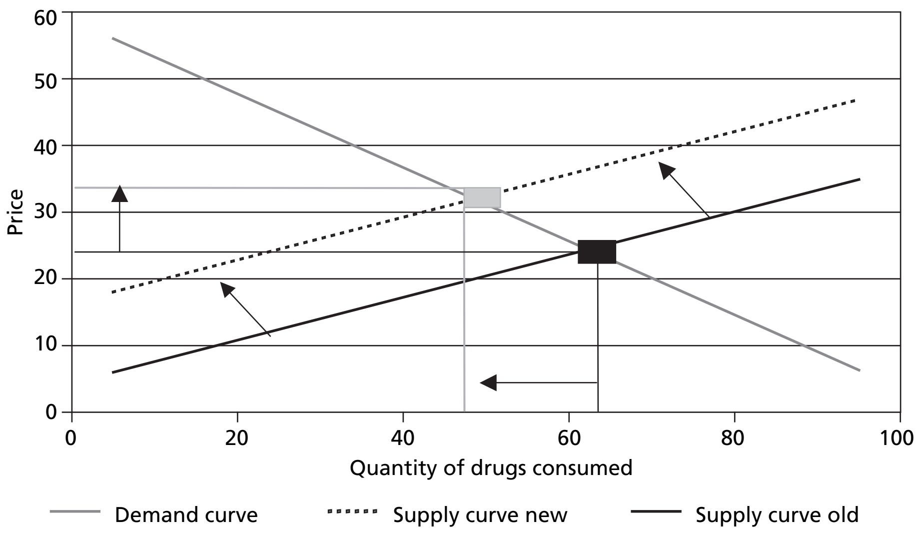


Source: Global Illicit Drug Trends 2003 (United Nations publication, Sales No. E.03.XI.5).

Thus, a key problem is that most enforcement interventions targeting drug trafficking take place below threshold levels that would prompt a change in traffickers' behaviour. Yet this does not mean that supply-side interventions are systematically failing: in fact they work indirectly. Increased levels of risk translate into higher drug prices, compensating for the higher risk. The higher drug prices have a positive impact on consumption in the sense that drug consumption, in general, is "price-sensitive" though not necessarily "price-elastic".* Higher drug prices, *ceteris paribus*, tend to lower drug consumption [14]. (Similarly, lower drug prices tend to raise consumption.) Thus, the overall amount of drugs consumed (and therefore "required" to be trafficked) can be expected to fall once prices increase (see figure III).

This is the case as long as the new risk-profit equilibrium remains intact under the changed conditions. However, the higher profit margins are a constant incentive for new participants, who have a lower degree of "risk-awareness", to enter the market. ("Illegal aliens" and marginalized youth, including children of second-generation immigrants, are particularly vulnerable in this regard.)

Figure III. Supply reduction intervention



*This is an important distinction: drug prices are considered "elastic" if a 1 per cent increase in prices would result in a more than 1 per cent decline in consumption; otherwise they are considered "inelastic". According to estimates provided by the Institute for Defense Analyses, cocaine use price-elasticity amounted to –0.63 (Drug Abuse Warning Network), –0.29 (Drug Use Forecasting), –0.6 (Smith Kline Beecham Chemical Laboratories), –0.38 (Treatment Episode Data Set), so that the authors proceeded with an assumed price-elasticity of –0.5 [10]. Caulkins, based on a review of the literature, arrives at substantially larger elasticities for cocaine use, ranging from –0.72 to –2.0 [11]. There is also some evidence suggesting that similar price-elasticities apply to heroin as well. One study carried out in three Swiss towns in the late 1990s revealed a heroin price-elasticity of -0.7 [12]. A previous study carried out by the National Bureau of Economic Research found for the United States price-elasticities of –0.9 for heroin and –0.55 for cocaine [13].

The elimination of a (potentially violent) drug trafficking network may well inspire other groups to fill the vacuum. This is a potential dilemma for law enforcement. Successful interventions increase risk and result in higher drug prices and thus larger profit margins. The net result is less trafficking in the short term. This does not however exclude the possibility that the supply curve will—eventually—fall back to its original position prior to the intervention, as other criminal groups become interested in the strong incentives for involvement in the drug market. The supply curve may even fall below its original position if a well structured trafficking network, which held a local monopoly, is removed from the market. In some cases, perceived risk is larger from rival drug trafficking gangs than from the police and the elimination of a violent group may well lower overall risk to operate in such a market, thus attracting many more groups to enter the market. This will increase competition as well as the overall supply of the market. (This seems to have been the case once the powerful Medellín cartel and then the Cali cartel were dismantled and many more smaller groups filled the gap [15].)

One could assume from this that law enforcement interventions are futile as they do not guarantee a long-term shift in the supply curve, but this is definitely not the case. Even temporary reductions in overall drug consumption are positive. If they do not occur, there is a potential danger of the development of a self-reinforcing drug epidemic. Once demand for certain drugs becomes "trendy", fuelled by sufficient supply, more and more people will promote the consumption of such drugs to finance their addiction. Drug consumption can thus grow exponentially, promoted by thousands of small-scale traffickers who actively contribute to the growth of the market, leading to a full-scale drug epidemic.

The analysis also reveals that one-time law enforcement actions will not be enough to ensure change. Constant efforts are required to keep the supply curve at the new equilibrium point of higher prices and thus less consumption. The model also suggests that any gradual increase in enforcement interventions, moving in parallel with growing profit margins, will not necessarily lead to an end of drug trafficking. The deterrent effects—as long as they are below the threshold levels—are then offset by the larger incentives reflected in the profit margins. The only limitation comes from the demand curve. Higher prices, prompted by the larger profits, will limit consumption.

Drug-specific characteristics

Regional characteristics

Geographical location is an important determinant of drug price. Prices usually rise along the trafficking chain from producer country to consumer country. Moreover, drug trafficking locations are often characterized by important spillovers. They tend to lower the retail drug prices in the respective markets, thus enabling a local consumer market to emerge.

The main indicator for such trafficking activities are seizures, which take place primarily in the transit countries in and around the main countries of production, as well as in traditional consumer countries of North America and Western Europe [16]. Differences in the global trafficking pattern can be explained, first of all, by differences in location and consumption. The main trafficking route of cocaine leads from the Andean countries to the United States; the main heroin trafficking routes from Afghanistan to Western Europe. Both heroin and cocaine trafficking are thus to a large extent interregional.

There are also important differences in the distribution patterns. While most cocaine is produced in Colombia and criminal groups from Colombia also play a key role in the world's largest cocaine consumer market (notably the east coast of the United States; the west coast and the states close to the southern border appear to be dominated by Mexican groups), the heroin trade is far more fragmented. Opium is typically produced by Pashtun villages in Afghanistan (and to a lesser extent by Tajik villages in the north). Opium is then sold at the local bazaars and shipped by traders to the borders or transported to local laboratories where it is processed into heroin. Heroin produced in northern Afghanistan usually leaves the country via Tajikistan. Once across the border, Tajik groups take over the business and sell the drugs across the region of the Commonwealth of Independent States, notably to the Russian Federation. Opium/heroin produced in eastern Afghanistan is trafficked across the border by Pashtun traders. In the case of opium produced in southern Afghanistan, specialized Baluchi traders take over once the opium has been shipped to the borders and ship the opium, morphine or heroin to the Islamic Republic of Iran, either directly or via Pakistan. Once in the western part of the Islamic Republic of Iran, Kurdish groups take over and ship the morphine/heroin across the border to Turkey. From eastern Turkey, opiates are usually shipped to Istanbul. Turkish/Kurdish groups and, increasingly, Albanian groups then transport the heroin to Western Europe, sometimes subcontracting local East Europeans [17]. Several depots exist in Eastern Europe. Many of the bulk deliveries head towards the Netherlands, from where they are then further distributed across Western Europe. In addition, there are some direct deliveries of heroin from Pakistan to the United Kingdom of Great Britain and Northern Irealnd [18], which may also explain the overall higher levels of purity found in the United Kingdom of Great Britain and Northern Ireland compared with those in the rest of Western Europe. The actual sales at the street level have in recent years been taken over increasingly by West Africans and, in some locations, by North Africans [19].

A different distribution system is typical for amphetamine-type stimulants (ATS). In this case, most trafficking is intraregional. Interregional trafficking is limited to the supply of the precursor chemicals [20]. (One exception here is methylenedioxymethamphetamine (MDMA, commonly known as Ecstasy), which in recent years has gained popularity outside Europe, though most production still takes place within Europe.) ATS do not need to be imported into Europe or the United States, but they are to a significant degree locally produced. ATS laboratories exist primarily in the Netherlands and neighbouring Belgium (often operated by groups from the Netherlands), as well as, to a lesser extent, in practically all other West European countries and increasingly in East European countries as well. Production of ATS in Europe is focused on amphetamine and MDMA. ATS manufacture in North America, by contrast, is dominated by methamphetamine, which is also the main ATS produced in East and South-East Asia [19].

With regard to trafficking in cannabis herb, there is a focus on Mexico, the United States and Africa, though practically all countries are affected by trafficking in and abuse of cannabis [21]. Trafficking in cannabis resin concerns mainly Europe (notably Spain), as well as Morocco and Pakistan [22].

Involvement of organized crime

The more organized crime is involved in drug trafficking, the higher are the overall dangers for society. The risks usually emerge from drug-related violence and the financial power of such criminal groups to corrupt the economic and the political system, leading to a crowding out of legitimate investment, inefficient capital allocation and eventually to lower economic growth and poorer living conditions. Ironically, a strong involvement of organized crime will, in general, lead to rather high drug prices (owing to reduced competition) and will thus actually contribute to lower levels of drug consumption than under competitive market conditions.

Based on qualitative information (such as intelligence reports but also press reports), the strongest involvement appears to be in the cocaine trade, followed by trade in opiates. This is also confirmed indirectly once drug seizure cases are analysed. The larger the seizure, the more efforts are likely to have gone into planning and financing such a drug trafficking operation.* Based on this indicator, the largest involvement of organized crime can be found for cocaine, followed by opiates and ATS, notably methamphetamine.

In the case of cocaine, the joint United Nations Office on Drugs and Crime/International Criminal Police Organization (Interpol)/Customs Cooperation Council (also known as the World Customs Organization) database on significant seizures shows a total of 36 cases equivalent to 1 ton or more (going up to 21 tons) for 2001. By contrast, there were only 8 seizures of opiates exceeding 1 ton reported in 2001 (5 seizures of opium, 2 seizures of heroin and 1 seizure of morphine) and 20 such cases (15 involving opium, 3 morphine and 2 heroin) in 2000. (All of the large opium seizures took place in the Islamic Republic of Iran and the morphine seizures in the Islamic Republic of Iran and Pakistan.) In the case of ATS, there were only four cases of methamphetamine seizures exceeding 1 ton in 2001 (all of them in China). No individual seizure case exceeding 1 ton was reported for either cannabis herb or cannabis resin for 2001 [23].

According to replies to the annual reports questionnaire of the United Nations Office on Drugs and Crime, the average size of a cocaine seizure amounted to 3.9 kilograms (kg) at the global level in 2001 and was thus significantly larger than the average heroin seizure case (0.17 kg) or the average size of a seizure of ATS (0.07 kg) in 2001 (including MDMA) [19].

---

*Of course, one should not exclude the possibility that highly sophisticated trafficking networks could also resort to a large number of small deliveries in order to improve their overall risk management.

Another possibility (see the table below) is to calculate the ratio of overall reported total seizure cases (T) to significant seizure cases (S). This is based on the assumption that the involvement of organized crime in "significant" seizure cases, reported by Member States to the United Nations Office on Drugs and Crime, Interpol or the World Customs Organization, is, in general, more important than for all seizure cases. The lower the ratio T:S, that is the higher the proportion of significant seizures (S) in total seizure cases (T), the higher the likelihood that organized crime is involved in such business activity. Using that indicator, the highest involvement of organized crime at the global level appears to be in trafficking cocaine, followed by heroin, cannabis resin, cannabis herb, amphetamines and MDMA.

Reported seizure cases, 2000 and 2001

	Cocaine		*Heroin*		*Cannabis resin*		*Cannabis herb*		*Amphetamines*		*MDMA*
	2000	*2001*	*2000*	*2001*	*2000*	*2001*	*2000*	*2001*	*2000*	*2001*	*2000*
T[a]	58 229	77 633	182 405	170 482	293 915	194 826	275 790	230 865	226 299	79 768	71 765S
S[b]	5 433	6 533	3 131	2 310	3 036	2 876	1 595	1 975	624	557	207
Ratio T:S	11		65		83		142		259		347

[a]Average total seizure cases (T) of 2000 and 2001.

[b]Average significant seizure cases (S) of 2000 and 2001.

Sources: United Nations Office on Drugs and Crime, annual reports questionnaire data and United Nations Office on Drugs and Crime/Interpol/Customs Cooperation Council significant seizures database.

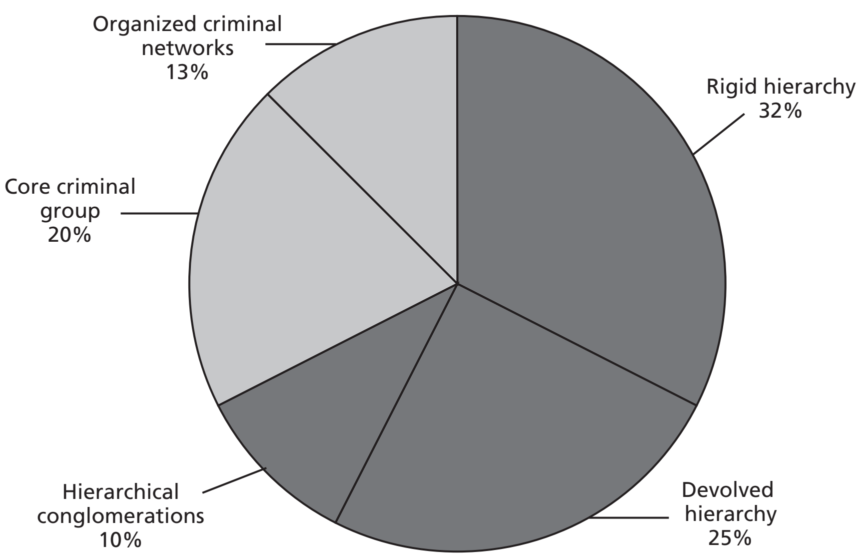
Not only the involvement of organized crime as such, but also the actual structure of the groups involved plays an important role in the price-setting behaviour of the players in a drug market. In a study conducted by the United Nations Office on Drugs and Crime in 2002, 40 transnational organized criminal groups, located on all continents,* were analysed with regard to their activities, structure, size, transborder operations and (ethnic and social) identity, violence, corruption,

*The study was based on information collected from Australia, Canada, China, Colombia, the Czech Republic, Germany, Italy, Japan, the Netherlands, the Russian Federation, South Africa, the United Kingdom, the United States, the Caribbean region, as well as on information on Albania, Belarus, Bulgaria, Lithuania and Ukraine obtained from a similar study by the United Nations Interregional Crime and Justice Research Institute (UNICRI). The 40 transnational organized criminal groups investigated in detail were located in Australia, Bulgaria, Canada, China, Colombia, Germany, Italy, Japan, Lithuania, Mexico, the Netherlands, the Russian Federation, South Africa, Ukraine, the United States and the Caribbean region. The following institutions and agencies were involved in the research process: the Australian Institute of Criminology, the Canadian Anti-Organized Crime Division, Ernst and Young in the Netherlands Antilles, the Universidad Nacional de Bogotá, the Institute of Criminology and Social Prevention in Prague, the German Bundeskriminalamt in Wiesbaden, the Direzione Centrale della Polizia Criminal in Rome, the National Police Academy in Tokyo, the Research and Documentation Centre of the Ministry of Justice in the Hague, the Academy for International Cooperation of the Ministry of the Interior in Moscow, the Institute of Security Studies in Cape Town, South Africa, the National Criminal Intelligence Service in London and the National Institute of Justice in Washington, D.C. UNICRI partners in the study were the University of Tirana, the European Humanities University of Belarus, the Varna Free University of Bulgaria, the Law Institute of the Ministry of Justice of Lithuania and the National Academy of Sciences of Ukraine.

political influence, penetration into the legitimate economy and cooperation with other organized criminal groups [24]. The study found, inter alia, that drug trafficking was still the key business activity for more than half of the organized criminal groups investigated. The drug trafficking activities of such groups were centred around cocaine, heroin, cannabis, methamphetamine and amphetamine. For a significant number of organized groups in which drug trafficking was not the key activity, it played, nonetheless, a role for overall income generation. Forty-three per cent of the organized criminal groups were reported to have been involved in just one key activity (which in almost 60 per cent of these cases was drug trafficking), though this did not exclude that a number of related criminal sub-activities were also carried out (e.g. money-laundering). A quarter of criminal groups operated in two or three crime sectors. Among organized criminal groups for which multiple key activities (more than three) were reported (a third of the groups), about 60 per cent were active in drug trafficking [25].

Though there seems to be a general trend towards smaller, looser and less hierarchically organized groups (one third were loosely organized in forms of core groups or networks), two thirds of the criminal groups studied still had a classical hierarchical-type structure (see figure IV and box). Such classical hierarchical structures—in combination with the actual size of such criminal groups—allow such groups to be active price-setters, while the more loosely organized forms of core groups and networks are usually "price-takers" (i.e. they simply react to the incentives provided via the drug prices in the market).

Figure IV. Structure of transnational organized criminal groups in 16 countries (N = 40)



Source: Centre for International Crime Prevention, "Assessing transnational organized crime: results of a pilot study of 40 selected organized criminal groups in 16 countries", *Trends in Organized Crime*, vol. 6, No. 2 (winter 2000).

Typology 1. hierarchy

Characterized by: *a single leader; a clearly defined hierarchy; a code of conduct (code of honour); strong systems of internal discipline; group known by a specific name; often strong social or ethnic identity; violence essential to activities; and often clear control over defined territory*.

The standard hierarchy was found to be the most common form of organized criminal group in the sample of 40 organized criminal groups. The standard size is some 10-50 members. Criminal groups in Asia, Eastern Europe and some groups in the Americas in particular fit this typology.

Typology 2. Regional hierarchy ("devolved hierarchy")

Characterized by: *a single leader; a line of command from the centre but a degree of autonomy at the regional level; a code of conduct (code of honour); strong systems of internal discipline; multiple activities; group known by a specific name; often strong social or ethnic identity; and violence essential to activities.*

Nowadays Asian (Japanese) and Italian organized criminal groups fit this organizational structure. The Italian organized criminal groups, for instance, have a hierarchical structure, headed by a single boss or an oligarchy. Most of these groups are based on a three-tiered organizational structure with a high level controlling a province (or a region), a middle level with representatives or families controlling a territory and a lower level of members executing the orders. In the case of several of the Asian groups operating in Australia, Japan or the United States, the overall business activities are overseen by a leader, but day-to-day business is left to "managers" who have significant levels of autonomy and are generally in full control of operations in specific geographical areas.

Typology 3. Clustered hierarchy ("hierarchical conglomerates")

Characterized by: *association of organized criminal groups with a governing or oversight body; high degree of autonomy for constituent groups but the cluster has a stronger identity than the constituent groups; often the result of a variety of individual criminal groups coming together to divide markets or regulate conflict between them; and over time, the cluster assumes some identity of its own*.

Relatively rare; some Italian, Russian and South African groups fit this profile.

Typology 4. Core group

Characterized by: *a core group surrounded by a loose network; a limited number of individuals form the core group (less than 20 persons); flat, horizontal organizational structure; seldom a particular social or ethnic identity (may include several nationalities); only in a limited number of cases known by a specific name; opportunistic, shifting to whatever activities promise most profits*.

Around the core group, there may be a large number of associate members or a network which are used from time to time depending on the criminal activity in question. Use of violence is less prominent than in the standard hierarchical groups. Such groups were mainly identified in Western Europe, Australia, Colombia and Mexico though as law enforcement pressure increases on hierarchical groups across the globe, this type of organization may well become one of the most common organizational forms.

Typology 5. Criminal networks

Characterized by: *activities of key individuals; prominence in the network is determined by contacts and/or skills; personal loyalties and ties are more important than social and/or ethnic identities; network connections survive coalescing around a series of criminal projects; low profile, seldom known by any name; and the network reforms after the exit of key individuals*.

Criminal networks are defined by the activities of key individuals who engage in illicit activity in often shifting alliances. Such individuals often do not regard themselves as being members of a group. Networks are formed around a key series of individuals (nodal points) through which most of the network connections run. It is likely that networks are more common than represented in the sample (as they are more difficult for the authorities to identify). They are considered to be a growing phenomenon. Even when key individuals are arrested, the network tends to reform itself quickly around new individuals and activities. Caribbean and West African criminal groups involved in cocaine and heroin trafficking, often fit the profile of such criminal networks, which from the outside may appear unstructured and not highly sophisticated, though in fact their operations are highly effective and enduring as such networks can quickly reform themselves around new players.

Source: Centre for International Crime Prevention, "Assessing transnational organized crime: results of a pilot study of 40 selected organized criminal groups in 16 countries", *Trends in Organized Crime*, vol. 6, No. 2 (winter 2000).

In terms of size, transnational organized criminal groups turned out to be smaller than expected, however. Most of the transnational organized criminal groups were found to have fewer than 100 members (63 per cent); 35 per cent were found to have a membership of 20-50 persons. This suggests that the price-setting capabilities of most organized criminal groups are in fact limited, as groups of this size have, in general, problems to achieve and defend a monopoly position in a drug market. There are, however, a few groups with several hundred members and a few groups with up to 1,000 regular and a total of 10,000 associate members [26]. Members are usually drawn from the same ethnic or social background (close to 60 per cent of the transnational organized criminal groups) [27].

Three quarters of the groups investigated were found to be also engaged in the legitimate economy. In almost half of the cases there was even extensive cross-over between legitimate and illegitimate activities, indicating clearly the high degree of penetration of the legitimate economy by such criminal groups. In the majority of cases, criminal activities took place in multiple countries [28].

Violence formed an essential part of the business activities in close to 60 per cent of the organized criminal groups investigated and in a further 25 per cent violence was used from time to time. For three quarters of the groups corruption was a key element in undertaking organized criminal activities, with either occasional or regular use. One third of the groups were said to have political influence at the local or the regional level [29].

Once the variables are correlated, a number of additional findings are revealed:

- The stronger the hierarchy, the more likely is the use of violence.
- Where trafficking in illicit drugs is regarded as either the primary activity of a group or an important core activity, the level of violence practised is generally much higher.
- There is also a positive correlation between involvement in the illicit drug trade and a strong ethnic basis in criminal groups.
- The stricter the hierarchy, the more likely a group is to have a strong ethnic or social basis.

In other words, the most violent groups are generally those which have a hierarchical structure, are characterized by strong social or ethnic identity and are involved in drug trafficking. In contrast, more loosely organized groups were found to be smaller in size, have no particular social or ethnical identity and use less violence. Thus, they are seldom seen as posing the same threat as hierarchical groups.

Against this background, dismantling the hierarchically organized Medellín and Cali cocaine cartels by the Colombian authorities in the early 1990s must be seen in a positive light—though it did not stop drug trafficking and de facto made possible an increase in the total number of organizations participating in this business. The new groups consist of tightly controlled core groups, assisted by a web of individuals engaged in auxiliary services. While drug trafficking operations were dominated in the past by 10-15 major organizations and their subsidiary groups, the illicit drug trade over the last few years is thought to have been dominated by 150-200 smaller organizations and many other groups made up of as few as 10 people [30].

One side-effect was increased competition and thus an ongoing pressure of cocaine prices to decline—even though enforcement efforts in the Americas, and notably in the United States, increased considerably.

Price-setting behaviour along the trafficking chain

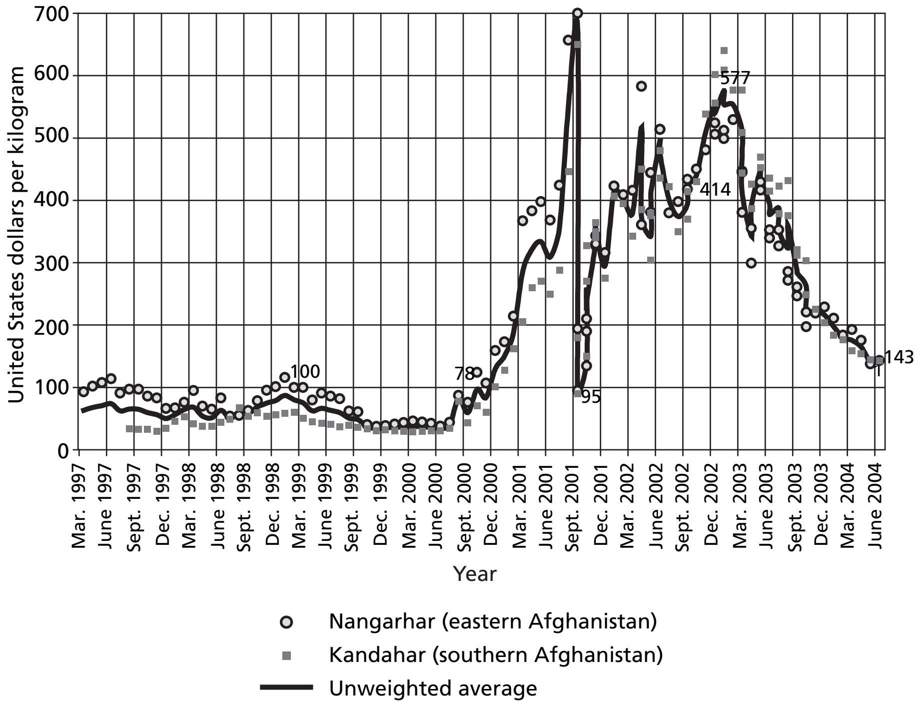
On the question of the price-setting behaviour of drug trafficking groups along the trafficking chain,* here the opium/heroin market will be investigated in more detail. That market is particularly interesting as major price changes have taken place in the main source country—Afghanistan—over the last few years. It is thus interesting to see how those price changes have affected the markets in neighbouring countries and in Western Europe.

Price changes in Afghanistan

From an initial analysis of changes in opium prices observed in Afghanistan, it can be seen that opium prices react strongly to actual changes in supply and to changes in expectations of future supply, which are often related to political factors that are seen to have a potentially important impact on future supply. Following reports of excessive rain at harvest time in 1998, opium prices started rising in Afghanistan, thus providing an additional incentive for farmers to grow opium. "Salaam arrangements" also played a role. (In such arrangements farmers sell some of their future opium harvest in advance to obtain cash immediately; the repayment of the cash loan is then in terms of opium.) Farmers who were not in a position to repay their loans with opium in 1998 often had to offer their creditors twice the amount of opium for repayment in 1999. As a result, Afghanistan had a bumper crop in 1999. Prices reacted immediately: they fell in eastern Afghanistan by almost two thirds by mid-2000 (from around $100 in early 1999 to $30 by mid-2000). Following the announcement by the Taliban of the ban on cultivation of opium poppy in mid-2000, the opium markets again reacted immediately. Prices rose 10-fold from mid-2000 ($30 per kg) to harvest time in 2001 ($300 per kg) and skyrocketed to levels close to $700 prior to 11 September 2001 as no end of the opium poppy ban was in sight (see figure V). This changed after 11 September 2001. Immediately following the attacks on the World Trade Center, traders in the Taliban-controlled areas tried to get rid of their stocks, fearing air attacks. Prices plummeted to levels around $90 as

*This section draws on experience that the United Nations Office on Drugs and Crime has acquired over the years in analysing the operations of the "opium economy" of the world's largest producer of opium, Afghanistan, and its impact on neighbouring countries and on the main final destinations, Europe and, to a far lesser extent, North America.

Figure V. Afghanistan: prices for opium, March 1997-June 2004



Source: United Nations Office on Drugs and Crime, Illicit Crop Monitoring Programme-AFG/F98, Monitoring of opium production in Afghanistan.

stocks were moved out of the country to neighbouring Pakistan [31]. Soon afterwards, prices started recovering, rising to around $350 per kg at harvest time in 2002 [32]. Despite a good harvest in 2002, prices continued upwards to levels around $580 per kg in January 2003. Speculative purchases by traders, following some eradication activities and notably following the announcement of the Karzai Government (in September 2002) that it would implement a new opium poppy ban in 2003, have been largely responsible for this price hike [33]. As the expectations of the market players of a harsh implementation of the opium poppy ban did not materialize, the good opium harvest of 2003 and the even better opium harvest of 2004 led again to a massive fall in opium prices to around $140 per kg in May 2004.

Price changes in neighbouring countries

As regards the question of how the price changes in Afghanistan affect prices in neighbouring countries, two opposing models have been proposed in the literature, the "additive" and the "multiplicative" price model. According to the additive price model, any increase in the price of a kg of opium in a producer country will be reflected by an increase of the same amount (in absolute terms) in a neighbouring country and ultimately in the countries of final destination. According to the multiplicative price model, the growth rate will be about the same along the trafficking chain.* In the first model it is assumed that traffickers are de facto paid a fixed price for their services, while in the second model it is assumed that profit margins, reflecting risk, remain unchanged.

These theoretical concepts have important policy implications. If one assumes that the drug markets operate along the lines of the additive price model, interventions by third parties in producer countries would have to be considered a waste of money. The impact of a $100 increase per kg of opium in Afghanistan, equivalent to a $0.1 increase per gram of opium or $1 per gram of heroin (assuming that 10 kg of opium are needed for 1 kg of heroin), would not really change the behaviour of drug users in consumer countries where a gram of heroin (at street purity) costs around $70 (Western Europe). By contrast, if the original opium price was around $50, an increase of $100 would be equivalent to a threefold increase in the producer price. If this were to translate into a threefold increase in consumer price, interventions at or close to source, driving up the consumer price, would be extremely efficient from an economic point of view in reducing drug consumption.

The massive changes in prices in Afghanistan provide an almost ideal opportunity to test the two opposing models for their practical relevance. For this purpose, the price developments in the Islamic Republic of Iran, Pakistan and Tajikistan, as well as in Western Europe, will be analysed. In order not to complicate the analysis too much, the following analysis will focus primarily on the impact of the opium price hike in Afghanistan in 2001 as a consequence of the opium poppy ban.

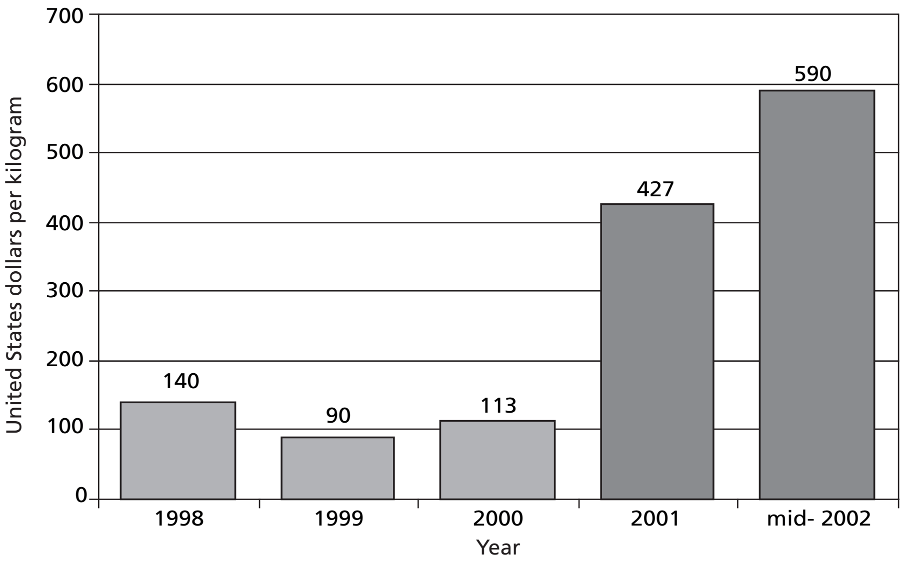
Price changes in Pakistan

Available information suggests that the largest outlet of opiates produced in Afghanistan is Pakistan [34]. Opium prices, based on average annual data, rose fivefold in Afghanistan in 2001 (from $60 per kg on average in 2000 to $300 per kg on average in 2001). The rise in Afghan opium prices prompted opium prices to rise fourfold in Pakistan, from slightly more than $100 per kg in 2000 ($113) to more than $400 per kg in 2001 ($427) (see figure VI). Based on the additive price model, a price rise of $240 per kg in Afghanistan should have resulted in prices of $353 in Pakistan, while a purely multiplicative model should have resulted in prices of $565. The actual opium price in Pakistan turned out to be in between—about 50 per cent higher than the expected results according to the additive price model and about 25 per cent less than the expected results according to the multiplicative model.

*Thus, if opium prices in Afghanistan rose from $30 to $300 or by $270 per kg between mid-2000 and mid-2001, the same increase should also be found in neighbouring countries. An opposing model is the multiplicative price model: if opium prices in Afghanistan increased 10-fold, a 10-fold increase in prices in neighbouring countries would be expected.

The results are more complex when it comes to heroin as additional factors are at work here. Statistics show a fall in heroin prices in Pakistan by about a third between 1998 and 2000 (see figure VII). This reflected—to some extent—Afghanistan's bumper harvests of 1999 and 2000. Opium prices fell over the same period in Pakistan; but their decline was far less pronounced (about 20 per cent) than the decline in heroin prices. In this case, both the additive and the multiplicative price-setting models fail to explain the price movement, as another factor played a more important role. The difference between the changes in the opium and the heroin prices is most probably due to the rising heroin manufacture capacity in Afghanistan during the late 1990s, which led to strongly falling heroin prices in Pakistan.

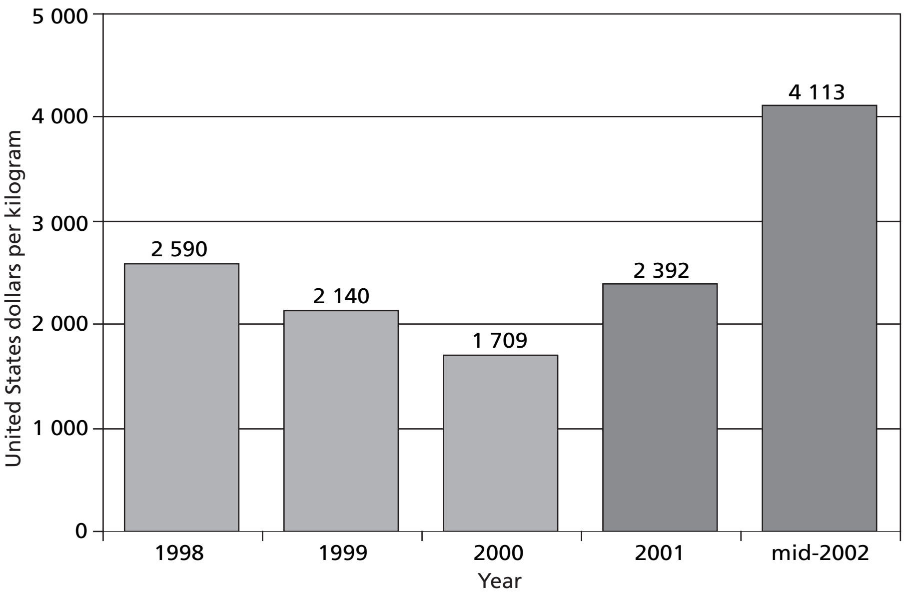
Figure VI. Pakistan: wholesale prices for opium, 1998-2002



Source: United Nations Office on Drugs and Crime, annual reports questionnaire data and Field Office.

Both the additive and the multiplicative price-setting models fail to explain subsequent price changes in Pakistan. When opium prices quadrupled in Pakistan in 2001, the overall increase of heroin wholesale prices was just 40 per cent (average 2001 data compared with average 2000 data). If prices of opium in Afghanistan had risen by $240 per kg over the period 2000-2001, they would have had to rise—according to the additive price model—by $2,400 per kg of heroin, assuming a 10:1 conversion ratio. If the transformation ratio of opium to heroin were changed to 7:1 (reflecting higher levels of laboratory efficiency and of the rather high morphine content of Afghan opium), the expected price rises would still have amounted to $1,700 per kg of heroin according to the additive model. The actual heroin price increase in Pakistan, however, amounted to less than $700 per kg. Likely alternative explanations for the actual changes in heroin prices can be seen in the existence of huge stocks of heroin—acting as a buffer—as well as changes in purity levels. (There were some reports that heroin purity deteriorated in the region in 2001.) In short, both the additive and the multiplicative price model failed to explain the price movements of heroin in Pakistan as additional factors apparently played a more important role.

Figure VII. Pakistan: wholesale prices for heroin, 1998-2002

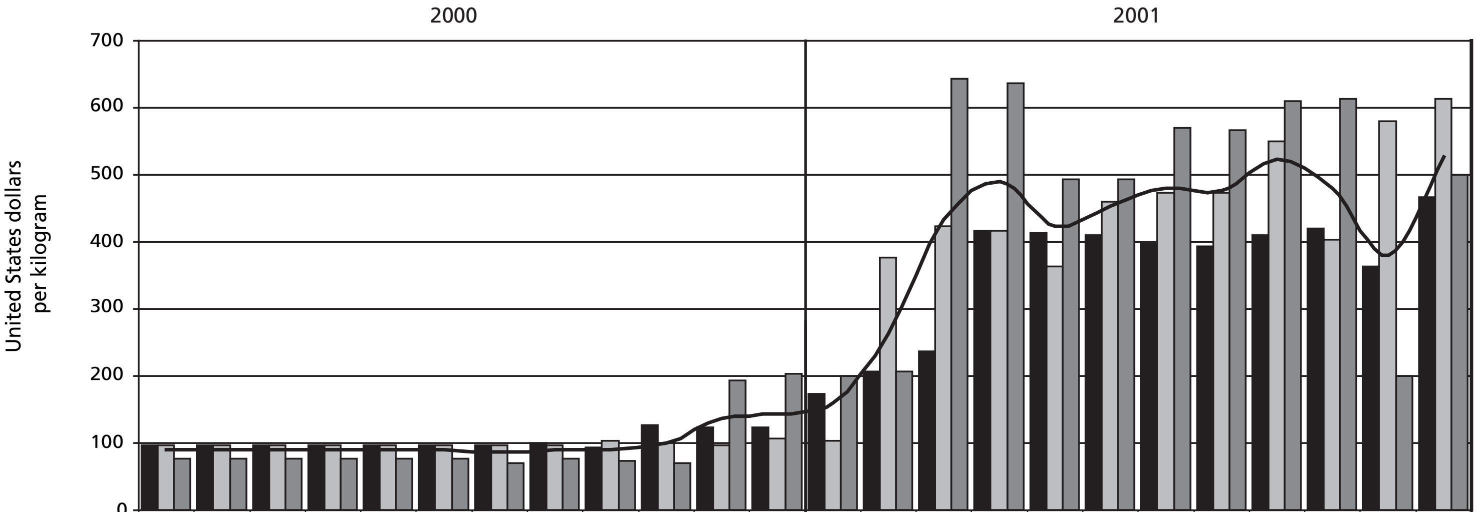


Source: United Nations Office on Drugs and Crime, annual reports questionnaire data and Field Office.

The basic results do not change significantly if the time frame is extended for an additional year. Between 2000 and 2002, the overall increase in heroin wholesale prices was two and a half times in Pakistan; over the same period, opium prices rose more than fivefold in Pakistan and increased more than six-fold in Afghanistan.

More detailed monthly data, available for 2000 and 2001, also confirm that picture. From the announcement of the poppy ban in Afghanistan in mid-2000 to harvest time in the spring of 2001, opium prices in Afghanistan actually rose 10-fold. The 10-fold increase in opium prices in Afghanistan led a four- to five-fold increase of opium prices in Pakistan in 2001 (see figure VIII). Average opium prices increased by about $400 per kg in Pakistan between January and December 2001. The outcome of the price rise was thus again in between the expected results of the additive and the multiplicative price model. Developments in opium prices in Pakistan showed that price rises turned out to be some 50 per cent more than expected by the additive price model, while the multiplier was about half as high as the multiplicative model would have predicted.

Figure VIII. Pakistan: prices for opium, 2000-2001

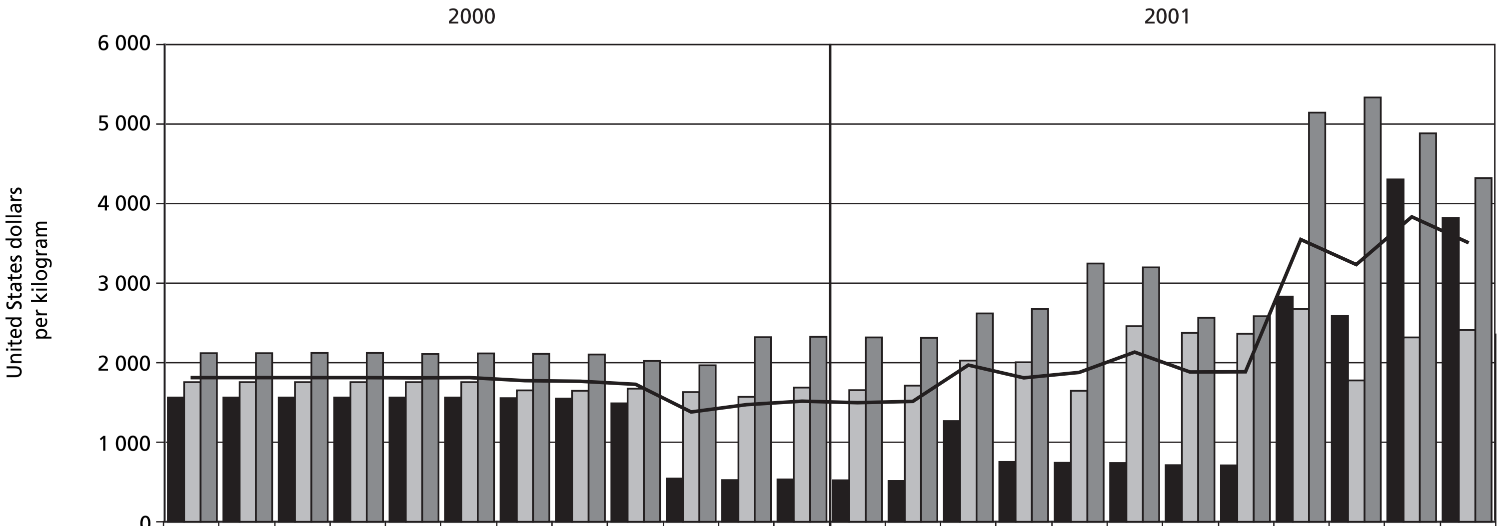


	2000												2001											
	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Peshawar	98	98	98	98	98	98	97	99	93	127	122	124	174	206	237	418	412	410	396	394	409	420	364	465
Quetta	98	98	98	98	98	98	97	97	102	100	96	107	105	377	422	418	363	459	475	473	551	404	579	615
Karachi	77	77	76	77	77	76	70	76	73	72	194	202	199	206	642	635	494	492	570	567	610	614	199	499
Average	91	91	90	91	91	90	88	90	89	99	137	144	159	263	434	490	423	454	480	478	523	480	381	526

Peshawar Quetta Karachi Average

Source: Global Illicit Drug Trends 2002 (United Nations publication, Sales No. E.02.XI.9).

Figure IX. Pakistan: wholesale prices for heroin, 2000-2001



2000

	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Peshawar	1 562	1 562	1 562	1 562	1 562	1 562	1 556	1 549	1 489	543	523	533
Quetta	1 757	1 757	1 757	1 757	1 757	1 757	1 653	1 646	1 675	1 630	1 570	1 687
Karachi	2 119	2 120	2 123	2 123	2 108	2 118	2 112	2 104	2 021	1 968	2 321	2 326
Average	1 812	1 813	1 814	1 814	1 809	1 812	1 773	1 767	1 729	1 380	1 472	1 515

2001

	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Peshawar	523	514	1 267	752	742	738	712	709	2 832	2 587	4 303	3 822
Quetta	1 655	1 713	2 027	2 005	1 648	2 460	2 375	2 363	2 674	1 778	2 317	2 409
Karachi	2 317	2 312	2 619	2 674	3 247	3 198	2 565	2 584	5 144	5 335	4 882	4 320
Average	1 498	1 513	1 971	1 810	1 879	2 132	1 884	1 885	3 550	3 234	3 834	3 517

Peshawar Quetta Karachi Average

Source: Global Illicit Drug Trends 2002 (United Nations publication, Sales No. E.02.XI.9).

If the analysis is extended to heroin, monthly data show price increases that clearly exceeded those seen in average annual data. Heroin prices—though not quintupling like the opium prices—actually doubled between the beginning and the end of the year 2001. The increase was, nonetheless, less than predicted by the additive or the multiplicative model (see figure IX). An increase in heroin prices in Pakistan was only observed in September, that is, six months after the price increases in opium started to be noticed. The explanation for this can be found in the existence of important heroin stocks, which acted as a buffer and thus prevented any increase in heroin prices for several months—and even when heroin prices did start rising, the stocks prevented them from skyrocketing. However, the quality of heroin was reported to have deteriorated. Thus, part of the actual price increases may have actually been caused by increased dilutions of the end product.

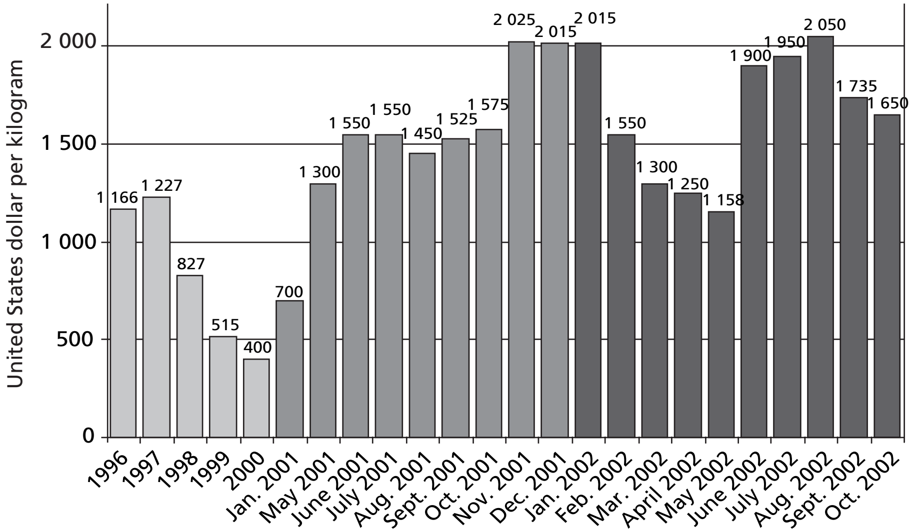
Price changes in the Islamic Republic of Iran

From Pakistan opiates are usually shipped to the Islamic Republic of Iran. There the price level of opiates is significantly higher than in Pakistan, which creates the main incentive for traffickers to engage in such an activity. In addition, opiates are also directly smuggled from Afghanistan into the Islamic Republic of Iran. Based on the analysis of the locations of individual seizures within the Islamic Republic of Iran (reported for 2001), it can be estimated that 66 per cent of the heroin and morphine enters the country via Pakistan and 34 per cent directly via Afghanistan [35].

Opium prices fell in the Islamic Republic of Iran between 1998 and 2000 by half, reflecting the bumper harvests of 1999 and 2000. Between 2000 and the end of 2001, opium prices quintupled in the Islamic Republic of Iran (see figure X).

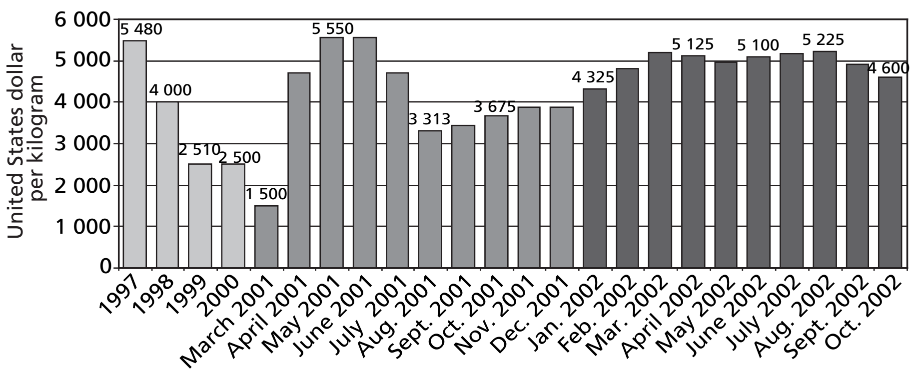
If the additive price model is applied, the price rise of about $270 in Afghanistan should have led to opium prices of around $670 ($400+$270) towards the end of 2001. If the price rises in Pakistan of about $400 were used, the additive price model would have predicted prices around $800 ($400+$400). The actual opium prices rose, however, to more than $2,000 per kg by December 2001. The price rises were thus twice as large as the additive price model would have predicted. Given a 10-fold increase of prices in Afghanistan and a four-to fivefold increase in prices in Pakistan and given the existing supply routes (66 per cent Pakistan, 34 per cent Afghanistan), the multiplicative model would have suggested a slightly more than sixfold increase of opium prices in the Islamic Republic of Iran to about $2,500 per kg. The actual opium prices turned out to be about 20 per cent less. In other words, actual opium prices in the Islamic Republic of Iran were again in between the additive and the multiplicative price model, but turned out to be far closer to the expected values of the multiplicative model than those of the additive model (see figure XI).

Figure X. Islamic Republic of Iran: wholesale prices for opium, 1996-2002



Source: The Opium Economy in Afghanistan: an International Problem (United Nations publication, Sales No. E.03.XI.6).

Figure XI. Islamic Republic of Iran: wholesale prices for heroin, 1997-October 2002



Source: United Nations Office on Drugs and Crime, Field Office.

This also meant that the profits per shipment of a kg of opium from Afghanistan to the Islamic Republic of Iran increased significantly though the profit margins declined. Buying a kg of opium in Afghanistan at $30-40 in mid-2000 and selling it at $400 in the Islamic Republic of Iran brought traffickers a gross profit of at least $360 per kg. The profit margins for shipping the opium from Afghanistan to the Islamic Republic of Iran thus promised a 10-fold increase. In mid-2002, a kg of opium could be bought in Afghanistan for about $350; selling it in the Islamic Republic of Iran at around $2,000 brought a profit of $1,650 per kg, that is, four times more than in 2000. The profit margins, however, fell from a ratio of 1:10 (Afghanistan:the Islamic Republic of Iran) to a ratio of 1:6 by mid-2002, and declined further to a ratio of 1:4 by October 2002.

One consequence of the far higher profits per unit of opium trafficked was that a larger number of small-scale traffickers participated in the business as the shipment of even small quantities became highly profitable. The larger number of participants, at the same time, led to increased competition and thus contributed to the reduction of profit margins. The average size of an opium seizure in the Islamic Republic of Iran in 2000 was 3.2 kg, falling to 1.9 kg in 2001, a decline of 40 per cent in just one year. The average size of a "significant opium seizure" as reported by the Iranian authorities to the United Nations Office on Drugs and Crime, Interpol and the World Customs Organization, fell from 142 kg in 1999 and 76 kg in 2000 to just 50 kg in 2001, a decline of 34 per cent over the period 2000-2001 (or 65 per cent over the period 1999-2001) and the average size of a seizure appears to have remained at the lower levels in 2002 as well. The average size of a significant opium seizure reported to the Office over the first eight months of 2002 amounted to 53 kg per shipment, 63 per cent less than the average size of a significant opium seizure in 1999.

Heroin prices, like opium prices, fell over the period 1998-2000 in the Islamic Republic of Iran and increased thereafter. Nonetheless, the increases in heroin prices in the Islamic Republic of Iran, like in Pakistan, were far more moderate than the increases in opium prices. Starting from a low of $1,500 per kg in March 2001, heroin prices rose to around $3,900 by December 2001, reflecting a two- to threefold increase, slightly more than in Pakistan, but far less than the increase in opium prices in Afghanistan, Pakistan or the Islamic Republic of Iran. Prices continued rising until August 2002 to about $5,200. By then the overall increase since March 2001 was by a factor of 3.5, but this was still less than the fivefold increase in opium prices over the same period in the Islamic Republic of Iran or the more than 10-fold price increase in opium prices in Afghanistan. The existence of heroin stocks may explain this phenomenon in part.

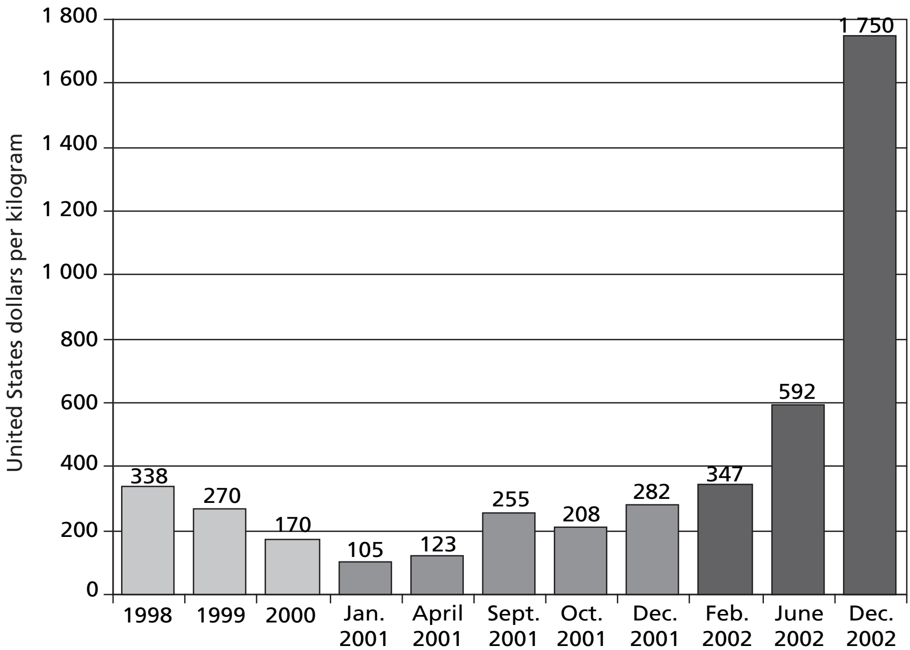
In addition, heroin purity declined in the Islamic Republic of Iran. While the authorities reported an average heroin purity in the Tehran street market of around 20 per cent in the late 1990s, purity levels fell to between 5 and 10 per cent in 2001. One negative side-effect of this were dilutions with all kinds of poisonous substances, causing the number of drug-related deaths to increase by 70 per cent in the Islamic Republic of Iran in 2001.

Price changes in Tajikistan

The main outlet for Afghanistan's opiate production in the north is Tajikistan. The basic picture is again similar to that of Pakistan or the Islamic Republic of

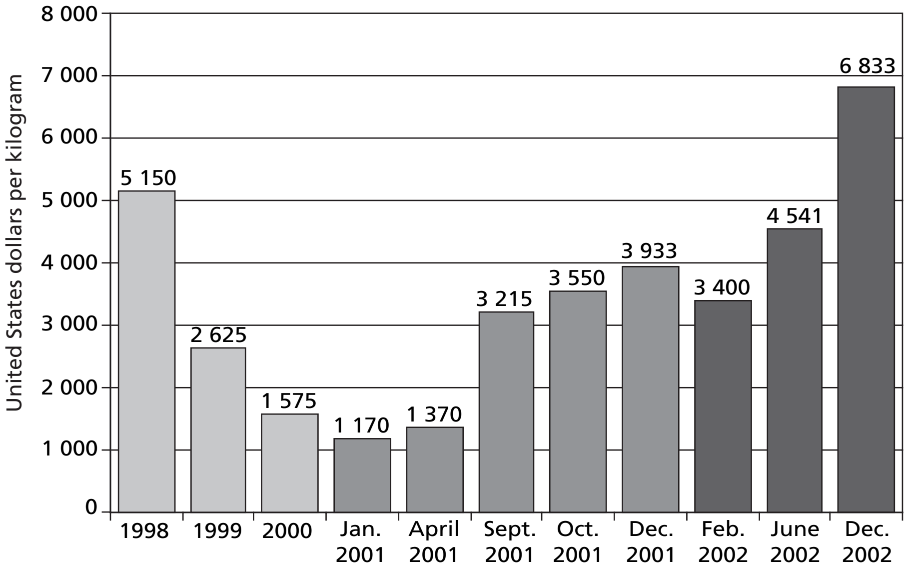
Iran. Both opium and heroin prices declined over the period 1998-2000 and increased in 2001 and 2002 (see figures XII and XIII). Nonetheless, there are some differences in price patterns, reflecting the lack of a unified opium market in Afghanistan in 2001. Tajikistan is mainly supplied with opiates produced in northern Afghanistan (notably Badakshan). No opium production ban was implemented in the northern provinces, which were outside the control of the Taliban regime. Thus, in the first months of 2001, practically no price rises were observed in Tajikistan. Price rises only started in September 2002. There was a less than threefold price rise between January and December 2001, while in Pakistan or the Islamic Republic of Iran prices rose about fivefold. However, as Afghanistan started to unify after the end of the Taliban regime, so did the drug market. Thus by mid-2002 opium prices had increased sixfold in Tajikistan, which was about the same growth rate as in Pakistan and an even higher rate than in the Islamic Republic of Iran (fivefold increase). Opium prices continued to rise steeply thereafter, reflecting ongoing price rises in neighbouring Afghanistan. In December 2002, wholesale prices for opium were six times as high as in December 2001 and 17 times as high as in January 2001.

Figure XII. Tajikistan: wholesale prices for opium, 1998-2002



Source: United Nations Office on Drugs and Crime, Field Office.

Figure XIII. Tajikistan: wholesale prices for heroin, 1998-2002



Source: United Nations Office on Drugs and Crime, Field Office.

Heroin prices rose threefold between January and December 2001 and almost fourfold between January 2001 and June 2002 in Tajikistan. There was thus hardly any difference in the growth rate of opium and heroin prices in 2001 (at a time when northern Afghanistan was still largely independent from the rest of the country). This shows clearly that unless special factors are at play (such as large stocks, changes in purity and changes in the manufacturing infrastructure), there is no reason why heroin prices would perform differently than opium prices. However, following an increasingly unified opiate market in Afghanistan in 2002, the same pattern was observed in Tajikistan as in Pakistan or the Islamic Republic of Iran. Rises in the price of opium exceeded those for heroin. While opium prices increased sixfold between December 2001 and December 2002 in Tajikistan, heroin prices were only some 70 per cent higher. Between January 2001 and December 2002 they increased sixfold, while opium prices were almost 17 times higher.

Summary of results of neighbouring countries

To sum up the discussion so far, it can be stated that a 10-fold increase in opium prices in Afghanistan resulted in:

- An increase of between four and five times in opium prices in the Islamic Republic of Iran and Pakistan in 2001.
- An increase of between five and six times by mid-2002 in the Islamic Republic of Iran, Pakistan and Tajikistan.

Increases in heroin prices were more moderate, apparently reflecting significant stocks in the region, as well as the tendency to reduce purity instead:

- Heroin prices rose between 2 and 3 times in neighbouring countries (the Islamic Republic of Iran, Pakistan and Tajikistan) in 2001 and between 2.5 and 4 times between early 2001 and mid-2002.

Impact of price changes on countries in Western Europe

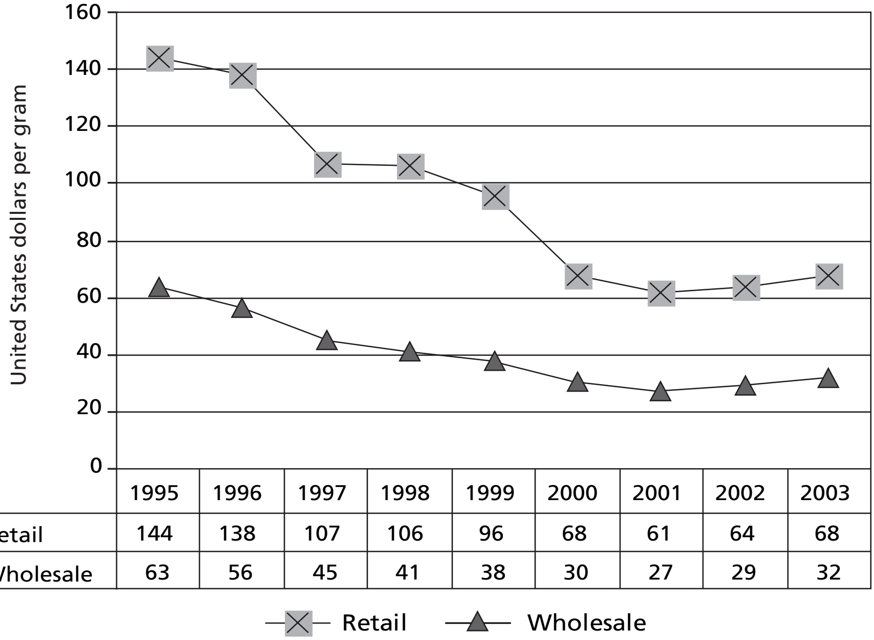
The next question is how price rises in the countries neighbouring Afghanistan translated into price changes in the European consumer markets.

Prices in Western Europe do not seem to have been much affected by the price rises in Afghanistan or the countries neighbouring Afghanistan (see figure XIV). Expressed in dollars, heroin wholesale prices (weighted by population as a proxy for market size) continued to fall by about 10 per cent in 2001, reflecting large heroin stocks along the trafficking routes, and increased by about 19 per cent over the period 2001-2003 to some $32,000 per kg. Heroin retail prices fell by about 10 per cent in 2001 and increased over the next two years by some 10 per cent to slightly less than $70 per gram, on average. Many of these changes, however, reflected shifts in the exchange rate between the dollar and the euro rather than actual changes of nominal heroin prices expressed in local currency. Expressed in euros, heroin prices declined only slightly in 2001 (−7 per cent), remained stable in 2002 and declined again slightly (−10 per cent) in 2003, apparently reflecting the resumption of large-scale opium production in Afghanistan in 2002 and 2003.

How can this apparent discrepancy between a doubling or tripling of heroin prices in the countries neighbouring Afghanistan versus largely stable heroin prices in Europe be explained? Is this the ultimate proof that neither the additive nor the multiplicative price model is correct in predicting price changes in the main consumer markets? Are there other factors at work that have a far stronger influence on the final outcome of the price-setting mechanism? One possible explanation could be a change in the demand curve, offsetting in large part the rise in the supply curve. In theory, such a shift in the demand curve could even lead to stable or even falling prices in the consumer markets despite higher producer prices (see figure XV).

This could be possible if a significant decline in overall consumption were to be observed. There have indeed been some indications of a decline of heroin abuse in Western Europe, as reported by Member States to the United Nations Office on Drugs and Crime. A number of treatment programmes set up in Western Europe, substituting heroin addicts with methadone and other synthetic opioids, appear to have contributed to a reduction in the demand for heroin. There also seems to have been a shift of younger drug users to drugs other than heroin. The average age of heroin users in Western Europe has been rising over the last few years. Heroin seizures declined in Western Europe by some 13 per cent in 2002 as compared with 2001, or by around 20 per cent as compared with 2000, possibly indicating that less heroin was trafficked (and thus probably less heroin was consumed) in the region.

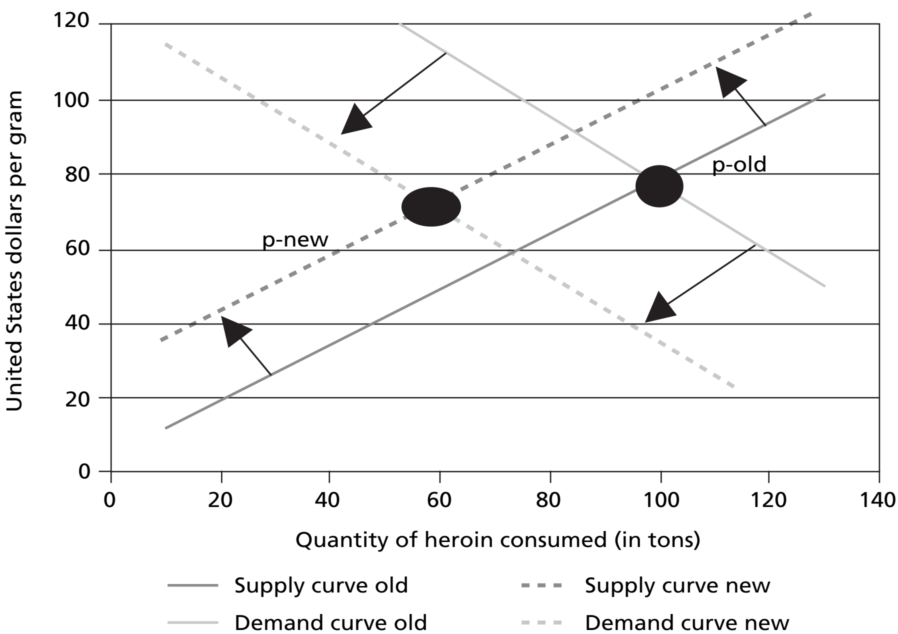
Figure XIV. Western Europe: prices for heroin (weighted by market size), 1995-2003 (not adjusted for purity)



	1995	1996	1997	1998	1999	2000	2001	2002	2003
Retail	144	138	107	106	96	68	61	64	68
Wholesale	63	56	45	41	38	30	27	29	32

Source: World Drug Report 2004 (United Nations publication, Sales No. E.04.XI.16), vol. 1.

Figure XV. Fall in drug prices despite an upward shift in the supply curve

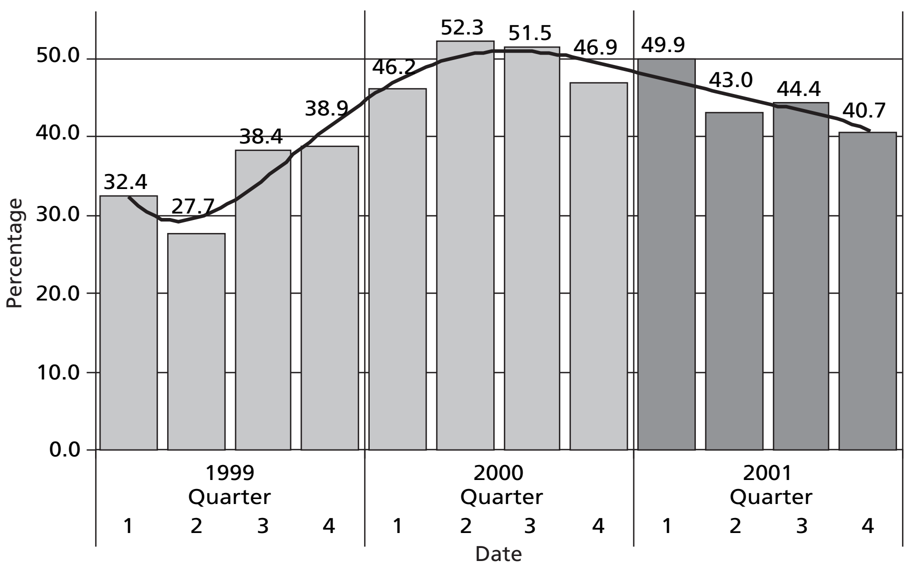


Nevertheless, it is still difficult to believe that in practice shifts in the demand curve in Western Europe would have been sufficient to neutralize the very strong price hikes of heroin reported from the countries neighbouring Afghanistan. In other words, the explanation of stable prices in Western Europe as the result of an offsetting shift in the demand curve may be part of the explanation, but it is unlikely to be the full explanation.

There is another important factor to be taken into account: changes in purity. As far as consumer markets are concerned, prices are not the main variable affected by changes in supply. Indeed, short-term market adjustments usually work via changes in purity rather than changes in prices.* Thus, the purity-adjusted prices rise or fall in line with supply conditions, while the nominal prices, expressed in local currency, tend to remain stable over longer periods of time. A practical reason—notably for retail sales—is that transactions have to be effected rather quickly at the street level to reduce risk. Thus round price figures are usually preferred and they are not changed on a daily basis. Thus, a short-term reduction in supply often results in lower purity levels rather than higher prices. The observed changes in purity levels in Europe seem to confirm such a reaction pattern:

- In *Turkey*, the overall purity of all heroin tested fell from 52 per cent in the third quarter of 2000 to close to 40 per cent in the fourth quarter of 2001 (see figure XVI).

Figure XVI. Turkey: average purity of heroin, 1999-2001[a]

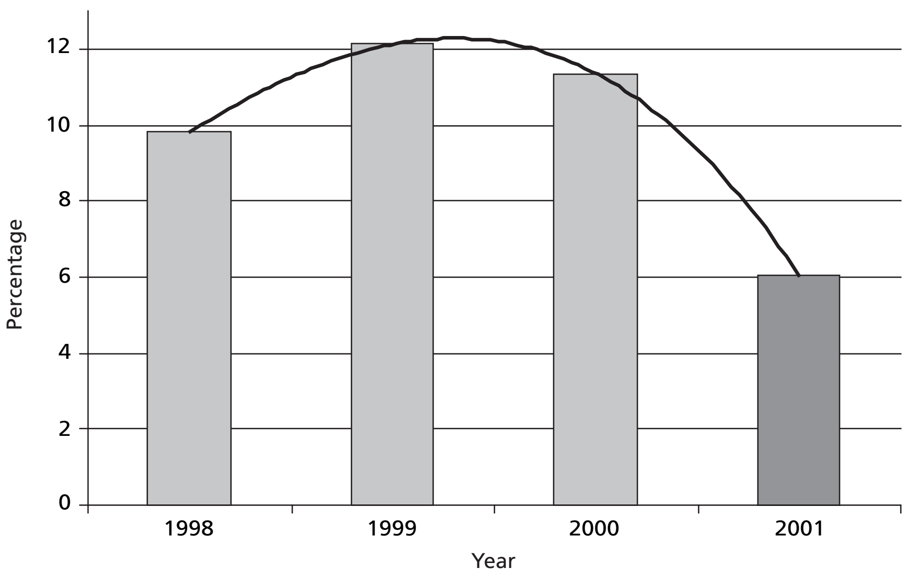


[a]Based on the forensic examination of 404 heroin seizure cases.
Source: Turkey, Ministry of the Interior/Gendarmerie, Narcotics Laboratory, Ankara.

*In addition, there have been reports that the quantities sold may differ, that is, a gram of heroin (or cocaine) can become less than a gram according to the metric system, though the nominal price per gram (whatever the true quantity) may well remain stable.

- In *Greece*, authorities reported that the purity of heroin on the retail market, which went up to 65 per cent in 2000, fell to between 8 per cent and 35 per cent in 2001.
- In *Slovakia*, the average purity of heroin on the local retail market was reported to have fallen from 12-50 per cent in 2000 to just 5-12 per cent in 2001.
- In the *Czech Republic*, authorities reported the average purity of heroin on the retail market to have fallen from 45-75 per cent in 2000 to 10-40 per cent in 2001.
- In *Estonia*, the typical purity of heroin was reported to have fallen from around 50 per cent in 2000 to 13 per cent in 2001.
- In *Lithuania*, the purity of heroin was reported to have declined in retail markets from 40-80 per cent in 2000 to just 0.1-10 per cent in 2001.
- In *France*, authorities reported that 11 per cent of all heroin seized and analysed had a purity of more than 50 per cent in 2000; in 2001 that proportion fell to 6 per cent [36]. The overall average purity of heroin was around 21 per cent in 2001 (see figure XVII).

Figure XVII. France: proportion of heroin tested with a purity level of more than 50 per cent,[a] 1998-2001

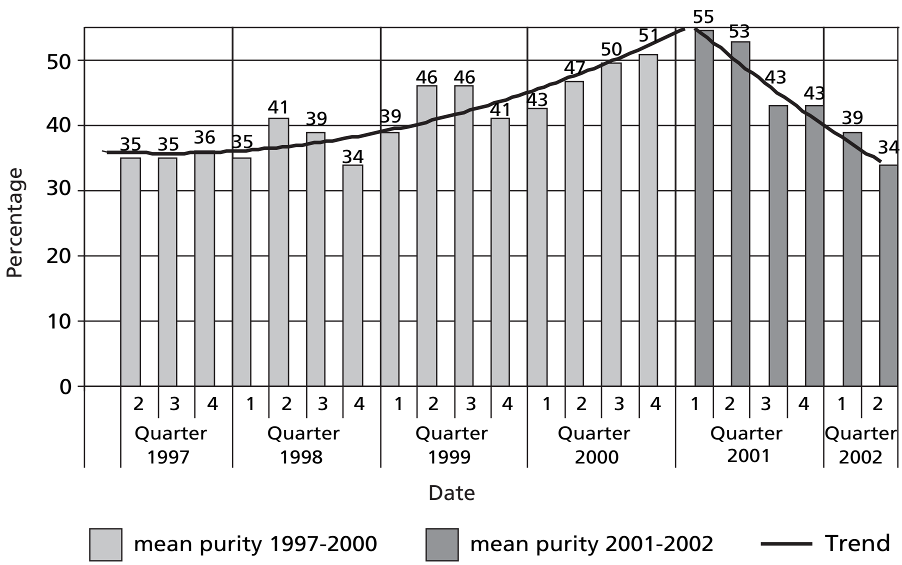


[a]Mean purity was 21 per cent in 2001.

Source: Laboratoire scientifique de la Police de Lyon and Laboratoire interrégional de la Douane de Paris, Observatoire française des drogues et des toxicomanies, *Indicateurs et tendances 2002.*

- In *Italy*, the average purity of heroin seized fell from 36 per cent in 2000 to 31 per cent in 2001 [37].
- In *Germany*, the average purity of heroin seizures of more than 1 kg fell from around 45 per cent in 2001 to some 27 per cent in 2002 (and to less than 10 per cent in 2003) [38].
- In the *United Kingdom*, tests carried out by forensic laboratories showed a decline in the average purity of heroin from 55 per cent in the first quarter of 2001 to, on average, 34 per cent in the second quarter of 2002 and to around 30 per cent by June 2002 (see figure XVIII).

Figure XVIII. United Kingdom of Great Britain and Northern Ireland: changes in the average purity of heroin, 1997-2002



Source: Forensic Science Service, United Kingdom.

According to the United Kingdom authorities' response to the annual reports questionnaire of the United Nations Office on Drugs and Crime, heroin prices in nominal terms declined by about 10 per cent, from £70 per gram ($106) in 2000 to £63 per gram ($90) in 2001. However, the change in purity between the first quarter of 2001 and mid-2002 was equivalent to an effective price rise of more than 80 per cent; a gram of pure heroin cost £127 ($186) at the beginning of 2001 and £210 ($318) at mid-2002 (assuming that the nominal heroin prices had remained largely stable between 2001 and mid-2002). That change would be equivalent to a net price increase of about 65 per cent in pound terms or 70 per cent in dollar terms.

In other words, a doubling (Pakistan) or tripling (Islamic Republic of Iran) of heroin prices in countries neighbouring Afghanistan in 2001 resulted in a de facto 70 per cent increase of heroin prices in West Europe's largest heroin market, though nominal heroin prices declined slightly. The multiplier model would have predicted a doubling or tripling of the prices in the United Kingdom. Based on the additive model, the price increases in the United Kingdom should not have been more than $2.4 per gram (Pakistan) or $3.6 per gram (Islamic Republic of Iran), equivalent to a price increase of between 2 per cent and 3 per cent in the heroin prices prevalent in 2000 ($106 per gram). The overall price increase in the heroin market in the United Kingdom was thus less than predicted by the multiplier model, but significantly more than the additive price model would have predicted.

These results have important policy implications. As long as the price-setting behaviour of illicit drug markets is not just additive but at least in part multiplicative, it seems to make economic sense to assist producer and transit countries with their enforcement efforts, as this will reduce the supply of drugs to consumer countries and help to decrease demand by raising drug prices in the consumer countries as well. In that context, it should also be borne in mind that interventions in and around the drug-producing countries are substantially less costly than in the countries of final destination. With one tenth of the United Kingdom budget at their disposal, the Iranian authorities, for instance, were in a position to take 30 times more opiates out of the market in the late 1990s.* Similarly high efficiency ratios of invested capital and returns, in terms of drugs seized, are also likely to apply to Tajikistan and other transit countries.

Summary and conclusions

Drug trafficking, in conceptual terms, can be analysed in relation to two key factors: profit and risk, as well as a number of enabling and protective factors. The key motivation is profit; the main limiting factor is risk. Prices are a key element in profit considerations. Law enforcement interventions, in general, will increase risk. Operating above certain thresholds, trafficking can be reduced. However, most enforcement interventions are below such thresholds. Under such

---

*In the Islamic Republic of Iran, the total drug control budget in 1998 amounted to 1,136 billion rials, of which 606 billion rials or $141 million (converted at the official United Nations exchange rate) were dedicated to supply reduction. For comparison, the total drug control budget of the United Kingdom—a country of similar size in terms of population and equipped with some of the most effective drug control organizations in Europe—amounted to £1.4 billion, of which £870 million, or $1.45 billion, were used for supply control purposes. The United Kingdom supply control budget was thus 10 times larger than that of the Islamic Republic of Iran. With that budget, the United Kingdom authorities succeeded in taking 1,346 kg of heroin out of the market in 1998. In the same year the Iranian authorities seized 2,895 kg of heroin, 22,291 kg of morphine and 154,454 kg of opium, which, expressed in heroin equivalents, amounted to 40.5 tons of opiates. Thus with just one tenth of the United Kingdom budget at their disposal, the Iranian authorities were in a position to take 30 times more opiates out of the market than their counterparts in a major European heroin consumer country.

conditions, increased risk will be reflected in increased profit margins and the higher profit margins will largely offset the higher risk. A new profit-risk equilibrium emerges. Nonetheless, as drug prices increase, consumption will decline, and so will trafficking. In other words, interventions on the supply side in general work indirectly.

In order to understand the price-setting behaviour of all the various players in the illicit drug markets, substance-specific issues as well as the organizational structures of criminal groups must also be taken into consideration. Location plays an important role: while much heroin and cocaine trafficking is interregional in nature, trafficking in synthetic drugs is largely intraregional. This has implications for the price structure. Interregional trafficking is, in general, more risky, leading to higher prices as compared with drugs that are only trafficked intraregionally. Interregional trafficking in synthetic drugs is often limited to trade in precursors that have been diverted from licit channels. The level of concentration of specific drug markets is another factor influencing price structure. Cocaine trafficking is still highly concentrated, though it is becoming less so. The heroin trade is far more fragmented, though it often operates along ethnic lines. In terms of involvement of organized crime, there are indications that the strongest involvement is in cocaine, followed by heroin, cannabis and trafficking in ATS. A study conducted by the United Nations Office on Drugs and Crime found that drug trafficking was the key business activity for more than half of the transnational organized criminal groups investigated. Though there is a general trend towards smaller, less hierarchically organized groups, two thirds of the criminal groups investigated still had a classic hierarchical structure. Such a hierarchical structure, in combination with an appropriate size and market concentration, are preconditions for criminal groups to be price-setters, reaping monopoly rents. Smaller groups, in contrast, are rather price-takers. The stronger the levels of hierarchy, the more likely a group is to use violence. The dismantling of the highly hierarchically organized Colombian drug cartels in the early 1990s thus reduced drug-related levels of violence and political interference, but overall trafficking did not decline, as a far larger number of small core groups, assisted by a web of individuals engaged in auxiliary services, replaced the old cartels. The new structure is more resilient to law enforcement interventions and has led to increased market competition. As a result, the downward pressure on cocaine prices continued in the 1990s, even though enforcement efforts were stepped up.

In terms of price-setting behaviour along the trafficking chain, an empirical analysis was conducted, testing the hypothesis of an additive price-setting behaviour of criminal groups versus a multiplicative price-setting behaviour. For the subsequent empirical analysis, the opium market in Afghanistan and the transit of opiates to consumer markets in Western Europe were investigated. The massive price changes in Afghanistan following the announcement of an opium poppy ban offered ideal, almost laboratory-type conditions for the analysis of the price-setting behaviour of drug trafficking groups. The poppy ban in Afghanistan in 2001 led to a 10-fold increase in poppy prices over the period between July 2000 and harvest time in 2001, from $30 per kg to $300 per kg, and, following some fluctuation, to a further increase to $350 per kg in mid-2002. By the end of 2002, opium prices amounted to $540 and were thus nine times higher than average annual prices in 2000 ($60). The result of the price hike was an increase in opium prices of between four and five times in neighbouring countries and an increase in heroin prices of between two and three times.

The price changes suggested that neither the multiplicative nor the additive price model were correct in predicting price movements. Actual price changes were in between, though, with regard to opium trafficking, leaning more towards the multiplicative price-setting model. Some interesting additional findings also emerged. As a result of rising opium prices in Afghanistan, profits per unit of drugs trafficked increased, prompting a larger number of small-scale traffickers to enter the business, thus increasing competition and reducing profit margins. In the case of heroin, the existence of huge stocks in the region appears to have biased the results.

Despite the doubling or tripling of heroin prices in the countries neighbouring Afghanistan in 2001, no price increases were reported from the transit countries or the countries of Western Europe that year and only moderate price increases were reported over the next two years (with wholesale prices—expressed in United States dollars—rising by some 20 per cent between 2001 and 2003). Theoretically, this could have been the result of a downward shift in the demand curve, offsetting to a large extent the upward shift in the supply curve. While the demand curve may well have shifted downward as a result of several demand reduction programmes in place, it is unlikely that it did this to the extent necessary to offset the marked upward shift in the supply curve. There is, however, an additional explanation at hand. While nominal heroin prices—expressed in dollars—remained stable or declined in 2001 and increased only slightly in 2002 and 2003, purity levels in several European countries were reported to have declined substantially following Afghanistan's opium poppy ban of 2001. Taking the data from the United Kingdom, the decline in purity more than offset the nominal price fall, so that the price of a hypothetical gram of 100 per cent pure heroin actually increased by about 70 per cent between the first quarter of 2001 and June 2002. The price rise was thus still less than the doubling or tripling of prices predicted by the multiplicative price model, but significantly higher than the expected price rise of 2-3 per cent according to the additive price model.

These results have important policy implications. As long as the price-setting behaviour of markets is not just additive but at least in part multiplicative, it makes economic sense to assist producer and transit countries with their law enforcement efforts. This reduces not only the supply of drugs to consumer countries but also helps to lower overall demand by raising drug prices. This must be also seen against the background that interventions in and around the drug-producing countries are far less costly than interventions in the countries of final destination.

References

1. Jonathan P. Caulkins, "The distribution and consumption of illicit drugs: some mathematical models and their policy implications", doctoral dissertation in Operations Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1990.
2. Jonathan P. Caulkins, *Developing Price Series for Cocaine*, MR-317-DPRC (RAND, Santa Monica, California, 1994).
3. Jonathan P. Caulkins, "Evaluating the effectiveness of interdiction and source country control", *Economics of the Narcotics Industry, Conference Proceedings, November 21-22, 1994*, RAND RP-410 (Santa Monica, California, 1994).
4. Barry D. Crane, *Deterrence Effects of Operation Frontier Shield* (Institute for Defense Analyses, Alexandria, Virginia, 1999), p. III-3-5.
5. *Global Illicit Drug Trends 2002* (United Nations publication, Sales No. E.02.XI.9).
6. United States Department of State, Bureau for International Narcotics and Law Enforcement Affairs, *International Narcotics Control Strategy Report* (March 2002).
7. United Nations Office on Drugs and Crime, Database for Estimates and Long-term Trend Analysis (DELTA).
8. *World Drug Report 2004* (United Nations publication, Sales No. E.04.XI.16), vol. 1, p. 108.
9. *Global Illicit Drug Trends 2002* . . . .
10. B. D. Crane, A. R. Rivolo and G. C. Comfort, *An Empirical Examination of Counterdrug Interdiction Program Effectiveness* (Institute for Defense Analyses, Alexandria, Virginia, 1997).
11. Jonathan P. Caulkins, "Qualitative analysis of drug prices", presented to the United Nations International Drug Control Programme, Vienna, 1998.
12. N. Braun, "Ökonomie der Sucht", research findings presented at a workshop organized by the Bayerische Akademie für Suchtfragen, 25 October 2002.
13. Henry Saffer and Frank Chaloupka, "The demand for illicit drugs", *Economic Inquiry*, vol. 37, No. 3 (1999), pp. 401-411.
14. *World Drug Report 1997* (New York, 1997), p. 125.
15. Ibid., p. 133.
16. Ibid., pp. 78-212.
17. *The Opium Economy in Afghanistan: an International Problem* (United Nations publication, Sales No. E.03.XI.6), pp. 53 and 54.
18. United Kingdom, Forensic Science Service, *Heroin Intelligence Database: Heroin Seized from 1 June 2000 to 31 May 2002* (London, 2002).
19. UNODC, annual reports questionnaire data.
20. UNDCP, *Amphetamine-type Stimulants: a Global Review* (Vienna, 1996).
21. *Global Illicit Drug Trends 2002* . . . , p. 129.
22. Ibid., p. 140.
23. United Nations Office on Drugs and Crime/Interpol/Customs Cooperation Council database on significant individual seizures.

24. Centre for International Crime Prevention, "Assessing transnational organized crime: results of a pilot survey of 40 selected organized criminal groups in 16 countries", *Trends in Organized Crime*, vol. 6, No. 2 (winter 2000).
25. Ibid., p. 26.
26. Ibid., appendix B, Overview of the 40 criminal groups surveyed.
27. Ibid., pp. 24 and 25.
28. Ibid., p. 29.
29. Ibid., pp. 27-30.
30. United Nations, Office for Drug Control and Crime Prevention, "Global trends in organized crime" (internal report), December 2000, pp. 32-36.
31. *Global Illicit Drug Trends 2002* . . . , p. 24.
32. United Nations, Office for Drug Control and Crime Prevention, *Afghanistan Opium Survey 2002* (October 2002), p. 36.
33. UNDCP, Field Office/Illicit Crop Monitoring Programme.
34. *The Opium Economy in Afghanistan: an International Problem* . . . (United Nations publication, Sales No. E.03.XI.6).
35. Ibid., p. 150.
36. Laboratoire scientifique de la Police de Lyon and Laboratoire interrégional de la Douane de Paris, *Indicateurs et tendances 2002* (Paris, 2002), p. 173.
37. Italy, Ministerio del Lavore de delle Politiche Sociali, *Report to the EMCDDA, Italy Drug Situation 2001* (Rome, 2002).
38. Germany, Bundeskriminalamt, *Bundeslagebild Rauschgift 2003* (Wiesbaden, 2004).

Illegal "lemons": price dispersion in cocaine and heroin markets*

P. REUTER

Professor, School of Public Policy and Department of Criminology, University of Maryland, College Park, Maryland, and Senior Economist, RAND Drug Policy Research Center, Santa Monica, California, United States of America

J. P. CAULKINS

Professor of Operations Research and Public Policy, H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States of America

ABSTRACT

The authors examine variability in the price and purity of cocaine and heroin using data gathered over a 14-year period by the System to Retrieve Information from Drug Evidence (STRIDE) of the United States Drug Enforcement Administration. The amount of variability is very great, larger than the variability for any of 15 legal goods for which comparable estimates are available. This raises the question of how such markets cope with the problem of purchasing "lemons"—undesirable or unreliable goods whose quality can not be determined ex ante. Repeated purchases may be an essential part of the answer.

Much, though not the majority, of variability in prices per pure gram comes from variation in purity. Conversely, the vast majority of variation in purity does not translate (with positive correlation) into variation in price per raw gram; instead, it translates (with negative correlation) into variation in price per pure gram. Thus, variability in price per pure gram is the "sum" of variability in price per raw gram plus variability in purity. The extent of this variability is curiously stable across drugs, market levels and time.

Introduction

Price dispersion is a characteristic of many markets. It is particularly high in markets for "lemons" (undesirable or unreliable products), such as used cars; in such markets, information concerning quality is costly to acquire. On the

*The research reported in the present article was supported in part by the National Institute of Justice of the United States Department of Justice and by the Office of National Drug Control Policy of the Executive Office of the President of the United States. Joel Feinleib provided valuable research assistance.

premise that extreme examples are instructive, the present article examines price and quality dispersion in illicit markets for cocaine and heroin. In those markets, as in Akerlof's classic "lemons" model [1], the buyer cannot observe the true quality of the goods at the time of purchase. However, unlike the classic "lemons" model, illicit drug sellers also usually have only incomplete knowledge of the quality of their goods, since they are not the manufacturers but have purchased the drugs themselves from higher-level dealers without having complete information about the quality. Furthermore, the authorities are actively engaged in suppressing the flow of information by, for example, making it risky for a seller to advertise.

The present article uses transaction-level data on over 145,000 individual illicit drug purchases and seizures; the data had been gathered over a 14-year period by United States Drug Enforcement Administration (DEA) laboratories. Analysis of the data revealed that such illicit markets are characterized by extremely high price and quality (purity) dispersion, apparently higher than that observed in any licit markets. Indeed, the dispersion is so high as to raise the question of how an illicit market sustains itself in the light of *(a)* opportunities and incentives for defrauding customers and *(b)* impediments to the dissemination of information. Hypotheses are offered about why illicit markets are able to function in settings that seem to invite the kind of persistent fraud that could lead to their demise.

Variation in price per gram unadjusted for purity is only weakly correlated with variation in purity, so a substantial minority of the dispersion in prices per pure gram (the true effective price) stems from dispersion in purity. Nevertheless, price dispersion falls in a narrower and more consistent range than does purity dispersion. The dispersion of prices is also very high across time, place and purchase quantity.

In the section below, the authors set the context by summarizing relevant findings on the sources of price dispersion in the literature on licit markets. In the next section those sources are related to characteristics of the illicit markets for cocaine and heroin. That is followed by a section on data and methods. The section after that presents the empirical findings. The last section summarizes findings, compare dispersion in licit markets and the illicit drug markets and discuss implications for economic theory and for drug control policy.

Insights from prior research on price dispersion and quality uncertainty

Since Stigler's seminal work [2], published in 1961, economists have observed that, even in highly competitive markets, heterogeneity in the willingness or ability of consumers to search for the "best buy" will allow ostensibly homogeneous goods to be sold for a range of prices in the same market.* While

*Even if every buyer has the same propensity to search, differences in initial information may be enough to generate price dispersion.

sellers who offer the competitive market price may attract consumers most willing to search for the "best buy", sellers who sell at higher prices can still attract consumers for whom finding the "best buy" is too costly [3-5].

In separate but related work, other researchers have investigated the dynamics of markets in which the quality of the goods to be sold may be well known to the seller but is largely unobserved by the buyer. In his classic article, Akerlof [1] described a "market for 'lemons'" in which such informational asymmetry leads to an oversupply of low-quality goods and, in extreme cases, to the disappearance of the market altogether as buyers refuse to enter it. Markets for so-called "experience goods", whose quality is only fully knowable after use, such as restaurant meals, used cars or illicit drugs, are especially vulnerable to the "lemons" principle.

The extreme case where markets actually dry up depends critically on the assumption that the seller's volume of sales is independent of its quality, which tends to be realistic only for markets in which a sale occurs precisely once [6]. The substitution of products of ever-lower quality may be reversed if sellers can make future gains by establishing their reputation for supplying reliable quality and garnering the goodwill of buyers. Given the great frequency with which individual buyers and sellers interact, the article suggests that such reputation effects may be important in cocaine and heroin markets. In licit markets, other mediating institutions may intervene to establish minimum quality standards or to publish data about product quality. These include the civil courts and the ability to sue for fraud, letters to the newspaper, the Better Business Bureau and, more recently, computer chat rooms or bulletin boards.

Prices can differ among sellers for a number of other, more predictable reasons, for example, because sellers have different production costs or supply sources. Numerous studies have documented the existence of price dispersion in markets for licit goods [5], including commercial airline tickets [7], auto insurance [8], fuel [9] and, more recently, books and compact discs (CDs) sold on the Internet [10]. There are very few studies that have examined empirically the "lemons" principle [11-13] or have explicitly focused on underlying measures of quality of goods sold where that quality is largely unobservable by the consumer. There are no known studies that have examined price dispersion in illicit markets.

Characterizing the cocaine and heroin markets

The literature reviewed above is relevant because illicit drugs are, ultimately, consumer goods and, like other goods in modern society, they are provided primarily through markets.

The markets for illicit drugs have distinctive characteristics. For example, the drugs are enormously valuable per unit weight: at retail, heroin costs 30 times as much per unit weight as gold [14].* These distinctive characteristics have implications. For example, smugglers can afford to employ sophisticated methods to conceal and transport even modest quantities of drugs. Retail purchases are in very small quantities, expressed as either doses or milligrams.

The present article is concerned with those characteristics of illicit drug markets which not only are distinctive but also pertain to price and quality dispersion. The most prominent are: the large number of distribution layers separating producers and consumers; the considerable uncertainty about product quality on the part of both buyers and sellers; the high cost (for both buyers and sellers) of searching and heterogeneity in the willingness of buyers to search for better prices; high and unpredictable turnover among buyers and sellers; and limited ability to signal quality.

Multi-stage distribution networks connecting producers and consumers

Cocaine and heroin distribution within consumer countries are almost purely brokerage activities.** In the United States of America, cocaine and heroin enter the country primarily in large shipments (in the case of cocaine, shipments containing 10-1,000 kilograms). They pass through about five intermediate transactions and are then sold in retail units of 0.1-1 gram. Hence, within destination market countries, all sellers are themselves also buyers. A fascinating implication of this for price and quality dispersion is that some considerations that pertain to buyers (notably their inability to reliably assay drug purity before purchase) may apply to sellers as well as to final consumers.

There are large quantity discounts as drugs move down these multi-tiered distribution networks [16]. The price of cocaine at the point of entry is 15-25 United States dollars per pure gram when sold in bundles of multiple kilograms; at the street level, the price is typically about $100 per pure gram [14].

Packaging and promotional activities add minimally to costs. Likewise, converting cocaine powder into crack costs very little as a proportion of the value of the drugs converted; this conversion can occur at any market level, but it usually occurs towards the lower end of the distribution chain and almost always within the United States. The drugs are sometimes diluted or adulterated with other ingredients, but the cost of those ingredients and the associated labour involved in "cutting" the drugs are minor relative to the value of the drugs and they are much less psychoactive than the drugs themselves. Hence, prices are most sensibly quoted in terms of prices per pure gram [17], and purity is the most important aspect of quality. There can be other aspects of quality. For

*The source of these high prices is presumably the substantial risks of arrest and of victimization by other market participants. Reuter, MacCoun and Murphy [15] estimated that in Washington, D.C., in 1988 a drug dealer faced a 22 per cent probability of incarceration. In addition, they estimated a 1 in 70 annual risk of being a homicide victim and a 1 in 14 risk of serious injury. Those violence-related risks have many sources, including lack of access to civil courts, ready access to guns, valuable property being held by young males and lack of written contracts.

*Cannabis is not included in this analysis because the transaction-level data available do not reveal the potency of the cannabis.

example, "black tar" heroin cannot be snorted whereas "No. 4 heroin" can be either snorted or injected. With respect to quality, the quantitative analysis presented in this article focuses exclusively on purity because it is the best-measured and most important aspect of quality for cocaine and heroin.

Uncertainty about quality (purity)

Transactions, particularly at the retail level, are frequently clandestine and hurried; purchase is the time of maximum exposure to law enforcement risks. In the United States, for example, purchases are usually made in round dollar figures ($10, $20) because even the act of returning change is a luxury that persons seeking to avoid detection cannot afford. The existence of standardized dollar purchase units has been observed for over four decades [18].

Though they are natural products, subject only to simple refining, cocaine and heroin are experience goods. The drugs are diluted as they move through the distribution system. At the time of purchase, the retail customer can make only an imperfect assessment of the quantity being purchased and has even less information about the chemical composition of the drug, including its purity and hence its psychoactive effect. Incomplete knowledge of quantity stems simply from the drug's extreme potency; hence, there are tiny quantities involved in retail transactions. One distinction between the retail and higher levels is that it is easier to weigh the larger quantities, both because of the settings in which they are purchased (more protection) and the greater quantities involved.

Not only is the purity of a drug not known at the time of its purchase, but it also may not even be estimated accurately at the time of its consumption. Some adulterants mimic the drug's physical effects (for example, numbing), and the user may have only a general notion of how much of the drug he or she actually consumed, since there is variability in response to the drug, depending on, inter alia, the time elapsed since the last drug ingestion and "set and setting" or expectation and context [19]. The user will make an assessment of the quality of the experience, but without certainty. Drug users are thus vulnerable to strategic manipulation by drug sellers, who can reduce the purity of the drugs by adding diluents and/or adulterants.*

What exactly is in street heroin, how pure is it and what are the effects of different "cuts" (adulterants or diluents)? The answers to those questions are the subject of much discussion on the street. In New York, assays of street-level heroin from a sample of 40 bags found that, in manufacturing the heroin, at least 27 types of adulterants and diluents had been used [18]. These observations are of retail transactions. At the wholesale level, there is in principle more

*Not all purity variation is strategic. The mixing of ingredients for retail distribution is done by hand and involves quantities so small that it is difficult to maintain consistent purity. Mixing 1 gram of heroin and 1 gram of lactose (a common diluent) and then attempting to put 10 milligrams of heroin in each 20-milligram packet probably leads to substantial variations in the purity of individual packages. Interested readers might wish to see for themselves just how difficult it is to manipulate 20 milligrams of salt.

opportunity and incentive to conduct systematic testing of purity. However, it is apparently not difficult to sell bundles of varying purity. The fact that testing often appears to be no more sophisticated than having a "taster" snort some to see if the experience is good suggests either a lack of sophistication in a trade that selects for other qualities (such as capacity for violence) or technological difficulties in testing (for example, the time required or ease of manipulation of purity within a bundle or set of bundles).* Fuentes [20], having researched Colombian cocaine-importing organizations in the United States, reported not having heard of any systematic testing (personal communication); refunds and replacements were available for bad shipments.** Accurate testing would require a mass spectrometer; not only would it be expensive to purchase such equipment, but it would also create substantial risk of disclosure.

In summary, sellers do not have more than a rough estimate of the purity of the drugs they are selling, however that distinguishes illicit drug markets from the classic "lemons" markets; information is asymmetric (the seller knows whether he or she "cut" the drug), but the person with better information about quality still has only very incomplete information about that quality.

High costs of search and heterogeneity in buyers' willingness to search

Comparison shopping for illicit drugs can entail significant costs. It is time-consuming and it also raises risks of arrest and/or violent victimization for both buyers and sellers.

Willingness to search may vary for a variety of reasons. Some are characteristics of the buyer (new or experienced; frequent or occasional; aversion to being arrested), others refer to the setting of the purchase (indoors or on the street; in the neighbourhood or in distant areas) and yet others represent "intra-individual" variation (whether "stoned" or sober; in regular resupply or while suffering withdrawal). Heavy drug users spend a larger portion of their income on drugs than do casual users and thus may have a strong rational incentive to seek out the best prices. However, the effects of heavy addiction (for example, more acute withdrawal symptoms) may, for various behavioural reasons (such as hyperbolic discounting), lead to a very abbreviated search [22].

Turnover of buyers and sellers

The range of prices and importance of quality uncertainty are affected by exogenous forces, but they may also reflect endogenous strategic interaction between buyers and sellers. A distinctive feature of the illicit drug market is that buyers

*As reported by a police lieutenant from Newark, New Jersey, "there is a rudimentary chemical test to see if the substance is in fact cocaine and not a fake substitute, such as milk sugar; however, there is apparently no interest in conducting even that test."

**These apparently refer to "intra-organizational" transactions. Manufacturers in Colombia shipped cocaine through a Mexican trans-shipment organization, which then delivered the drug to specific Colombian distribution groups within the United States.

and sellers frequently exit the market suddenly, without giving prior notice to others, as a result of arrest/incarceration or injury/death. Perhaps as many as one half of all heroin retailers are arrested in the course of a year for selling drugs or for other crimes; injury and their own drug habit may make them frequently unavailable at other times. Turnover among personnel in licit enterprises (such as restaurants) can also be high (for example, among waiters), but often the enterprises themselves continue to exist. In contrast, many sellers of illicit drugs are independent operators, and even drug-selling "organizations" have less institutional structure and culture with which to transcend turnover at all but the very lowest staff levels.*

Perhaps as a consequence, regular buyers of cocaine and heroin usually have more than one supplier; Riley [21], for example, found that, depending on the city, heroin buyers know an average of 10-20 sellers from whom they purchase the drug. High turnover among illicit drug sellers and buyers** should reduce the value of investing in a reputation for "honest" dealing; in strategic games of repeated interaction, for example, whether it is optimal to "cooperate" (in this case, sell goods of the expected quality) or "defect" (sell goods whose quality is lower than what is expected) depends on the probability that the game will terminate after a given move [23]. To the extent that buyers and sellers do not have continuing relationships, the incentive to sell low-quality goods should increase. Hence, all other things being equal, tougher law enforcement may increase dispersion by reducing the value of reputation.***

Limited ability to signal quality

Branding is common for licit experience goods. Consider, for example, the brand recognition strategies of restaurant chains. Meaningful branding is rare for illicit drugs perhaps because of the transience of selling organizations and the inability to create legally binding claims to product quality. The exception is retail heroin sellers who stamp "brand names" on their heroin bags. Goldstein and others [25] identified 400 heroin "brands" found in New York between 1975 and 1982. Wendel and Curtis [18] reported that one addict had lovingly collected 175 "brand" labels. A single organization might even sell five or six different "brands" simultaneously.

---

*Prior research suggests that illicit drug markets are rarely subject to monopoly or cartel arrangements; exclusion is too difficult. There may be monopolies over very small geographical areas (a few blocks), but cocaine and heroin markets are generally characterized as either competitive or monopolistically competitive.

**Buyer exits reduce the value of reputation, since some of that is customer-specific. Moreover, each buyer may, through referrals, lead to other buyers.

***However, tough law enforcement may lead to the market transforming from a street market into a social-network-based market [24]. Tough law enforcement may drive dealing indoors not purely through law enforcement "swamping" [22]; however, it may also exacerbate the "lemon" problem to the point where the stranger-to-stranger market is not sustainable. Users then revert to trading only with known partners, and it is easier for that to occur within routine activities in social networks than in a place-based market.

However, it is unclear how much meaningful information is conveyed by such stamps. As Wendel and Curtis [18] observed: "The principle of product recognition, however, is undermined by the frequent manipulation of quality and many stamps last only a few days before being replaced. To compensate for this instability and create the illusion that users have choice, many distributors (particularly large organizations which could afford to do so) simultaneously issue several stamps. Users are aware that different stamps do not necessarily mean different heroin and that one of the bags might often be better than the rest."

Similarly, Simon and Burns ([26], p. 65) have noted that, in Baltimore, Maryland, "labels are stamped right on the glassine packet . . . Free testers are tossed out every morning as word-of-mouth advertising for the coming package, and the touts are constantly trumpeting blue-light specials: two for the price of one, or a free vial of coke with every dime of dope." However, the transience of brand reputation is clear: "A product gets a reputation at the beginning of its run, but by the end, the cut takes over and the quality drops precipitously" ([26], pp. 79-80).*

The authors are not aware of any branding of cocaine powder at the retail level. Crack vials sometimes have caps of different colours, paralleling heroin stamps, and cannabis sellers sometimes describe verbally the source of their supply ("Colombian gold"), but there is little if anything to back up such claims. One reason stamps are not used more widely is that logos increase sellers' risks, since they help police connect a particular user and a seller, a problem that is exacerbated by intensified pressure [18].

Branding may also occur on the heroin market at higher levels. Though there are no data available for the United States, the Australian Bureau of Criminal Intelligence has reported on the branding of kilogram bundles of heroin imported from South-East Asia; at least one dozen brands have been found in the last decade. This is consistent with the claim that testing is difficult and thus purity is unobserved even for high-level transactions.

Data and methods

The data analysed in the present article are from the DEA System to Retrieve Information from Drug Evidence (STRIDE) [27], covering the years 1984-1997. About 75 per cent of the records are the product of investigations involving DEA and the Federal Bureau of Investigation (FBI). Most of the remaining 25 per cent of the records are from investigations of the Metropolitan Police Department of the District of Columbia, the jurisdiction that consists of the capital, Washington. Each observation includes information on the amount paid and the date, city, quantity, identity and purity of the drug, as well as the number of packages into which the drug was divided. A description of the data, including their strengths and weaknesses, is presented below. The results are presented

*This might suggest the value of early purchase, while information has value; cash-constrained customers with little self-control and urgent demands will have trouble doing so.

in easy-to-read univariate tables, although multivariate modelling suggests that the distinctions described below along one dimension persist even after controlling for the other dimensions.

The STRIDE database has both great strengths and great weaknesses. Its greatest advantages are that it contains a large amount of transaction-level data from throughout the United States, and it has recorded the data in a consistent manner over many years. Its principal disadvantage is that it is an administrative data set, not one collected for research purposes. For example, it is unbalanced by city and by size of purchase. It is in no sense a representative (let alone random) sample of the market. A National Research Council panel [28] has strongly criticized STRIDE; the use of STRIDE data has also been questioned by Horowitz [29]. However, STRIDE has been used to develop national and local price series that have performed well in a number of uses, such as explaining cannabis use among students in their final year of secondary school [30, 31].

STRIDE data need careful editing. For example, it may be that the collection procedures of DEA and the Metropolitan Police Department of the District of Columbia differ in ways that make it misleading to combine observations from those two sources. Indeed, there is some evidence that the Metropolitan Police are "smarter shoppers" than are DEA agents in Washington, D.C., in the sense that they may pay less for a given amount of drug. Nevertheless, it is clear that the STRIDE price data are not just noise; they are highly correlated with data from other sources, such as the number of drug-related emergency room mentions across cities [31, 32]. That is not to say that the variables are highly reliable; however they may contain some information and, in a field such as drug control policy, where data are so scarce, some information is better than none.

STRIDE does not represent all segments of the market equally. In particular, at the retail end, DEA undercover agents are not known customers of those from whom they purchase drugs, though they may simulate experienced users by their appearance and bargaining behaviour. This perhaps exposes them to higher than average probability of fraud. That is, STRIDE purchases (and perhaps, to a lesser extent, seizures) may have lower average purity than the market as a whole.

As in most administrative data sets, some information results from recording, coding and/or transcription errors. Excluded from this analysis are observations where crucial information, such as the weight of the drug, was missing and where values were so extreme as to be implausible. Table 1 describes the initial sample size and the number of observations excluded according to specific criteria for each of three drug types (cocaine hydrochloride, cocaine base and heroin) for the years 1984-1997. Specifically excluded are all observations from outside the United States, all observations where the raw weight was less than 0.02 grams, all transactions where the expenditure was less than $10 and all observations where the price per pure gram (cost/(weight x purity)) was less than $10 or greater than $5,000 for cocaine ($10,000 for heroin). The last exclusion was conservative in that it reduced measured variability but excluded a small number of wholly implausible combinations of price and purity.

Table 1. Sample characteristics, 1984-1997

Category	*Cocaine hydrochloride*	*Cocaine base*	*Heroin*
Total observations	99 388	66 693	48 348
Exclusions			
Foreign	–12	—	–11
Raw weight <0.02 g	–4 471	–4 095	–2 796
Expenditure <$10	–41	-36	–16
Price per pure gram <$10	–234	–154	–56
Price per pure gram >$500	–581	–837	–1 713
Total exclusions	–5 339	–5 122	–4 592
Percentage excluded	(–5.40%)	(–7.70%)	(–9.48%)
Total included observations (seizures and purchases)	94 226	61 632	45 017
Total included seizures	63 249	39 022	28 740
0% purity	–3 085	–2 111	–2 270
≤ 2% purity	–3 233	–2 211	–2 766
Total included purchases	30 977	22 610	16 277
0% purity	—	—	–1 051
≤ 2% purity	–6	–5	–1 375
Local (District of Columbia) police seizures and purchases			
Seizures	–5 400	–23 535	–8 446
Purchases	–1 228	–6 400	–1 806
Total	–6 628	–29 935	–10 252
Total included observations, excluding local (District of Columbia) seizures and purchases and excluding ≤ 2% purity	84 598	30 111	31 659
Total included purchases, excluding local (District of Columbia) seizures and purchases and excluding ≤ 2% purity	29 743	16 205	13 254

The above-described procedure resulted in the exclusion of 5.4 per cent of the cocaine hydrochloride observations, 7.7 per cent of the cocaine base observations and 6.9 per cent of the heroin observations. Most of the exclusions were made because the recorded weight was less than 0.02 grams. In addition, for analyses where data from many localities were grouped together, data acquired by the Metropolitan Police Department of the District of Columbia were excluded. The main reason was that the data gathered by the Metropolitan Police Department accounted for such a significant share of the total data set (7 per cent of the data for cocaine hydrochloride, 48 per cent of the data for cocaine base and 22 per cent of the data for heroin) that any aggregate summary statistics of the whole sample would be too heavily weighted towards the features of the market in the District of Columbia, including those features which might arise because the Metropolitan Police Department, in acquiring its samples, might be using techniques differing from those used by federal agencies. The remaining data set was more evenly weighted across cities and only contained samples acquired by federal law enforcement agencies, namely FBI and DEA, including those acquired in Washington, D.C. Excluding the data described above, the Metropolitan Police Department samples and those with a purity of less than 2 per cent, there were 84,598 cocaine hydrochloride records (of which 35 per cent involved purchases), 30,111 cocaine base records (of which 54 per cent involved purchases) and 31,659 heroin records (of which 42 per cent involved purchases).

The main focus of the analysis presented in this article is the degree of variability in the purity and price per pure gram for both cocaine and heroin. Cocaine appears in two forms: "cocaine hydrochloride", usually referred to as cocaine powder, and "cocaine base", most commonly in the form of "crack".* For analysis of purity the authors included data on purchases and seizures, but only data on purchases could be used to derive prices per pure gram.

For a significant proportion of seizures, the recorded purity level was "zero", but it was impossible to accurately distinguish samples actually containing no psychoactive substance from those cases where there was merely a failure to assay the purity of the samples or to record it properly. Again, the authors adopted a conservative approach in excluding samples where the recorded purity level was less than 2 per cent, in order to avoid both samples with "zero" purity and those with extremely low values that might be transcription errors. The net effect was that the analysis should slightly understate the true variability of purity.

STRIDE records the number of separate packages into which each seizure or purchase was divided. For purposes of analysing purity by the weight of the transaction, the authors assigned a level of purity to the average weight per package but not to the total weight. However, for subsequent analyses of retail quantities and prices by year or across cities, the total weight was used.

*"Crack" and "cocaine base", the codes used in STRIDE, are not identical. Free-based cocaine might be classified as "base", and that probably accounts for most of the few observations in this category in the early 1980s. By the mid-1980s, crack, being a more efficient mode of obtaining a quick and intense high, had supplanted earlier forms of cocaine base.

Methods

Calculation of pure gram price

Price was measured in terms of dollars per pure milligram (actual recorded expenditure ÷ (recorded weight x purity)). That value was an approximation, since total weight might have hedonic value as well: 200 milligrams of pure cocaine contained in 500 milligrams might provide a slightly different experience than if it were contained in 1 gram. It was assumed that that is a second-order effect.

Measures of variability

There is no single convention for describing the variability of a particular statistical distribution, especially when that distribution is not symmetric, nor is there a single convention in the empirical price dispersion literature. Therefore, a number of measures were analysed for both purity and pure gram price, as well as the full shape of the underlying distribution not readily captured in the summary statistics; these included the coefficient of variation (standard deviation divided by the mean), the interquartile range divided by its median, the standard deviation of the logged values (which, especially for prices, tended to be normally distributed) and the Gini coefficient. There was a very high degree of correlation between the various measures. The Gini coefficient is emphasized because it provides a convenient interpretation: twice the coefficient gives the absolute difference in price (purity) as a proportion of the mean price (purity) to be expected from two observations drawn at random from the population.

In order to reduce the influence of extreme outliers that might be the result of administrative errors in the data (implausible combinations of weight, purity and expenditure not already excluded by the conservative procedures described above), measures of variability for the price per pure gram were calculated only for the range of values between the 5th and 95th percentile. That, again, would tend to understate price dispersion and thus, was conservative with respect to the central conclusion that dispersion was very large.

Findings

Substantial price and purity variation for cocaine and heroin

The main descriptive finding was that the markets for cocaine and heroin were characterized by extremely high dispersion in purity and price, particularly when compared with licit goods. Table 2 contrasts the coefficient of variation in prices for cocaine and heroin with those for licit goods as reported in a number of studies.

Table 2. Comparison of price dispersion measures

Category	*Product*	*Coefficient of variation*
Panel A. Various studies: reported coefficient of variation in retail prices		
Pratt, Wise and Zeckhauser (1979):		
Boston area[a]	Paint	0.071
	Mufflers	0.173
	Cameras	0.090
	Fuel oil	0.289
	Stationery	0.060
	Lumber	0.130
Treno and others (1990): California[b]	Beer	0.190
	Wine	0.340
	Spirits	0.240
Walsh and Whelan (1999):		
Irish grocery stores, 1992-1995[c]	Mineral water	0.077
	Sugar	0.107
	Coffee	0.126
	Bacon	0.058
	Tea	0.038
	Dog food	0.144
Panel B. Illicit drug markets, STRIDE data for the period 1984-1997		
Average coefficient of variation in retail prices observed in any given year	Cocaine hydrochloride	0.755
	Cocaine base	0.476
	Heroin: "high-purity cities"	0.661
	Heroin: "low-purity cities"	0.630
Average coefficient of variation in purity of retail quantities observed in any given year	Cocaine hydrochloride	0.331
	Cocaine base	0.218
	Heroin: "high-purity cities"	0.520
	Heroin: "low-purity cities"	0.800

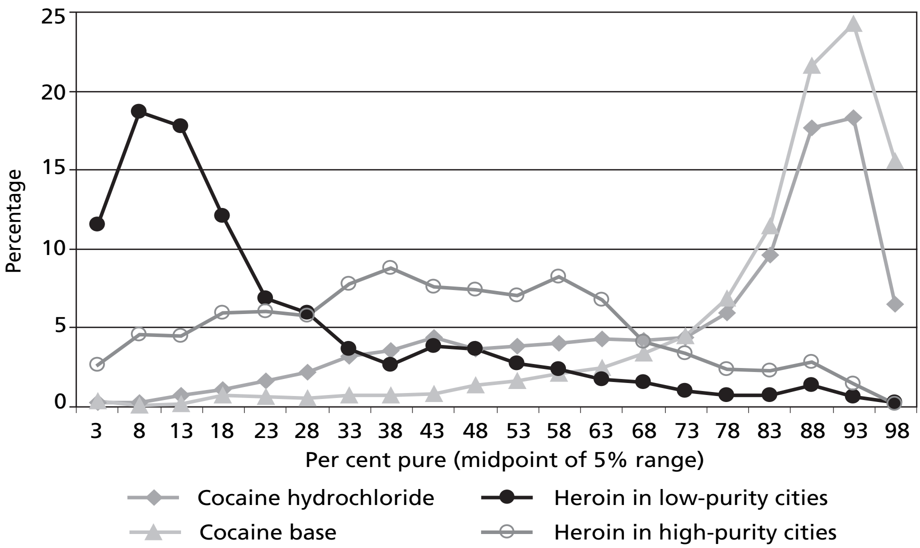
[a]John Pratt, David Wise and Richard Zeckhauser, "Price differences in almost competitive markets", *Quarterly Journal of Economics,* vol. 93, No. 2 (1979), pp. 189-211.

[b]A. J. Treno and others, "Alcohol beverage price spectra: opportunities for substitution", *Alcoholism: Experimental Research,* vol. 17, No. 3 (1993), pp. 675-798.

[c]P. P. Walsh and C. Whelan, "Modelling price dispersion as an outcome of competition in the Irish grocery market", *Journal of Industrial Economics,* vol. 43, No. 3 (1999), pp. 325-343.

The data for both cocaine and heroin were divided into two groups. STRIDE distinguishes between "cocaine hydrochloride" (cocaine powder) and "cocaine base". For the years analysed, "cocaine base" was primarily crack. The heroin observations are divided into two groups by location, because there were striking differences in purity across two groups of cities. As shown in figure I, the interquartile range of heroin purity in "low-purity cities" was 6-29 per cent, whereas for "high-purity cities" the range was 29-59 per cent.

Figure I. Distribution of purity, retail quantities, by drug type, 1987-1991



The licit goods included some that were very homogeneous (e.g. sugar and tea) and others that were highly differentiated (e.g. wine and cameras). The price of wine was expressed in dollars per ounce of ethanol, but there was little variation in the alcohol content of different wines; the differences were in observable and moderately well-signalled quality. Low-priced table wine is not sold in the same market as an expensive bottle of fine wine from a well-known wine-producing area.

None of the licit goods except sugar could claim to be as undifferentiated as cocaine and heroin if their purity were observable. Yet only the highest coefficient, 0.340 for wine in California, approached the range of coefficients of variation for any illicit drug. The lowest observed average coefficient of variation for illicit drugs was higher than those for all 15 licit goods.

Few empirical studies have dealt explicitly with variability of quality, so it was hard to compare the range in purity of illicit drugs with the counterpart dimensions of other goods. Indeed many licit goods, including alcohol and prescription drugs, have quite strict controls on the purity of active ingredients, thus constraining quality variation. On the other hand, some reports on unregulated but extremely popular herbal supplements, such as St. John's wort or gingko biloba, suggest that the possible range of ingredient quality is quite high [33].

The amount of variability in the purity of illicit drugs is striking. The degree of variability in purity was highlighted by analysing the distribution of purity for a typical drug purchase of $100 (the modal expenditure in the STRIDE data). Figure I shows the histogram for each of the four types of illicit drugs for the period 1987-1991. The distributions for the various drugs are different, but all show large variability.

Much, though not the majority, of the variability in prices per pure gram comes from variation in purity. Conversely, the vast majority of variation in purity does not translate (with positive correlation) into variation in price per raw gram; instead, it translates (with negative correlation) into variation in price per pure gram. Looking at it another way, price per raw gram is highly variable and is very weakly correlated with purity, so variability in price per pure gram is the "sum" of variability in price per raw gram plus variability in purity. This is consistent with the expected purity hypothesis of Caulkins [34].

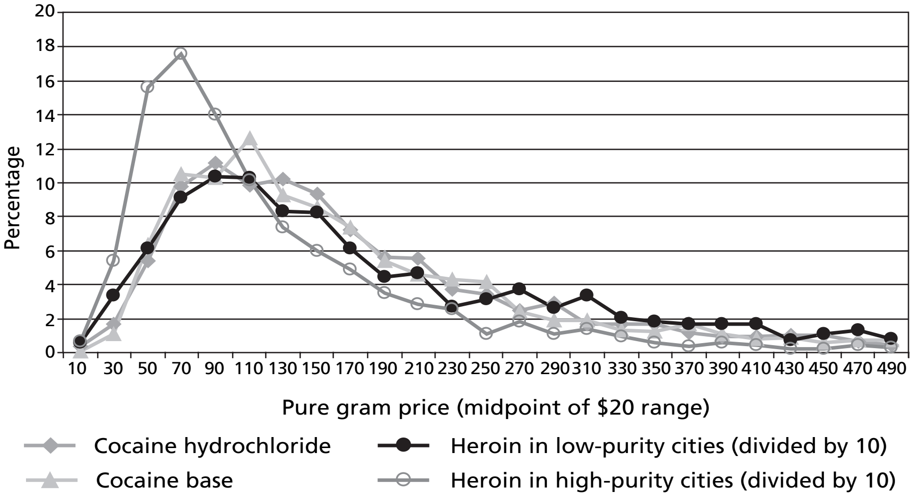
The variability in price and purity might simply be an artefact of combining heterogeneous but much less variable profiles of distributions from many cities. Or, if there were significant seasonal factors influencing purity, measuring over a single year or several years might again combine distinct but much less variable periods. Tests for seasonal and yearly effects, as well as day-of-week effects, in purity and price found none. Moreover, the variability did not seem to come primarily from systematic "within-period" price variation (i.e. price spikes).

Another omitted variable explanation was that there might be multiple distinct markets within a single city that were being combined, but *(a)* the lack of correlation between purity and price per raw gram undermines theories of distinct low-price/low-purity and high-price/high-purity markets and *(b)* the lack of correlation across cities between price variability and city size cast some doubt on the idea that this market aggregation story was the true explanation for most of the apparent variability. If it were, larger cities would be expected to display greater variance because they could sustain more different markets.

Curious stability in the amount of variation in illicit drug prices

There were noticeable differences in the variability of purity by drug, market level, time and place. In contrast, price dispersion, while very high, was highly stable over transaction size, time and drug (though not necessarily city). Particularly striking was the closeness of the retail price distributions for all four drugs (heroin being divided between "low-purity" and "high-purity cities") and their approximation to a log-normal curve. (Compare figure II with figure I.)

Figure II. Distribution of pure gram prices, retail quantities, by drug type, 1987-1991



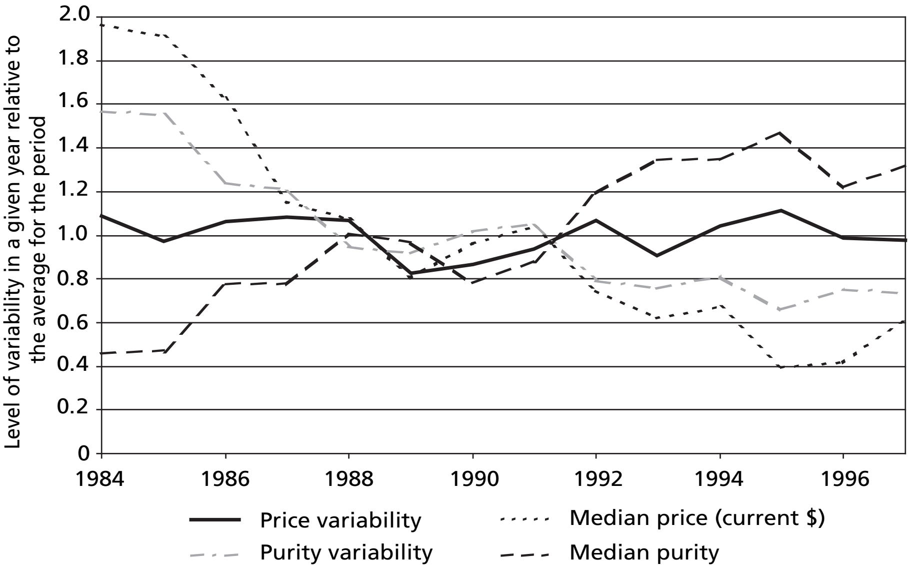
Purity was not consistently less or more variable than price: crack purities were less variable than any price measure; heroin purities in "low-purity cities" throughout the 1980s were consistently higher.* Instead, for all three categories of drugs in powder form (heroin of both low and high purity and cocaine) in almost all years, the Gini coefficient for price variation was between 0.28 and 0.38. For crack, the coefficient was almost always between 0.2 and 0.3.** For purity, there was low and stable variability for crack (0.08-0.15), moderate and stable variability for cocaine powder (0.15-0.24) and declining variability for heroin of high purity (0.47-0.21) and even for heroin of low purity (0.47-0.37 or so).

The stability of the Gini coefficients for price variability was all the more striking when considering how much changed in illicit drug markets during the period under review; for example, purity varied enormously for the cities with high-purity heroin, increasing from nearly 20 per cent to 60-70 per cent purity.

*The differences in the variability of purity across drug types may reflect the nature of that particular drug and its appeal. Heroin purity is truly heterogeneous, while crack is fairly uniform, perhaps because drugs in powder form are easier to dilute than are "rocks" of crack. This may help to explain the popularity of each drug. Potential crack users can be fairly confident about the potency of their drug of choice. Potential heroin users, however, must contend with the strong possibility that the drugs they buy will be either too potent, leading to risk of overdose, or insufficiently potent, leading to continued craving. The emergence of heroin of very high but less variable purity may have helped boost the popularity of that drug in the late 1990s.

**More specifically, there was little variation across years for heroin in either "high-purity cities" or "low-purity cities"; the range of Gini coefficients was 0.280-0.381. Cocaine base also showed moderate variation and no trend after 1987; the range was 0.212-0.342. The large coefficients were from earlier years, when crack was just entering the market. From 1988 to 1997, the range was only 0.212-0.296. For cocaine powder, there was substantial variation (0.289-0.449) but no trend. Cocaine in both forms showed a spike for 1990, the year of the crackdown on the Medellin cartel. Though heroin prices increased at that time, the variability did not.

Figure III. Heroin in "high-purity" cities: stable variability in price, 1984-1997



That change in median purity apparently greatly affected the variability in heroin purity and the median price per pure gram of heroin, which fell by 75 per cent even before adjusting for inflation, but the variability in heroin prices was stable throughout. (See figure III.)

The exception to this rule was crack. Variability in crack prices tended to decline between 1985 and 1997, while variability in crack purity showed no consistent trend, first decreasing and then rebounding.

It was also not the case that variability was constant across locations; for example, among cities with adequate data (50 or more observations between 1987 and 1991), Gini coefficients were tightly clustered across locations for heroin prices (0.218-0.308), cocaine purity and both price and purity of crack. Gini coefficients were diverse for cocaine prices (0.092-0.555) and heroin purity (0.171-0.308).* A striking and unexpected result, however, was the overall tendency for the Gini coefficient for variation in price to be quite consistent across drugs, time and, as shown below, market level.

Differences (and lack of differences) across market levels

There is not one market for illicit drugs but many markets, differentiated by transaction size. Transactions involving 1 kilogram of cocaine (valued at roughly $25,000) differed from transactions involving 0.2 (pure) gram (valued at roughly

*There was no consistency in price variability across cities for different drugs, with a rank correlation in price variation of only 0.33 between cocaine powder and crack and 0.02 and 0.03 for cocaine powder and heroin and for crack and heroin, respectively.

$20) in terms of location, type of participants, stakes and many other aspects. It was hypothesized that there would be less dispersion in both prices and purity at higher market levels, where incentives and opportunities for verifying the quality of the drugs or the seller were greater. The expectations were generally confirmed for purity but not for price.

Moving down the distribution chain from larger wholesale quantities to small retail amounts, the median purity of heroin and cocaine powder declined as various diluents and fillers were added to the product. The degree of variability increased commensurately (see table 3).*

Table 3. Price and purity levels and variability by market level

	Gini coefficient				*Median*			
Weight category	*Cocaine powder*	*Cocaine base*	*Heroin of low purity*	*Heroin of high purity*	*Cocaine powder*	*Cocaine base*	*Heroin of low purity*	*Heroin of high purity*
Price per pure gram					*United States dollars*			
1kg+/1kg+	0.226	0.363	—	—	22	34	—	—
125-1,000g/100-1,000g	0.228	0.235	0.365	0.306	32	50	396	301
35-125g/10-100g	0.216	0.188	0.338	0.311	52	51	614	440
4-35g/1-10g	0.253	0.234	0.339	0.318	69	71	1 166	785
1-4g / 0.5-1g	0.300	0.268	0.353	0.348	132	125	1 507	892
<1g / <0.5g	0.395	0.329	—	—	209	167	1 235	1 005
Purity					*Percentage*			
1kg+/1kg+	0.063	0.283	0.205	0.185	89	64	74	83
125-1,000g/100-1,000g	0.099	0.176	0.342	0.248	87	75	51	62
35-125g/10-100g	0.139	0.140	0.415	0.312	84	82	28	47
4-35g / 1-10g	0.156	0.122	0.475	0.289	81	85	20	55
1-4g / 0.5-1g	0.171	0.099	0.441	0.369	81	88	20	36
<1g / <0.5g	0.173	0.095	0.444	0.289	83	89	14	43

*The moderate levels of heroin purity observed at high wholesale levels contradicted the orthodoxy about heroin markets. Since bulk is a principal source of law enforcement risk, high-level dealers have an incentive to minimize exposure by transacting in high-purity drugs. Yet there was clear evidence of "cutting" close to the import levels of the heroin trade. In the data analysed, cities with high-purity heroin, 10 per cent of observations involving 1 kilogram or more had a purity level of less than 27 per cent. This has also been reported in recent studies of markets in Frankfurt and Milan [35] and in Britain [36]. Cocaine shows much less purity variation at this level; the 10th percentile is 74 per cent and the 90th percentile is 93 per cent. On the other hand, except for the first two levels of low-purity heroin distribution, the drop in median purity was never consistent with an image of most dealers cutting drugs with 1 unit of diluent per unit of drugs, or even 1 unit of diluent per 2 or 3 units of drugs. Dilution is, on average, much less extreme.

Curiously, the median purity of cocaine base increased as it moved down the distribution chain and also became significantly less variable. That might be the result of STRIDE not differentiating between crack per se and other forms of base, such as the base produced as an intermediate product at the source, the Andean subregion, which might be more common at higher market levels.

What was striking, however, was that, despite the vast differences in the total dollar expenditure at various levels in the distribution chain, the amount of price dispersion at each market level was roughly the same for heroin in both "low-purity" and "high-purity cities". The price per pure gram was as variable at the level of 100-1,000 grams as for quantities of 20-1,000 grams. For both forms of cocaine there was more variation at the retail level than at higher levels, but not enormously more.

Relationship between variability of purity and pure gram prices

The authors hypothesized that demand for an illicit drug would be negatively related to the variability of its purity, since unobserved quality variation would create undesirable uncertainty. For retail quantities of each drug type and combination of year and city, the authors calculated the median, mean and standard deviation of purity and price (or log price) and then the aggregate coefficient of variation. Using a fixed effect estimator to control for city and excluding city years with fewer than 10 observations, the authors estimated the effect of median purity level and variability of purity on expected median of pure gram price (and log pure gram price).

It was found that median price in a given city and year was negatively related to median purity. For quantities at retail levels, increasing purity decreased the pure gram equivalent price, confirming the conventional wisdom.*

The effect of purity variability was to increase pure gram prices, contrary to expectations. The same model was estimated separately for cocaine powder, crack cocaine and heroin because they have varying degrees of purity dispersion, heroin and crack being at the extremes. The relationship of purity variability on price was still positive but more significantly attenuated for the much more highly variable heroin than for the relatively reliable crack supply. Perhaps where purity is highly variable, suppliers are more able to get by selling lower-quality products, which are effectively higher priced in pure gram equivalents, and this swamps any tendency for variability in quality to lessen demand.

Discussion: why markets for illegal "lemons" survive

Perhaps the most fundamental question raised by heroin and cocaine markets is how they survive at all in the face of such great uncertainty about product quality and price. Given all the factors that allow for cheating in any individual

---

*For heroin there may be a mechanical element to the relationship between purity and price. It is difficult to handle quantities smaller than 50 milligrams. If the price of heroin is $1,000 per gram, a "dime bag" (a bag that costs $10) may contain only 10 milligrams of heroin and thus requires a purity of no less than 20 per cent.

transaction, including imprecise assessment of the quality after consumption, why are there not more frequent "rip-offs", particularly in a data set such as STRIDE, which is primarily composed of purchases made in the context of new, as opposed to long-standing, customer-supplier relationships? Clearly "rip-offs" are technologically possible. STRIDE contains zero-purity observations, and Simon and Burns even describe sellers who specialize in fraud [26, p. 69]: "They stand where they want, sell what they want, and risk only the rage of their victims or in a rare instance, the ire of a street dealer whose business reputation suffers from proximity". Likewise, buyers can sometimes "rip off" sellers by stealing the drugs without paying. Why does such behaviour not become so common that it destroys the market, as in the classic problem of the "lemons"?

One hypothesis, oddly enough, is that, even among criminals, trust may be the critical factor. Despite the high rates of turnover, these are markets in which repeat business is the norm and is highly valued.

The following model with plausible parameters illustrates this point. It is assumed that an addict makes twice-daily purchases, say, 600 in the course of a year. It is also assumed that, consistent with Riley [23], the addict has 15 suppliers. The buyer then purchases an average of 40 times annually from each seller. By a simple Bernoulli model, even if there is a one-third probability of each buyer and seller exiting the market in the course of the year (reflecting the cumulative effects of incarceration, ill health and violence from other participants), the probability that this dyad, after making a purchase, will transact again within a year is 0.98.* In this important respect, illicit drug transactions differ from the classic "lemons" market in which each purchase is the sole transaction involving that seller. This high probability of a "repeated game" may be sufficient to induce cooperation.

In the classic "repeated game" model, when players cooperate the interactions are of uniformly high quality. However, in this case, cooperation does not mean always selling high-quality drugs because the seller has "imperfect" knowledge and, hence, "imperfect" control over quality. Sellers can decide to "rip off" someone, but they cannot simply decide to sell high-quality drugs. It is physically impossible for them to sell drugs with a purity that is higher than that of the drugs they receive from their supplier, and they have at best imperfect knowledge of when that purity is substandard.

Hence, even a cooperating seller provides a wide range of qualities. (Similarly, even a cooperating buyer reports a range of experiences for transactions of a given quality.)** Precisely because the buyer makes so many purchases from the same seller, no single transaction is taken to provide much information about quality or cooperativeness. The seller may aim to provide, in the course of the 10 transactions that occur in a quarter, a distribution of quality and price that is consistent with that in the market in general, but he

*$(0.98)^{40} = 0.446$, which is essentially the same as $(2/3)(2/3) = 4/9 = 0.444$.

**Ethnographers report that complaints are not uncommon and that compensation is sometimes made [37].

or she will have only a very rough estimate of what that distribution is, both because there are no institutions to collect such data and because no buyer or seller can report purity and price per pure gram for a given transaction.

On the other hand, the inevitability of the dispersion in quality affords the seller the opportunity to sometimes cut the drugs a little more. (Similarly, the inevitability of quality dispersion and the seller's "imperfect" knowledge gives customers an opportunity to grumble a little more than is truly justified about poor quality.) So some dispersion is unavoidable, and that dispersion creates incentives and opportunities for occasional further dilution as the drugs move from one stage to the next in the distribution chain;* however, excessive "cutting" or outright fraud are constrained by the "repeated game" character of the transactions.

Turning this explanation on its head may also help explain another paradox of illicit drug markets. Many buyers purchase small quantities of drugs with great frequency even though there are enormous quantity discounts to be had. For example, the hypothetical addict in this example might make 600 purchases of $20 each, for a total annual expenditure of $12,000. Given the typical quantity discounts available [16], the same addict could probably purchase 12 times as much of the drug per transaction for 7 times the cost (for example, weekly purchases of $140 would yield the same amount of drugs for a total of $7,000 per year). The usual explanation given for not taking advantage of such opportunities to save 40 per cent is that addicts are cash-constrained and/or cannot be relied on to keep inventory. Both stories are entirely plausible, but a third reason may be that bundling purchases into weekly rather than twice-daily loads erodes the "repeated game" character of the transactions and creates too great an incentive for fraud.

This account of markets surviving despite extreme variability because transactions are "repeated games" is not entirely satisfactory for two reasons. First, the numbers are predicated on an image of retail purchases. Wholesale purchases show similar price dispersion even though transactions are much less frequent. Perhaps, however, the same basic story holds true because it is more expensive to search for alternative transaction partners at higher levels, so customers in high-level transactions may divide their annual number of transactions over a much smaller number of alternative suppliers. Weekly transactions spread over three suppliers still give dyadic relationships a 95 per cent probability of repeat business, even if both customer and seller face the same high risk (probability: one third) of having to exit the market in the course of the year.

The few studies published on high-level dealers (see, for example, Adler [38] and Reuter and Haaga [39]) focus on the number of customers each supplier has rather than on the number of suppliers a customer has, but conversations with experienced investigators suggest that the chains are thin in both directions.

*It is easy to tell stories about why one party's preference for immediate rewards over delayed gratification may vary; for example, the customer may be in withdrawal or either party may owe money to someone who is about to employ violence as a collection tactic.

Dealers at the middle or high level may have few potential sources and without the kind of market places available to retailers, they have strong incentives to avoid further search. Relationships are ongoing [20] and repeat transactions are expected.

The second challenge to this model is that not only addicts but also infrequent users and even first-time buyers can purchase drugs. Many first-time and infrequent buyers may purchase from friends, co-workers or others with whom they have an ongoing relationship.* Such transactions are, for practical purposes, rounds in a repeated game, even though the next round may involve who picks up a restaurant tab or whether one party lends a tool to the other.

Stereotypical anonymous street markets catering to middle-class casual users who (infrequently) drive in from the suburbs in their imported cars are not in fact the norm—depictions in the media notwithstanding—but they do exist. Perhaps anonymous place-based markets only survive where formal organizations (such as gangs) or informal norms (for example, where all sellers are from the neighbourhood and grew up together) enforce quality standards by punishing sellers who defect by offering goods of substandard quality. That is, the reputation for quality may be associated with the place and its norms, not the individuals, and some coordinating mechanism enforces compliance with the norms of that place. The available evidence does not support the view that illicit drug markets in general are highly organized, but it is less difficult to accept the idea that this one special type of market has some such organization, whether formal or informal.

Not all illicit markets have high price dispersion. Illegal numbers banks (common in cities in the north-eastern and mid-western parts of the United States prior to the introduction of state lotteries) had similar pay-out rates, typically 600 to 1 for a three-digit bet; in cities the range was 550-650 to 1. Sports bookmakers also used standard pricing schemes [40]. There may have been unmeasured quality variation in terms of the reliability of large pay-outs** among the banks and bookmakers, but that was surely slight compared with what has been reported here for cocaine and heroin. Neither market was subject to much pressure from law enforcement,*** and transactions did not have to be particularly hurried.

The high price variability observed in illicit drug markets is thus not just a function of illegality but probably the combination of illegality and a number of

---

*The authors are not aware of data on sources for those who use cocaine and heroin infrequently (as opposed to heavier users, who show up, for example, in criminal justice samples), but in the 2001 National Household Survey on Drug Abuse (now called the National Survey on Drug Use and Health) such questions are asked of cannabis users in the United States. It turns out that 80 per cent of past-year cannabis users most recently acquired their cannabis from a friend and another 9 per cent from a relative (authors' analysis).

**Numbers banks also would offer lower pay-off rates for a few frequently played numbers; there may have been variation among banks in how many numbers were "cut" and how much the pay-off was reduced for those numbers. Fewer than 50 numbers were cut and the reductions might be only 25 per cent; this still left the range of effective pay-out rates constrained.

***Arrest was a common experience for numbers sellers, but almost none of them received even brief jail terms [40].

other factors, including the characteristics of those who select into those markets (for example, short planning horizon, situational urgency), the difficulty of ascertaining product quality and law enforcement pressure.

In summary, this paper has documented a number of striking empirical regularities in cocaine and heroin markets, most notably the very substantial dispersion in purity and prices. As is often the case, new empirical information answers some questions but raises others that require further research. Cocaine and heroin markets are obviously a topic of interest to scholars and practitioners in the field of drug policy. They also present interesting opportunities for those who study market problems related to imperfect information and its effects on product quality and price, as well as dispersion in price and quality. The present article represents a first effort to explore these aspects of illicit drug markets.

References

1. George A. Akerlof, "The market for 'lemons': quality uncertainty and the market mechanism", *Quarterly Journal of Economics*, vol. 84, No. 3 (1970), pp. 488-500.
2. George Stigler, "The economics of information", *Journal of Political Economy*, vol. 69, 1961, pp. 213-225.
3. Steven Salop and Joseph Stiglitz, "Bargains and ripoffs: a model of monopolistically competitive price dispersion", *Review of Economic Studies*, vol. 44, No. 3 (1977), pp. 493-510.
4. Steven Salop, "The noisy monopolist: imperfect information, price dispersion and price discrimination", *Review of Economic Studies*, vol. 44, No. 3 (1977), pp. 393-406.
5. John Pratt, David Wise and Richard Zeckhauser, "Price differences in almost competitive markets", *Quarterly Journal of Economics*, vol. 93, No. 2 (1979), pp. 189-211.
6. Tomas von Ungern-Sternberg and Carl Christian von Weizsacker, "The supply of quality on a market for 'experience goods'", *Journal of Industrial Economics*, vol. 33, No. 4 (1985), pp. 531-540.
7. Severin Borenstein and Nancy Rose, "Competition and price dispersion in the U.S. airline industry", *Journal of Political Economy*, vol. 102, No. 4 (1994), pp. 653-683.
8. Bev Dahlby and Douglas S. West, "Price dispersion in an automobile insurance market", *Journal of Political Economy*, vol. 94, No. 2 (1986), pp. 418-438.
9. Howard Marvel, "The economics of information and retail gasoline price behavior: an empirical analysis", *Journal of Political Economy*, vol. 84, No. 5 (1976), pp. 1033-1060.
10. Michael Smith, Joseph Bailey and Erik Brynjolfsson, "Understanding digital markets: review and assessment", *Understanding the Digital Economy*, Erik Brynjolfsson and Brian Kahin, eds. (Cambridge, Massachusetts, MIT Press, 2000), pp. 99-136.
11. Eric Bond, "A direct test of the 'lemons' model: the market for used pickup trucks", *American Economic Review*, vol. 72, No. 4 (1982), pp. 836-840.

12. Eric Bond, "Test of the lemons model: reply", *American Economic Review*, vol. 74, No. 4 (1984), pp. 801-804.

13. Michael Pratt and George Hoffer, "Test of the lemons model: comment", *American Economic Review*, vol. 74, No. 4 (1984), pp. 798-800.

14. Jonathan P. Caulkins and others, *The Price and Purity of Illicit Drugs: 1981 through the Second Quarter of 2003* (Santa Monica, California, RAND, 2004).

15. P. Reuter, R. MacCoun, and P. Murphy, *Money from Crime: a Study of the Economics of Drug Dealing in Washington, D.C.* (Santa Monica, California, RAND, 1990).

16. Jonathan P. Caulkins and Rema Padman, "Quantity discounts and quality premia for illicit drugs", *Journal of the American Statistical Association*, vol. 88, No. 423 (1993), pp. 748-757.

17. Jonathan P. Caulkins and Peter Reuter, "What price data tell us about drug markets", *Journal of Drug Issues*, vol. 28, No. 3 (1998), pp. 593-612.

18. T. Wendel and R. Curtis, "The heraldry of heroin: 'dope stamps' and the dynamics of drug markets in New York City", *Journal of Drug Issues*, vol. 30, No. 2 (2000), pp. 225-260.

19. N. Zinberg, *Drug, Set and Setting: the Basis for Controlled Intoxicant Use* (New Haven, Connecticut, Yale University Press, 1984).

20. Joseph Fuentes, "The life of a cell: managerial practice and strategy in Colombian cocaine distribution in the United States", Ph.D. thesis, City University of New York, 1998.

21. K. J. Riley, *Crack, Powder Cocaine, and Heroin: Drug Purchase and Use Patterns in Six U.S. Cities* (Washington, D.C., National Institute of Justice and Office of National Drug Control Policy, 1997).

22. M. A. R. Kleiman, "Enforcement swamping: a positive-feedback mechanism in rates of illicit activity", *Mathematical and Computer Modelling*, vol. 17, No. 2 (1993), pp. 65-75.

23. Robert Axelrod, *The Evolution of Cooperation* (New York, Basic Books, 1984).

24. Jonathan P. Caulkins, "The evolution of drug initiation: from social networks to public markets", *Optimization, Dynamics and Economic Analysis: Essays in Honor of Gustav Feichtinger*, E. Dockner and others, eds. (Heidelberg, Physica-Verlag, 2000), pp. 353-367.

25. P. J. Goldstein and others, "The marketing of street heroin in New York City", *Journal of Drug Issues*, vol. 14, No. 3 (1984), pp. 553-566.

26. D. Simon and E. Burns, *The Corner: a Year in the Life of an Inner-City Neighborhood* (New York, Broadway Books, 1997).

27. Richard S. Frank, "Drugs of abuse: data collection systems of DEA and recent trends", *Journal of Analytical Toxicology*, vol. 11, No. 6 (1987), pp. 237-241.

28. C. Manski, J. Pepper and C. Petrie, eds., *Data Needs for Monitoring Drug Problems* (Washington, D.C., National Academy Press, 2001).

29. Joel L. Horowitz, "Should the DEA's STRIDE data be used for economic analyses of markets for illegal drugs?", *Journal of the American Statistical Association*, vol. 96, No. 456 (2001), pp. 1254-1271.

30. Jonathan P. Caulkins, "Can supply factors suppress marijuana use by youth?", *Drug Policy Analysis Bulletin*, No. 7, June 1999, pp. 3-5.

31. Jonathan P. Caulkins, "Drug prices and emergency department mentions for cocaine and heroin", *American Journal of Public Health*, vol. 91, No. 9 (2001), pp. 1446-1448.
32. W. Rhodes, R. Hyatt and P. Scheiman, "The price of cocaine, heroin and marijuana, 1981-1993", *Journal of Drug Issues*, vol. 24, No. 3 (1994), pp. 383-402.
33. "No minor mix-up", *Los Angeles Times*, 10 January 2000.
34. Jonathan P. Caulkins, *Developing Price Series for Cocaine* (Santa Monica, California, RAND, 1994).
35. L. Paoli, *Drug Markets in Frankfurt and Milan* (Freiburg, Germany, Max Planck Institute, 2000).
36. G. Pearson and D. Hobbs, *Middle Market Drug Distribution*, Home Office Research Study 227 (London, Home Office, 2001).
37. H. Brownstein, S. Crimmins and B. Spunt, "A conceptual framework for operationalizing the relationship between violence and drug market stability", *Contemporary Drug Problems*, vol. 27, 2000, pp. 867-890.
38. Patricia Adler, *Wheeling and Dealing: an Ethnography of an Upper-Level Drug Dealing and Smuggling Community* (New York, Columbia University Press, 1985).
39. Peter Reuter and John Haaga, *The Organization of High-level Drug Markets: an Exploratory Study* (Santa Monica, California, RAND, 1989).
40. Peter Reuter, *Disorganized Crime* (Cambridge, Massachusetts, MIT Press, 1983).

UNITED NATIONS
Office on Drugs and Crime

Vienna International Centre, P.O. Box 500, 1400 Vienna, Austria
Tel: (+43-1) 26060-0, Fax: (+43-1) 26060-5866, www.unodc.org

Printed in Austria
V.05-87390—January 2006—2,015

United Nations publication
Sales No. E.06.XI.6
ISBN 92-1-148211-9
ISSN 0007-523X

9 789211 482119