1-1-1996

Exploring the Underground Economy: Studies of Illegal and Unreported Activity

Susan Pozo, Editor
Western Michigan University

Follow this and additional works at: https://research.upjohn.org/up_press

Part of the [Labor Economics Commons](#)

Citation

Pozo, Susan, ed. 1996. Exploring the Underground Economy: Studies of Illegal and Unreported Activity. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. <https://doi.org/10.17848/9780880994279>

This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 International License](#).

This title is brought to you by the Upjohn Institute. For more information, please contact repository@upjohn.org.

***Exploring
the***

**UNDERGROUND
ECONOMY**

Susan Pozo, Editor

Exploring the Underground Economy

Studies of Illegal and Unreported Activity

Susan Pozo
Editor

1996

**W.E. Upjohn Institute for Employment Research
Kalamazoo, Michigan 49006-4686**

Library of Congress Cataloging-in-Publication Data

Exploring the underground economy : studies of illegal and unreported activity / Susan Pozo, editor.

p. cm.

Includes bibliographical references and index.

Six essays based on presentations made at the 30th annual Public Lecture Series sponsored by the Department of Economics at Western Michigan University.

ISBN 0-88099-165-8 (pbk. : alk. paper). — ISBN 0-88099-166-6 (cloth : alk. paper)

1. Informal sector (Economics)—United States—Congresses.
2. Informal sector (Economics)—Taxation—United States—Congresses.
3. Tax evasion—United States—Congresses. I. Pozo, Susan.

HD2346.U5E96 1996

330—dc20

95-50921
CIP

Copyright © 1996

W. E. Upjohn Institute for Employment Research
300 S. Westnedge Avenue
Kalamazoo, Michigan 49007-4686

The facts presented in this study and the observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of the W. E. Upjohn Institute for Employment Research.

Cover design by J. R. Underhill

Index prepared by Shirley Kessel.

Printed in the United States of America.

PREFACE

The papers in this volume were presented during the thirtieth annual Public Lecture Series sponsored by the Department of Economics at Western Michigan University. The Lecture Series was made possible through the financial support of the W.E. Upjohn Institute for Employment Research and the College of Arts and Sciences of Western Michigan University.

Several of my colleagues in the Economics Department and at the Upjohn Institute assisted with the progress of the Lecture Series, including Raymond E. Zelder, William D. Bosshardt (now at Florida Atlantic University), and Timothy J. Bartik. I thank Randall W. Eberts and Robert G. Spiegelman of the Upjohn Institute for their continued support of the Series. Judith K. Gentry and Natalie M. Lagoni produced the book expeditiously. I am grateful to all for their help, as well as to the authors for their cooperation.

Susan Pozo
Kalamazoo, Michigan
January 1996

Contents

Introduction	1
Susan Pozo <i>Western Michigan University</i>	
Overseas Holdings of U.S. Currency and the Underground Economy	5
Edgar L. Feige <i>University of Wisconsin—Madison</i>	
The Mismeasurement of Illegal Drug Markets	63
Peter Reuter <i>University of Maryland</i>	
The Supply of Youths to Crime	81
Richard B. Freeman <i>Harvard University</i> and <i>National Bureau of Economic Research</i>	
Explaining Tax Compliance	103
James Alm <i>University of Colorado at Boulder</i>	
Beating the System?	129
Ann Dryden Witte <i>Florida International University</i> and <i>Wellesley College</i>	
The Informal Economy	147
Alejandro Portes <i>Johns Hopkins University</i>	

Introduction

Susan Pozo

Western Michigan University

Ask an academician, a government worker, a policy analyst, a politician, and a member of the general public what the underground economy is all about and you will surely get a wide variety of answers. This diversity is reflected in this volume, where six different papers about the underground economy—its many facets, its various consequences, and a multitude of interpretations—are presented.

Our fascination with the underground economy is as many-faceted as are our definitions of it, but it is safe to assume that we are intrigued with the underground economy in part because it is difficult to directly observe and directly measure. Whereas one can be reasonably assured of estimates of automobile production in the United States in a given year, we are less confident of the estimates offered for tax compliance, domestic marijuana consumption, the volume of illegal gambling, or the extensiveness of barter trade. Hence we dig deeper and ask, how were these numbers obtained? If alternative assumptions are made, how would the estimates change? Will a different methodological approach confirm or contradict these estimates?

Because of its nature, because participants in the underground economy hide their participation, direct measurement of the magnitude of underground economic production and participation is usually not possible. We normally have to find other methods, often relying on indirect measures to find traces of underground economic activity. In this book, alternative methods are used to track the underground economy, ranging from monetary measures to direct survey methods to experimental techniques.

The chapters by Edgar Feige and by Peter Reuter focus on measuring the size and volume of underground economic activity. Feige uses monetary techniques to measure the volume of unreported economic activity taking place. Previous studies estimating the size of the underground economy, which is presumed to bear some relationship to the level of unreported income, have used the currency ratio method. In

this method increases in the currency-to-checkable-deposit ratio are assumed to be associated with increases in unreported income because individuals find it easier to camouflage unreported income if it is kept as cash. However, Feige notes that the estimates offered thus far using the currency-to-deposit ratio or some variant of it have assumed that all U.S. currency circulates domestically. We are continually reminded, however, that U.S. currency plays an important role in many foreign countries, as evidenced in part by its use in black and parallel-dollar-currency markets. Feige extensively examines what is known about U.S. currency holdings abroad to derive estimates of the proportion of U.S. currency that circulates domestically. With currency estimates taking into account the volume of overseas holding of U.S. currency, Feige is able to obtain and report estimates of the size of the underground economy using a variant of the currency-to-deposit ratio method that better reflects U.S.-only underground economic activity.

Reuter, on the other hand, takes a hard look at a specific subsector of the underground economy—the illegal drug market. This market is, to the best of our knowledge, the largest of all the illegal markets, and thus its size is of concern to those interested in learning about the extensiveness of the underground economy.

Reuter suggests that U.S. government estimates pertaining to the size of this market are grossly inadequate and inconsistent. He cleverly shows how individually scrutinized estimates can be shown to be implausible and how the collective estimates of related markets or variables within the market are inconsistent. Why are such poorly engineered figures offered to the public? Overall, they do not offer guidance for good policy making and hence do not result in alterations, improvements, or development of drug policy. They are often offered instead to legitimize policy decisions that are derived independent of good estimates.

Reuter notes that good estimates are important and that they could serve a valuable function in evaluating, for example, the effectiveness of policies to reduce illegal drug consumption or the scope of drug markets.

The papers by Richard Freeman and James Alm are concerned with criminal behavior of individuals. Freeman concentrates on youth, in an attempt to gain an understanding of what leads to criminal behavior. What policies could be implemented if we wish to reduce criminal

activity among our youth? Since criminal activity during youth seems to be associated with subsequent behavior and economic attainment, it is especially important to understand how our current economic policies may encourage or deter youth from engaging in crime. Freeman obtains data from the National Longitudinal Survey of Youth to undertake his investigation and arrive at conclusions regarding participation in underground activities.

Alm looks at a different category of criminal activity—tax evasion. He too asks how economic policy affects the degree of tax cheating that takes place. To get around the problem of getting individuals to truthfully admit the degree to which they cheat on taxes, Alm employs experimental methods to track tax compliance. Under these carefully controlled methods and with careful experimental design he is able to generalize how changes in the variables in the hands of policy makers can alter behavior, causing the general compliance rate to rise or causing tax cheating to increase.

The papers by Ann Witte and Alejandro Portes serve as an overview of the underground economy in the Americas. Witte's paper points to some of the ambivalent attitudes we have had regarding the underground economy in the United States over the past several decades. In many cases we argue that certain behaviors and participation in the underground economy are not acceptable. We do not always seem entirely convinced of this position, however, and we end up engaging in less enforcement. Witte suggests that what constitutes the underground economy has changed over time. Variations in enforcement serves as a precursor to real changes in institutions and laws pertaining to underground activity. There is significant fluidity in the definition of what is legal and what is not, and participation in the underground economy does not necessarily mean that one is "beating the system."

Portes adds another dimension or perspective to the underground economy discussion. This results from his studies of underground economies in Latin America. While there is reasonable consensus of what constitutes the underground economy in the industrialized economies, the consensus is less broad for those who study developing economies. This may be because of the more diverse form taken by the underground economy in less-developed countries, and perhaps because there is a less clear distinction between what we term formal and informal production in these economies. The linkages between the

underground and aboveground economies are stronger in developing economies, making for more difficult delineation of the two categories of production.

Despite the additional challenges faced by researchers who measure the underground economy for developing nations, Portes offers estimates of the magnitude of this sector for many of the Latin American economies over the past several decades. He employs several different methodologies and concludes that there is considerable informal production taking place in these countries.

Though the underground economy is a somewhat elusive and variously defined concept, these studies suggest that a great deal is known. By using various approaches, economists and others have actually learned much about unrecorded and illegal production. Despite the diversity of approaches, many common conclusions seem to emerge. First, the underground economy is significant in size. Second, growth and persistence in the underground economy are not relegated to a particular historical period, geographic region of the world, or stage of development. Both developed and developing economies seem to be faced with significant underground sectors. Third, economic policy can play a significant role in determining whether economic agents participate to a greater or lesser degree in the underground economy.

Nonetheless, despite knowing a great deal about the underground economy, this collection of papers also brings to our attention that much has yet to be learned. Because of the diverse nature of underground economic activities, a variety of approaches is probably the best route to better understand the underground. Unlike many economic inquiries, it may be more appropriate and fruitful to employ economists (and noneconomists) from almost all the different subspecialties to add to our body of knowledge about the underground economy. Labor economics, macroeconomics, public finance, criminology, experimental economics, urban economics, and international economics all have contributed and have more contributions to make so that we can better understand the underground economy. The studies presented here attest to this.

Overseas Holdings of U.S. Currency and the Underground Economy

Edgar L. Feige

University of Wisconsin—Madison

Many public policy decisions require analytical and empirical knowledge concerning the size, growth, causes, and consequences of the “underground economy.” This paper seeks to clarify the meaning of underground activity, updates various discrepancies and fiscal estimates of its size and growth, and examines the empirical implications of new evidence concerning the growing use of U.S. currency throughout the world for indirect monetary estimates of the underground economy in the United States.

The popular term “underground economy” is inexact, covering a wide range of economic activities that encompass the production and distribution of illegal goods and services, as well as legal activities whose concealment from or misrepresentation to governmental authorities involves tax evasion or benefit fraud. Given the diversity of hidden activities, it is necessary to develop a taxonomy of “underground economies” that identifies specific types of underground behaviors and suggests appropriate methods for estimating their size and their unique economic implications. The general penchant to hide underground economic activities often precludes direct observation of their occurrence, necessitating the use of indirect measures that seek to uncover the footprints of hidden activities left behind in the sands of observable economic indicators. Currency, being an anonymous medium of exchange, is regarded as the preferred means of payment for economic transactions that economic actors seek to conceal. As such, cash stocks and flows are a natural starting point in the search for knowledge concerning the underground economy. The total amount of currency in circulation is also one of the best-measured macroeconomic indicators, since the production and distribution of currency by the government is strictly monitored and carefully recorded.¹ However, our knowledge concerning the location and circulation of the public’s holdings of U.S.

currency is meager. There are no reliable estimates of the varying amounts of U.S. currency circulating overseas, therefore no way of determining the domestic money supply and its change over time. This paper develops alternative means of calculating the proportions of currency held domestically and overseas and presents estimates of net outflows of U.S. currency over time. New estimates of domestic holding of U.S. currency are then used to reestimate the size and growth of the domestic underground economy.

One of the most puzzling macroeconomic anomalies is the huge amount of U.S. currency outstanding (\$390 billion) and its surprisingly persistent growth. Despite widespread predictions of the advent of the cashless society, and decades of cash-saving financial innovations, the *per capita* holdings of U.S. currency increased from \$160 in 1961 to \$1,450 by the end of 1994. Adjusting for inflation, real per capita currency increased by 70 percent, and the proportion of the M1 money supply composed of currency rose from 20 percent to 30 percent. Roughly 60 percent of the outstanding stock of currency is now in the form of \$100 bills.

The suggestion that the average American family of four now holds \$5,800 in currency, of which \$3,480 is in the form of \$100 bills, appears to be implausible. Moreover, since the average turnover rate of currency is estimated to be fifty times per year, the average American would have to be making *pro rata* cash transactions of \$72,500 per year, a figure which is plainly unbelievable. Federal Reserve surveys (Avery et al. 1986, 1987) of currency usage by American households determined that adult U.S. residents admit to holding only 12 percent of the nation's currency in circulation outside the banking system. Allowing for U.S. business holdings of currency, the whereabouts of more than 80 percent of the nation's currency supply is presently unknown. These anomalous findings give rise to the "currency enigma" (Feige 1990b, 1994a), which consists of a stock and a flow component. Our inability to identify the holders and location of a large fraction of the U.S. currency stock gives rise to a \$300 billion problem of "missing currency" (Sprenkle 1993). This missing stock of currency is used as a means of payment for the purchase of goods and services. If all of the missing currency turns over at the same rate as that estimated for U.S. households, it would effect a flow of "missing payments" amounting to almost \$15 trillion.

Two complementary hypotheses are put forward as possible explanations of the currency enigma. Some fraction of the missing currency may in fact be held by U.S. households to conduct unreported transactions in the U.S. underground economy. A considerably larger portion of the missing currency is more likely to be held abroad by foreigners who conduct transactions with U.S. dollars to effect payments in their own countries and as a store of value. This paper examines the extent to which the currency enigma can be resolved by appeal to both the underground economy hypothesis and the “world dollarization” hypothesis.

The first section of the paper presents a taxonomic framework for defining different types of underground activities, reviews alternative methods of estimation, and updates available estimates of various “underground economies” in the United States.² The following three sections describe alternative methods for estimating the amount of U.S. currency held abroad and present preliminary estimates of overseas currency holdings. These sections present a monetary demographic model (MDM) and a note ratio model (NRM) that can be employed to obtain indirect estimates of the proportions of currency held domestically and overseas, as well as direct estimates of net outflows of U.S. currency. To anticipate the results, a factor model composite measure of overseas currency flows suggests that foreigners do appear to hold a large fraction of U.S. currency, perhaps as much as 40 percent. The final section examines the implications of overseas currency holdings for the measurement of the domestic underground economy.

Defining and Measuring Underground Economies

The early literature on the underground economy lacked an accepted taxonomy for classifying various underground activities. Underground activities were variously described as subterranean, irregular or informal, hidden, grey, shadow, clandestine, parallel, and black, but these descriptions were rarely augmented with explicit definitions that aided analytic and empirical investigation of the underlying phenomena. It is now well understood that there are a variety of underground economies spanning both planned and market economies, be they developed or

developing. Agents engaged in underground activities circumvent, escape, or are excluded from the institutional system of rules, rights, regulations, and enforcement penalties that govern formal agents engaged in production and exchange. Different types of underground activities are distinguished according to the particular institutional rules they violate. Employing this criterion, we identify four specific types of underground economic activities: illegal; unreported; and unrecorded; and informal. The metric for measuring the dimensions of each underground activity is the aggregate income generated by the activity. Figure 1 presents a taxonomy of underground economies.

The *illegal* economy consists of the income produced by those economic activities pursued in violation of legal statutes defining the scope of legitimate forms of commerce. The most notable illegal activities are the production and distribution of prohibited substances (e.g., drug traffic) and services such as prostitution, pornography, and black market currency exchange. Estimates of income produced from illegal activities are typically derived from crime-related statistics and range from \$70 to \$100 billion. In 1982, unreported income from drugs and gambling was estimated to be roughly \$26 billion (Abt Associates 1984, pp. 62, 112), and the retail value of drugs sold in the United States in 1990 was estimated to be roughly \$40 billion (Office of National Drug Control Policy 1991, p. 5).

The *unreported* economy consists of those economic activities that circumvent or evade fiscal rules as codified in the tax code. A summary measure of the unreported economy is the amount of unreported income, namely, the amount of income that should legally be reported to the tax authority but is not so reported. Since illegal income is taxable, the unreported economy includes both legal and illegal source income that is not properly reported to the fiscal authority. A complementary measure of the unreported economy is the "gross tax gap," namely, the difference between the amount of tax revenues legally due the fiscal authority and the amount of tax revenue voluntarily paid.

The "net tax gap" represents the difference between the amount of revenue due and the amount actually collected. The difference between the gross and net measures represents the revenues collected as a direct result of enforcement activities. Benefit fraud, comprising false claims for benefits (welfare or unemployment payments) or subsidies to

Figure 1.

which the individual is not legally entitled, should formally be included in “tax gap” measures.

The *unrecorded* economy consists of those economic activities that circumvent the institutional rules that define the reporting requirements of government statistical agencies. A summary measure of the unrecorded economy is the amount of unrecorded income, namely, the amount of income that should (under existing rules and conventions) be recorded in national accounting systems (e.g., National Income and Product Accounts) but is not so recorded. Unrecorded income represents a discrepancy between total income or output and the actual amount of income or output captured or enumerated by the statistical accounting system designed to measure economic activity. Since national accounting conventions differ with respect to their inclusion of illegal incomes, unrecorded income may or may not include components of the illegal sector.

The term “informal economy” has been used so frequently, and inconsistently, in the development literature that it requires special attention. The *informal* economy comprises those economic activities that circumvent the costs and are excluded from the benefits and rights incorporated in the laws and administrative rules covering property relationships, commercial licensing, labor contracts, torts, financial credit, and social security systems. A summary measure of the informal economy is the income generated by economic agents who operate informally.

Estimating the size of various underground economies remains, at best, an inexact science. However, more precise definition of alternative underground economies has reduced the tendency to compare disparate measures, and improvements in tax compliance and monetary methodologies is narrowing the range of comparable estimates.

Updated Estimates of Unreported Income in the United States

Since underground economic activity typically subjects the participant to the risk of some form of sanction if discovered, agents engaged in the activity have an incentive to conceal their involvement. This propensity for secrecy creates special problems for the social science observer who attempts to measure underground behavior. A variety of direct and indirect measurements of various types of underground

activity have been proposed and each has well-known limitations (Feige 1989). Earlier empirical efforts to measure the size and growth of underground activities revealed that underground economies were large enough to be of economic significance and had grown considerably during the later half of the 1960s and throughout much of the decade of the 1970s. Costly regulation, rising tax rates, and a growing distrust of government were cited as the primary causes of increased underground activity. The conservative politics of the 1980s sought to reverse these trends by reducing government regulations, decreasing the burden of taxation, and restoring a greater sense of trust and confidence in the government by reforming the tax system and reducing what were perceived to be wasteful government expenditures. One of the questions we seek to examine is whether these efforts had any effect on reducing the size and growth of the underground economy.

Various macroeconomic measures have been mentioned as possible indicators of underground activities. These include the adjusted gross income (AGI) gap discrepancy measure produced by the Bureau of Economic Analysis (BEA); the audit-based discrepancy measures of unreported taxable income produced by the Internal Revenue Service (IRS); and estimates of unreported income derived from various specifications of monetary models. These measures are reviewed and updated.

Discrepancy Measures

The U.S. government produces two discrepancy measures that are often cited as indicators of underground activity. The first of these, compiled by the Bureau of Economic Analysis, calculates the discrepancy between adjusted gross income as reported to the Internal Revenue Service and an independent estimate of AGI derived from National Income and Product Accounts (NIPA) estimates of personal income. This AGI gap is not officially acknowledged as a measure of the underground economy; however, with several qualifications (Carson 1984; Feige 1989), the AGI gap can be interpreted as a lower bound measure of noncompliance in the reporting of taxable income.

Figure 2 displays the AGI gap estimates published by the BEA in 1985 and the most recently revised estimates. The latest government figures reveal that the earlier gap estimates had been much too low,

requiring upward revision of \$115 billion in 1983. By 1992, the AGI gap had risen to \$500 billion. As a percentage of AGI, the gap reached its peak of 16.1 percent in 1987 and then fell to an estimated 14 percent of AGI in 1992.

Figure 2. Adjusted Gross Income (AGI) Gap
Bureau of Economic Analysis Estimates

An alternative discrepancy measure of unreported income is prepared by the IRS on the basis of their Taxpayer Compliance Measurement Program (TCMP). Responding to reports of a large underground economy based on monetary estimation methods, the IRS undertook a series of studies (IRS 1979, 1981, 1983) to examine the extent of non-compliance with U.S. tax laws. Their first study (IRS 1979) concluded that in 1976, between \$75 and \$100 billion of legal source income was not properly reported on individual tax returns. The IRS estimated that the resulting revenue loss was between \$12 and \$17 billion. In addition, unreported illegal source income was estimated to be between \$25 and \$35 billion, with an added revenue loss of \$6 to \$8 billion. In their 1983 report, the estimate for legal source unreported income in 1976 was increased by \$30 billion and the associated estimated loss of

tax revenue was more than double the initial estimate. On the other hand, the 1983 report slashed the estimate of illegal source income to only \$13 billion and cut the corresponding revenue loss from the illegal sector to roughly \$4 billion.

Feige (1989) demonstrated the sensitivity of the results from the early TCMP studies to small variations in the questionable set of assumptions required to estimate magnitude of noncompliance. The admission by the IRS that 1981 total unreported income amounted to some \$283 billion and a corresponding revenue loss of \$90 billion led the BEA to undertake a major revision of the NIPA accounts. The BEA's 1985 "comprehensive revision" raised estimated personal income for 1984 by \$100 billion. Although their comprehensive revision included changes in definitions and statistical methods, the single most important element of the revision was the adjustment for income previously unrecorded due to understated tax source data. For 1984, the personal income adjustment for unrecorded wages, salaries, and nonfarm proprietor incomes amounted to \$101 billion, demonstrating the empirical connection between unreported and unrecorded income.

The latest IRS estimates of unreported income (IRS 1988) are based on their TCMP audits of tax returns in the years 1973, 1976, 1979, and 1982 and include estimates of unreported income and the corresponding losses in tax revenue projected out to 1992. The IRS (1988) estimates of unreported legal source income for the years 1973-1982 are presented in figure 3, along with their projected unreported income for the years 1983-1992.

In each of the audit years, a sample of roughly 55,000 tax filers was subjected to examination by IRS auditors who attempted to determine the amounts of income that should have been reported and the amounts that were unreported. Final estimates of unreported income of filers and nonfilers for those years were obtained by combining information from audits, information returns, and special surveys. The IRS projections for the period 1985-1992 were based on Office of Management and Budget forecasts of personal income combined with an assumption of constant rates of noncompliance between 1982 and 1992. The projections also assume that tax payer behavior was unaffected by the tax reforms enacted in 1986. Figure 4 depicts the discrepancy between projected AGI used in the 1988 IRS report and the actual AGI reported on subsequent tax returns.

**Figure 3. Unreported Legal Source Income
Internal Revenue Service Estimates**

**Figure 4. IRS Projection Error in AGI
Projected - Actual AGI**

By 1992, actual reported AGI fell more than \$500 billion short of the IRS projections. The overestimates of projected reportable income and the assumption that compliance rates were unaffected by tax cuts and tax reforms suggest that the IRS projections of unreported income are overstated.³

Whereas the earlier IRS studies reported estimates of both legal and illegal source unreported income, the 1988 study is limited to estimates of unreported legal source income. The IRS (1983) study estimated that in 1981, illegal source income amounted to \$34.2 billion, or roughly 15 percent of the revised legal source estimate for 1981. If illegal income remained at roughly the same percentage as legal income, it would add an additional \$88 billion of unreported illegal source income to the estimate of \$585 billion of unreported legal source income for 1992.⁴

Figure 5 reports alternative IRS estimates of the “gross tax gap” on legal source income of individuals and corporations. The gross tax gap

**Figure 5. Estimates of the Gross Tax Gap
Legal Source Income**

overstates the amount of revenue lost to the government due to non-compliance to the extent that IRS enforcement activities collect some

of the amounts due. The yield from these enforcement activities was estimated to be \$15.4 billion in 1981, \$18.9 billion in 1984, and \$21.9 billion in 1987 (IRS 1990, table 2, p. 10). The gross tax gap understates the loss of revenue to the government because it excludes revenue lost on illegal source income as well as revenue losses from noncompliance with other federal taxes including employment, excise, gift, and estate taxes and customs duties. For the year 1987, income taxes accounted for only 56 percent of federal budget receipts. Another 36 percent came from employment taxes and 5 percent from gift, estate, and excise taxes. There is virtually no information available on the revenue losses from noncompliance with these other important revenue sources, nor are there estimates of the amount of government expenditures wasted as a result of benefit fraud.

Currency Ratio Models

The most common method for estimating the size of the unreported economy relies on some variant of the general currency ratio model described in Feige (1989). The most restrictive specification of the currency ratio model (Cagan 1958; Gutmann 1977) assumes that currency is the exclusive medium of exchange for unreported transactions, that the ratio of currency to checkable deposits is only affected by the growth of unreported transactions, that the income velocities of reported and unreported transactions are identical, and that in some base period, unreported income was zero so that the observed base period currency deposit ratio serves as a proxy for the desired currency ratio in the official economy.⁵

Figure 6 displays estimated unreported income as a percent of recorded AGI as obtained from the simple currency ratio model under the assumptions that in 1940 there was no unreported income and that all currency outside of the banking system was held by the domestic public. As has been noted in earlier studies, the ratio of unreported income rose sharply during World War II and then declined and remained relatively stable until the early 1960s. Unreported income then grew from less than 5 percent of AGI in 1960 to 15 percent by 1980.

The percentage of unreported income reached a plateau during the early 1980s and actually declined around the time of the 1986 Tax

Reform Act. The percent of unreported income then rose steeply between 1987 and 1991. Figure 6 also presents the results of a more general specification of the currency ratio model.⁶ The general currency ratio model (GCR) employs the IRS estimate of unreported income for 1973 as an appropriate benchmark; it assumes that 75 percent of unreported income transactions are effected by currency and that the remaining 25 percent are effected by checkable deposits. The resulting estimates display a time path similar to that of the more restrictive estimates, however the percent of unreported activities is considerably higher in all periods.

Figure 7 displays three estimates of total unreported income from both legal and illegal sources for the period 1972 through 1993.⁷ The IRS projections are remarkably similar to those obtained by the simple currency ratio model suggesting that by 1991, total unreported income amounted to roughly \$650 billion or 17 percent of reported AGI.

Assuming that this unreported income had been subject to a marginal income tax rate of 20 percent, \$130 billion of income tax revenues would have escaped government collection, roughly 62 percent of the federal budget deficit.

The GCR model results suggest that unreported income grew gradually during the first half of the 1980s, declined in the mid-1980s and then resumed its growth until the early 1990s. Unreported income appears to have doubled during the last half of the decade and exceeded \$1 trillion during the early 1990s. All of these currency ratio model estimates are predicated on the assumption that U.S. currency is exclusively used to effect domestic transactions in either the official or the underground economy. There is, however, a growing body of anecdotal evidence suggesting that U.S. currency also circulates as a medium of exchange in foreign countries. If a large and perhaps variable fraction of U.S. currency is held outside of the U.S., this would tend to overstate the size of the domestic underground economy as estimated by conventional currency ratio models.

Federal Reserve (Avery et al. 1986, 1987) Surveys of Currency and Transaction Account Usage (SCTAU) reinforce the notion that a substantial portion of U.S. currency holdings cannot be accounted for by the behavior of U.S. households. In both 1984 and 1986, SCTAU determined that U.S. households admitted to holding at most 12 percent of the nation's currency supply. Since business firms are very concerned

Figure 6. Unreported Income as Percent of AGI
Currency Ratio Models

Figure 7. Estimated Total Unreported Income
Currency Ratio Models and IRS Estimate

with efficient cash management in order to minimize interest losses associated with cash inventories, they are likely to hold considerably smaller cash inventories than households. The scant evidence on U.S. currency holdings by business firms (Anderson 1977; Sumner 1990) suggests that domestic firms hold less than 3 percent of currency in circulation.

A conservative estimate of the stock of currency required to sustain the estimated cash payments made in the unreported economy can be obtained by employing the IRS projection that total unreported income in 1992 amounted to \$675 billion. Assuming that roughly 75 percent of this unreported income is effected by cash and taking currency turnover (the income velocity of currency) to be roughly fifty turnovers per year yields a stock of currency used for underground transactions that is less than 4 percent of currency in circulation with the public.⁸ In short, since U.S. households admit to holding only 12 percent of the nation's currency in circulation, firms hold roughly 3 percent, and underground transactions absorb another 4 percent, the ownership of roughly 80 percent of circulating U.S. currency is currently unknown. This anomaly of missing currency gives rise to the stock component of the "currency enigma" (Feige 1994a). A similar problem arises with attempts to allocate the flow of payments sustained by the outstanding currency stock to different sectors of the economy. The admitted household holdings of 12 percent of U.S. currency circulating outside of the banking system give rise to an estimated volume of cash payments of \$1.7 trillion for 1992, roughly 41 percent of recorded personal consumption expenditures. Estimated cash holdings of business firms analogously give rise to some \$400 billion of intermediate payments, amounting to roughly 7 percent of total intermediate payments. An additional \$675 billion of cash payments can be allocated to underground transactions. If the stock of the remaining "missing" currency circulates at roughly the same rate as currency held by U.S. households, it would give rise to an additional volume of unaccounted cash payments in excess of \$10 trillion.

Several hypotheses have been put forward to explain these monetary anomalies. One explanation is that the U.S. underground economy is substantially larger than current estimates suggest, and that domestic holdings of U.S. currency are much larger than the amounts of currency households admit to holding on currency surveys. The more

plausible “world dollarization” hypothesis suggests that a substantial fraction of U.S. currency is held abroad by residents of other nations. The complementary “hoarding” hypothesis suggests that overseas holdings of U.S. currency are largely held as stores of value rather than being used as a medium of exchange. The dollarization hypothesis requires independent estimates of the fraction of U.S. currency held abroad. The hoarding hypothesis requires evidence confirming that overseas holdings of U.S. currency circulate at slower rates than domestic currency holdings.

Anecdotal reports of U.S. currency circulating in parts of Latin America, the Middle East, Eastern Europe, and Russia are widespread, as are suggestions that foreign demand for U.S. currency can fluctuate quite dramatically. Since the size, variability, and velocity of foreign holdings of U.S. currency have important implications for the measurement of the domestic underground economy and for the conduct of domestic monetary policy, we turn our attention to efforts to locate the “missing” U.S. currency.

Indirect Methods for Estimating Foreign Holdings of U.S. Currency

Although there is much anecdotal evidence concerning the widespread use of U.S. currency overseas, until very recently there was no systematic research undertaken to determine the location of the U.S. currency supply. Several different approaches are presently being explored to estimate both the stock of foreign holdings of U.S. currency and the migration of currency into and out of the country.⁹

The introduction in 1996 of a newly designed U.S. currency series with modern counterfeit protection provides a unique opportunity to establish a currency census system (CCS), which *while fully preserving the anonymity of individual currency usage*, could monitor key characteristics of the currency population. Relevant information on the life cycle of a currency note could be electronically captured as notes are routinely and anonymously processed by high-speed sorting machines when they are issued and returned to the Federal Reserve Banks. The proposed CCS would eliminate any burden of human

reporting. By maintaining automated records of a note's age, quality, birthplace, location, and final redemption, a CCS information system would provide the data required to construct currency migration matrices and all other demographic characteristics of the note population.¹⁰ In the absence of a location-specific enumeration of currency holdings, it is necessary to rely on direct sources of information concerning inflows and outflows of U.S. currency as well as on indirect methods for estimating the changing proportion of U.S. notes held overseas.

The Monetary Demography Model

This section develops a model capable of estimating the proportions of notes held domestically and overseas. Given estimates of these proportions and knowledge of the total stock of currency outstanding, one can estimate the overseas stock of currency and also obtain estimates of net currency outflows from the changes in foreign holdings over time. The proportion of U.S. currency circulating domestically (β^d) and the proportion circulating overseas (β^o) can be estimated from a simple monetary demography model (MDM) of the currency population.

Consider the general problem of estimating the proportions P_1 and P_2 of members in two subpopulations C_1 and C_2 which comprise the total population C . Let X represent any characteristic of the total population C , and X_1 and X_2 respectively represent the corresponding characteristic of the subpopulations C_1 and C_2 . The unknown proportions P_1 and P_2 can then be derived from the following equation:

$$(Eq. 1) \quad X = P_1 \cdot X_1 + P_2 \cdot X_2.$$

Since $P_1 + P_2 = 1$, it follows that:

$$(Eq. 2) \quad \begin{aligned} P_1 &= (X - X_2) / (X_1 - X_2) \quad \text{and} \\ P_2 &= (X_1 - X) / (X_1 - X_2). \end{aligned}$$

A meaningful solution for the parameters P_1 and P_2 exists so long as the characteristics of the subpopulations are different ($X_1 \neq X_2$), and the calculated proportions lie between 0 and 1. For example, if we know that the average weekly income (X) of the entire U.S. population is \$600 and the average weekly income of the population located east of

the Mississippi is \$800 (X_1) and the average weekly income of the population west of the Mississippi is \$500 (X_2), it follows that one-third of the population lives in the eastern portion of the country and two-thirds of the population lives in the western portion.

This simple demographic model can be generalized to find the proportion of members in any exhaustive set of subpopulations so long as the characteristics describing the subpopulations are different. In the foregoing example, the characteristic measure was simply a scalar representing the mean of the income distribution for the entire population and for each of the subpopulations. More generally, the characteristics (X , X_1 , and X_2) may be vectors representing univariate or multivariate distributions of the total and subpopulation characteristics. Examples of measurable characteristics which might be employed to estimate the MDM are the age, quality, velocity, denomination, series, or seasonal characteristics of the total and subpopulations.

Given estimates of any note population characteristic X , and the corresponding domestic (X^d) and overseas (X^o) note characteristics, the proportion of notes circulating domestically (β^d) can be estimated as:

$$(\text{Eq. 3}) \quad (\beta^d) = (X - X^o) / (X^d - X^o).$$

The MDM employing age and quality characteristics

Applying general demographic concepts to currency populations leads naturally to a consideration of possible differences in the age and quality characteristics of denomination-specific notes circulating domestically and overseas. Estimates of the age, quality, and quality by age distributions of the domestic and overseas subpopulations were obtained from a special study conducted by the Federal Reserve.¹¹ Based on a sample of some four million individual notes, note quality was ascertained by recording light reflectivity measures from an optical desitometer that scanned individual notes during routine note processing by high-speed sorting machines located at Federal Reserve Banks in each of the twelve Federal Reserve Districts. Individual note serial numbers were recorded for a subsample of approximately 150,000 notes drawn from domestic and overseas sources in order to determine the date on which each note was sent by the Bureau of Engraving and Printing to a Federal Reserve Bank. An inventory model was then employed to estimate the date the note was actually placed

into circulation, thereby establishing its date of birth. Each note's age was then determined as the difference between the sample date and the note's date of birth. Therefore, for each note denomination (\$1; \$5; \$10; \$20; \$50; \$100) it was possible to construct the univariate age and quality distributions for notes sampled domestically and overseas.

Casual observation suggests that domestic notes are likely to be used predominantly as a medium of exchange, whereas overseas notes are more likely to be held as a store of value. Thus for any denomination, it was expected that the univariate age and quality distributions of domestic and overseas notes and the corresponding bivariate quality by age distributions would differ greatly. Domestic notes, sampled on their return to the Federal Reserve, were expected to be relatively younger than overseas notes and generally of poorer quality for a given age. Given these expected differences in domestic and overseas characteristics, age, quality, and quality by age distributions were thought to be promising characteristics for estimating the proportion of notes held overseas.

Surprisingly, preliminary analysis of the quality and quality by age distributions of the domestic and overseas samples revealed that they were not sufficiently different to obtain robust estimates of the proportions of notes held domestically and overseas. Initial efforts to estimate the MDM were therefore based on differences in the univariate age distributions between overseas and domestic notes for each specific note denomination population. Denomination-specific age distributions for the entire note population were derived from FR-160 data on note births and deaths (redemptions) combined with estimates of average note lifetimes.

Employing the age characteristics of the relevant populations, the problem then is to estimate the proportion of U.S. currency circulating domestically (β^d) and the proportion circulating overseas ($\beta^o = 1 - \beta^d$) from the MDM(A) specified for each denomination as follows:

$$(\text{Eq. 4}) \quad A = \beta^d A^d + (1 - \beta^d) A^o, \text{ and}$$

$$\beta^d = (A - A^o) / (A^d - A^o)$$

where A , A^d and A^o represent the denomination-specific age distributions for the total, domestic, and overseas note populations, respec-

tively. Estimates of the proportion of notes of different denominations circulating abroad in mid-1989 were then obtained from estimates of the overall, domestic, and overseas age distribution of the notes.¹²

Table 1 presents the resulting denomination-specific estimates of the proportion of notes held overseas. The MDM model estimated for the

Table 1. Estimates of the Demographic Model
 $\{MDM[C^{den l}; C^{den l d}; C^{den l o}]; (A)\}$
Age Distribution Characteristics

Denomination	Estimated share of each denomination overseas 1989 (percent)	\$ value of notes in circulation 1989 (billions)	\$ value of notes overseas 1989 (billions)	Denomination composition of overseas currency (percent)
\$1	35.7	4.00	1.43	1.57
\$5	37.6	4.99	1.88	2.07
\$10	19.3 - 35.0	10.25	1.98 - 3.59	3.95
\$20	49.0 - 69.3	54.67	21.86 - 37.89	24.09
\$50	45.3	26.25	11.89	13.10
\$100	51.1	98.17	50.12	55.21
Total	45.8 - 53.0	198.33	89.2 - 106.8	100.0

age distribution characteristics suggests that in 1989, between 45.8 and 53 percent of the U.S. currency stock was held overseas. Of these overseas holdings, 68.3 percent was held in the form of large denomination bills (\$100 and \$50); approximately 28 percent was held in mid-sized denomination bills (\$20 and \$10) and 3.6 percent was held in the form of small denomination bills (\$5 and \$1).

The MDM employing seasonality, series, and coin/bill ratio characteristics

Porter and Judson (1995) employ several variants of the MDM to estimate the proportion of currency held overseas by exploiting

assumed differences in the seasonality, series, and coin/bill ratio characteristics of domestic and overseas holdings of U.S. currency.

Since the seasonal component characteristic of the total U.S. currency population (S) is directly measurable, but the seasonal characteristics of the domestic (S^d) and foreign (S^o) are unobservable, Porter and Judson assume that for the period 1947-1994, the seasonal component of domestic U.S. currency holdings is identical to the observed seasonal pattern of the Canadian currency supply.¹³ They furthermore assume that there is no significant seasonal component in the foreign demand for U.S. currency, so that the seasonal characteristic of overseas holding of U.S. currency (S^o) can be assumed to be equal to unity. The seasonal variant of the MDM(S) model can then be estimated from the equation:

$$(Eq. 5) \quad S = \beta^d S^d + (1 - \beta^d) S^o$$

where the seasonal characteristics are time dependent and:

$$S = S^{US}$$

$$S^d \approx S^{CAN}$$

$$S^o \approx 1.$$

From Equation (3), it follows that the domestic share of currency holdings (β^d) is estimated as:

$$(Eq. 6) \quad \beta^d = (S - 1) / (S^d - 1).$$

Table 2 presents Porter and Judson's reported estimates of the denomination-specific share of U.S. currency held overseas for 1989. The denomination-specific MDM(S) yields an overall estimated proportion of currency abroad of 62.4 percent for 1989 compared to the estimated range of 45.8 to 53.0 percent from the MDM(A) age characteristic model. Concerning the composition of the currency held abroad, the MDM(S) results suggest that 67.8 percent of foreign holdings are in the form of large denomination bills, and correspondingly 29.7 percent and 2.5 percent are held in the form of mid-sized and small denomination notes.

A second variant of the MDM estimated by Porter and Judson exploits differences between the series composition characteristic (SR) of domestic and overseas notes to estimate the proportion of \$100 and

\$50 circulating abroad. In 1991, the Federal Reserve introduced a 1990 series note which was distinguished from the pre-1990 series notes in circulation by a polyester strip and microprinting to deter counterfeiting. Let the series composition characteristic (SR) be the proportion of the circulating note population (N) made up of new 1990 series notes (N^{90}) so that $(SR) = N^{90} / N$.

Table 2. Estimates of the Demographic Model
{MDM[$C^{den i}$; $C^{den i d}$; $C^{den i o}$];(S)}
Annual Seasonal Characteristics

Denomination	Estimated share of each denomination overseas 1989 (percent)	\$ value of outstanding notes 1989 (billions)	\$ value of notes overseas 1989 (billions)	Denomination composition of overseas currency (percent)
\$1	10.0	4.00	0.40	0.32
\$5	54.2	4.99	2.70	2.18
\$10	44.2	10.25	4.53	3.66
\$20	59.1	54.67	32.31	26.09
\$50	50.7	26.25	13.31	10.75
\$100	72.0	98.17	70.60	57.00
Total	62.4	198.33	123.85	100.0

Since the series composition of the total currency population is known, Porter and Judson require estimates of (SR) for both the domestic and overseas components. They assume that the series composition of "overseas" notes is adequately proxied by an estimate of the series composition of notes processed by the New York Federal Reserve, and that an estimate of the series composition of the notes processed by all other Federal Reserve Banks adequately reflects the "domestic" composition.¹⁴ The $MDM(SR)$ for series composition can then be represented as:

$$(Eq. 8) \quad SR = \beta^d SR^d + (1 - \beta^d) SR^o$$

where:

$(SR) = N^{90} / N$ is known and, by assumption,

$$SR^d \approx SR^{NonNY}$$

$$SR^o \approx SR^{NY}$$

The proportion of notes held domestically can be estimated as:

$$(Eq. 9) \quad \beta^d = (SR - SR^o) / (SR^d - SR^o) \approx (SR - SR^{NY}) / (SR^{NonNY} - SR^{NY}).$$

Porter and Judson employ two different procedures for estimating the domestic and overseas series composition characteristics, and table 3 presents their upper and lower bound estimates for the \$50 and \$100 denominations.

Table 3. Estimates of the Demographic Model

$\{MDM[C^{den1}; C^{den1d}; C^{den1o}]; (SR)\}$

$\{MDM[C; C^d; C^o]; (C/N)\}$

Series and Coin or Coin/Note Ratio Characteristics

	Estimated share of each denomination overseas 1994 (percent)	Estimated share of each denomination overseas 1994 (percent)	Estimated share of each denomination overseas 1994 (percent)	Estimated share of overseas currency 1989 (percent)
Denomination	Lower bound series model	Upper bound series model	Average series model	Coin ratio
\$1-\$20	n.a.	n.a.	n.a.	n.a.
\$50	28.0	48.0	38.0	n.a.
\$100	55.6	70.7	63.2	n.a.
Total	n.a.	n.a.	n.a.	20.9

A third variant of the MDM employs the ratio of coins to notes as the characteristic distinguishing domestic and overseas holdings of currency. Since the coin/note ratio of the total U.S. currency population

is directly observable, it remains to identify the coin/note ratio of domestic and overseas holdings. Porter and Judson assume that the domestic coin ratio can be represented by Canada's coin/note ratio and that the overseas ratio is identically zero since virtually no U.S. coin is held overseas. Let C/N represent the population ratio of coins to notes, $(C/N)^d$ represent the domestic coin ratio, and $(C/N)^o$ the overseas coin ratio. If β^d represents the fraction of U.S. currency held domestically, then it follows from equation (3) that the $MDM(C/N)$ can be represented as:

$$(Eq. 10) \quad (C/N) \approx \beta^d (C/N)^d + (1 - \beta^d) (C/N)^o.$$

By assumption,

$$(C/N)^d \approx (C/N)^{CAN} \text{ and}$$

$$(C/N)^o \approx 0$$

therefore, (Equation 3) reduces to:

$$(Eq. 11) \quad \beta^d = (C/N) / (C/N)^{CAN}.$$

As displayed in table 3, Porter and Judson's $MDM(C/N)$ estimate of the share of U.S. currency held abroad in 1989 is 20.9 percent.¹⁵

In order to examine the robustness of the $MDM(S)$ results presented by Porter and Judson, the model was reestimated employing the X11 ARIMA method for calculating the multiplicative seasonal component of notes in circulation for the United States and Canada.¹⁶ Our reestimation of the $MDM(S)$ confirms the Porter and Judson findings that the model is incapable of producing sensible estimates at monthly or quarterly frequencies. In particular, monthly and quarterly estimates of the overseas share of U.S. currency reveal a strong seasonal component, suggesting that the assumption that $(S^o \approx 1)$ may be unsustainable. Even annual time series estimates of the overseas shares obtained from the annual average of monthly seasonal components is quite different from that derived by Porter and Judson, employing a seasonal amplitude metric of the difference between the December and February sea-

sonals. Figure 8 presents the Porter and Judson time series of the estimated share overseas [MDM(S):DEC-FEB] and the corresponding estimate based on the average of monthly seasonal components [MDM(S):Monthly Average]. The figure also includes the range of 1989 point estimates from the age characteristic model [MDM(A1)] and [MDM(A2)], the overseas shares derived from the coin ratio model [MDM(C/N)], and the average share of \$100 notes derived from the series characteristic model [MDM(SR)].

As displayed in figure 8, the monetary demography models produce a wide range of estimates of the overseas share of U.S. currency and different temporal patterns of the change in overseas holdings. Given the diversity of results, and the strong assumptions required to produce them, it is difficult to place much confidence in the foregoing findings.

Figure 8. Share of U.S. Currency Overseas
Monetary Demography Model Estimates

The age characteristic model required the elimination of sample outliers before convergence could be obtained. The coin ratio model produces negative overseas shares for the period 1972-82, and the seasonal characteristic estimates produce implausible results at monthly and quarterly frequencies. Both the seasonal and serial characteristic models require strong assumptions concerning the unob-

served domestic and overseas characteristic specifications. Given these difficulties, we now turn to several alternative approaches for estimating the share of U.S. currency held abroad.

Note Ratio Models

The Note Ratio Model (NRM) provides an alternative means for obtaining indirect estimates of the share of currency held abroad. The known amount of U.S. notes in circulation (N) can be decomposed into the unknown amount of notes in domestic circulation (N^d) and the unknown amount circulating overseas (N^o). Let Z denote any variable that is assumed to affect the demand for notes. Then,

$$(Eq. 12) \quad N/Z = N^d/Z + N^o/Z.$$

As with the MDM models, assume that the domestic U.S. ratio can be proxied by the same ratio in Canada so that:

$$N^d/Z \approx (N/Z)^{CAN}.$$

Substituting the Canadian ratio $[(N/Z)^{CAN}]$ into Equation (12), multiplying through by Z and dividing both sides by N yields a solution for the unknown fraction of notes overseas (β^o).

$$(Eq. 13) \quad (\beta^o) = N^o/N = [N - (N/Z)^{CAN} \cdot Z]/N.$$

The simple note ratio model (NRM) is estimated for several variants where Z alternatively represents:

- (i) Personal Consumption Expenditures (PCE)
- (ii) Personal Disposable Income (PDI)
- (iii) Population (POP) x Consumer Price Index (CPI).

Figure 9 presents the estimated share of U.S. currency held overseas from each of the variants of the note ratio model: NRM(PCE), NRM(PDI), and NRM(POP, CPI). The results suggest that the share of U.S. notes held overseas declined for almost a decade between the early 1960s and the early 1970s and then rose significantly over the

next two decades. The peak in overseas holdings appears to have occurred in 1990 when roughly 30 to 35 percent of U.S. notes in circulation are estimated to be held abroad. The time series of estimated shares of currency held abroad derived from the note ratio models yield markedly lower shares abroad than those estimated by the seasonal variant of the MDM and higher shares than those estimated by the MDM(C/N).

**Figure 9. Share of U.S. Currency Overseas
Note Ratio Models**

Table 4 presents the correlation matrix of annual estimates of the share of currency overseas obtained from each of the indirect methods. The correlation matrix reveals relatively high correlations between all of the NRM estimates and the MDM(C/N) estimate of overseas shares. Comparing the MDM(S) estimate obtained from annual averages of the estimated monthly seasonal components with the MDM(S):Dec.-Feb. estimate (employing the difference between December and February seasonal components) reveals that the two alternative methods of computing the seasonal estimates yield very different results. The cor-

Table 4. Correlation Matrix of Indirect Annual Estimates of Share of Currency Overseas: 1962-1994

	MDM(C/N)	MDM(S)	MDM(S):Dec.-Feb.	NRM(PCE)	NRM(PDI)	NRM(POP,CPI)
MDM(C/N)	1.000	0.079	0.601	0.879	0.744	0.835
MDM(S)	0.079	1.000	0.393	-0.136	-0.026	-0.301
MDM(S):Dec.-Feb.	0.601	0.393	1.000	0.678	0.864	0.231
NRM(PCE)	0.879	-0.136	0.678	1.000	0.916	0.828
NRM(PDI)	0.744	-0.026	0.864	0.916	1.000	0.552
NRM(POP,CPI)	0.835	-0.301	0.231	0.828	0.552	1.000

relation between the two seasonal model estimates is only .393, suggesting that the model is quite sensitive to the arbitrary choice of a metric. The MDM(S) displays low and negative correlations with the other estimates, whereas the smoothed MDM(S):Dec.-Feb. series displays positive correlations with the other estimates.

Indirect Estimates of Net Outflows of U.S. Currency

Given the wide range of estimates of the share of currency abroad produced by different models, we turn our attention to estimating the net outflows of currency implied by each of the MDM and NRM models. Given the known total stock of notes in circulation and indirect estimates of the share of currency abroad, we obtain year-end estimates of the total stock of currency estimated to be held abroad.¹⁷ The difference in these estimated year-end overseas stocks yields estimates of the annual net outflows of currency from the U.S.

Table 5 displays the correlation matrix of the estimated net outflows derived from each of the indirect methods, and figure 10 displays the estimated net outflows. The net outflow estimates from the different

**Figure 10. Annual Net Outflows of U.S. Currency
NRM and MDM Estimates**

Table 5. Correlation Matrix of Annual Estimates of Net Outflows of Currency Overseas: 1962-1994

	MDM(C/N)	MDM(S)	MDM(S):Dec.-Feb.	NRM(PCE)	NRM(PDI)	NRM(POP,CPI)
MDM(C/N)	1.000	0.884	0.859	0.770	0.660	0.872
MDM(S)	0.884	1.000	0.990	0.793	0.732	0.853
MDM(S):Dec.-Feb.	0.859	0.990	1.000	0.791	0.758	0.847
NRM(PCE)	0.770	0.793	0.791	1.000	0.726	0.763
NRM(PDI)	0.660	0.732	0.758	0.726	1.000	0.769
NRM(POP,CPI)	0.872	0.853	0.847	0.763	0.769	1.000

Table 6. Means and Standard Deviations of Indirect Annual Net Outflow Estimates (\$ Millions)

	MDM(C/N)	MDM(S)	MDM(S):Dec.-Feb.	NRM(PCE)	NRM(PDI)	NRM(POP,CPI)
Mean	3,704	7,041	7,122	3,148	2,834	3,297
Std. Dev.	7,938	6,627	6,593	5,502	3,653	5,279

models appear to be more highly correlated with one another than the estimated net outflows. The net outflow estimates from the different models appear to be more highly correlated with one another than the share estimates, suggesting that the indirect methods may produce more accurate estimates of outflows than shares abroad. The major difference between the estimated net outflows is their magnitudes. As displayed in table 6, the MDM seasonal estimates yield average annual outflows that exceed the highest estimate produced by the other models by more than \$3.3 billion per year. Given these disparities, we turn our attention to direct estimates of net outflows of U.S. currency.

Direct Estimates of Net Outflows of U.S. Currency

Large shipments of U.S. currency into and out of the United States are typically handled by a small number of commercial banks that specialize in the business of bulk currency transport. These large currency shipments have been informally reported to the Federal Reserve Bank of New York cash office since 1988. Although the period spanned by these confidential estimates is short, and the data are not comprehensive, being limited to major shippers operating in the New York Federal Reserve District, they provide useful information on a portion of bulk cash shipments to and from the United States.¹⁸ Interviews with Federal Reserve officials suggest that much of the currency employed for bulk overseas currency shipments by the major transporting banks is supplied by the New York Federal Reserve Bank in the form of \$100 denomination notes. Employing FR-160 data, Feige (1994a) observed a correlation of .979 between the net value of \$100 denomination notes (NYNET) injected into circulation by the New York Federal Reserve and the net amount of currency shipped overseas as recorded in the confidential data series on bulk shipments abroad informally collected by the New York Federal Reserve Bank. Feige (1994a) employed (NYNET) as a proxy for net outflows of bulk currency shipments for the period 1974-1988, and this proxy was subsequently used by Porter and Judson (1995) as a proxy measure of total currency flows overseas.

The second direct measure of currency inflows and outflows is collected as part of the regulatory responsibility of the U.S. Customs Ser-

vice. Enacted in October 1970, the Currency and Foreign Transactions Reporting Act (also known as the "Bank Secrecy Act") required persons or institutions importing or exporting currency or other monetary instruments in amounts exceeding \$5,000, to file a Report of International Transportation of Currency or Monetary Instruments. These reports, commonly known as CMIRs, have been collected by the U.S. Customs Service since 1977. In 1980, the required reporting limit was raised to \$10,000. Although the CMIR data system was established with the aim of recording individual instances of cross border inflows and outflows of currency and monetary instruments, its micro data components can be usefully aggregated to study the size, origin, and destination of cross border currency flows. Since its inception, the CMIR data system has collected 2.3 million inbound filings and more than 300,000 outbound filings. With the cooperation of the U.S. Customs Service and the U.S. Treasury Department's Financial Crimes Enforcement Network (FinCEN), the information contained in the millions of accumulated confidential individual CMIR forms was combined by a specially designed algorithm that aggregated currency inflows and outflows by mode of transport and by origin and destination.

The CMIR data system represents the most comprehensive source of direct information on currency flows. It differs from the informal Federal Reserve data system (proxied by NYNET) in several important respects. It contains all reported currency inflows and outflows whether physically carried by individuals or shipped by financial institutions. The only excluded transactions are those that fall below the reporting requirements, direct shipments by Federal Reserve Banks, and shipments that circumvent the legal reporting requirements. The CMIR data are therefore more inclusive than the Federal Reserve informal series, which is limited to bank shipments to and from the New York Federal Reserve District by large bulk shippers.

Figure 11 displays the net outflows of U.S. currency as derived from the two direct sources of information pertaining to net outflows. The figure contains the NYNET proxy for the Federal Reserve informal series as well as the total net outflows as reported in the CMIR reports (CTNET). Figure 11 also displays two key components of the CMIR source data, namely the series of transported net outflows reported by financial institutions (CSNET) and the net outflows physically carried

by individuals (CCNET). The total net outflow series derived from the CMIR reports differs markedly from both the NYNET series and the net outflow series derived from indirect methods. In particular, CTNET suggests a significantly lower volume of net outflows than any of the other estimates. The source of the difference is largely due to the individual component (CCNET), which suggests that individuals physically transported larger amounts of currency into the United States than out of the United States for all periods except 1990. Since individual transactions are not recorded in the informal Federal Reserve data, it is necessary to examine more closely the conceptual relationship between CMIR direct source information on net outflows and the information on net outflows obtained by other direct and indirect means.

**Figure 11. Direct Estimates of Net Outflows
NYNET and CMIR Annual Estimates**

*Conceptual and Empirical Comparisons
of Direct Measures of Currency Flows*

Table 7 presents a conceptual comparison of the coverage of the CMIR reporting system and the Federal Reserve reporting system. The

table reveals that there are major differences in the conceptual content and coverage of the CMIR and FED currency flow data systems. In order to derive meaningful comparisons between the information content of the two reporting systems, it is first necessary to obtain measures of currency inflows and outflows that are conceptually comparable between the two data systems.¹⁹ An algorithm was therefore written to aggregate the micro CMIR data records into currency flows that would be comparable to FED data in both content and coverage.

Table 7. Content Comparison of CMIR and FED Currency Flow Reporting Systems

Characteristics	CMIR	Federal Reserve
Level of aggregation	Individual transactions	Aggregate transactions of large bulk shippers
Number of records	2.3 million inflow records 300,000 outflow records	Approximately 62,000 records (7 years x 12 months x 8 banks x 93 countries)
Time periods	Monthly—1977-1994	Monthly—1988-1994
Private institutions	All reporting banks	Major New York banks
Federal Reserve banks	Not included	New York Federal Reserve
Individuals	Reported cross border currency transport $\geq$ \$5,000 (pre-1980); $\geq$ \$10,000 (1980-1994)	Not reported
Domestic coverage	Entire U.S. (all Federal Reserve Districts)	New York Federal Reserve Branch
Overseas coverage ^a	220 countries	93 countries

a. The countries reported in the two data systems do not match exactly. Differences reflect temporal name changes and different levels of country aggregation (i.e., United Kingdom vs. England and Scotland). A separate comparison algorithm was written which resolves these difficulties and creates two country sets: those included in both the Federal Reserve and CMIR data systems and those included in the CMIR system but not in the FED system.

Table 8 (A and B) presents an aggregation scheme for the CMIR data system that creates observations on particular currency inflows

Table 8(A). CMIR-Currency Inflow Matrix: Notation

Mode of transport	Shipped by bank			Carried by individual			Total CMIR
Destination district	Country of origin						
	FED countries	All other countries	Total shipped	FED countries	All other countries	Total carried	
NY district	CSINYF	CSINYO	CSINY	CCINYF	CCINYO	CCINY	CTINY
All other districts	CSIOF	CSIOO	CSIO	CCIOF	CCIOO	CCIO	CTIO
Total CMIR	CSIF	CSIO	CSI	CCIF	CCIO	CCI	CTI

Table 8(B). CMIR-Currency Outflow Matrix: Notation

Mode of transport	Shipped by bank			Carried by individual			Total CMIR		
District of origin	Destination country								
	FED countries	All other countries	Total shipped	FED countries	All other countries	Total carried			
	NY district	CSONYF	CSONYO	CSONY	CCONYF	CCONYO		CCONY	CTONY
	All other districts	CSoOF	CSOOO	CSO	CCOOF	CCOOO		CCOO	CTOO
Total CMIR	CSoF	CSOO	CSO	CCOF	CCOO	CCO	CTO		

and outflows that are conceptually comparable to the currency inflows and outflows believed to be captured in the Federal Reserve data system.²⁰ Individual CMIR currency inflow and outflow transaction reports are aggregated by mode of cross border transport and by the geographic origin and destination of the currency flow. Thus, for currency inflows, total CMIR inflows (*CTI*) were first disaggregated by mode of transport into transactions involving shipment by a financial institution (*CSI*) and transactions involving currency physically carried by individuals (*CCI*). Shipped inflows (*CSI*) were then further disaggregated by country of origin and Federal Reserve District of destination. It was then possible to recombine the shipped inflows into four exhaustive categories reflecting:

1. All shipments whose country of origin was contained in the FED data system and whose destination was the New York Federal Reserve District (*CSINYF*);
2. All shipments whose country of origin was not in the FED data system but whose destination was the New York Federal Reserve District (*CSINYO*);
3. All shipments whose country of origin was included in the FED data system but whose destination district was not New York (*CSIOF*); and
4. All shipments whose country or origin was not in the FED data system and whose district of destination was not New York (*CSIOO*).

Table 9 (A and B) presents the same inflow and outflow matrix for the FED data system, indicating that the only conceptually comparable currency flows captured by both data systems is *CSINYF=FSI* and *CSONYF=FSO*. None of the other flows captured by the CMIR data is available in the FED data. The Federal Reserve data system excludes currency shipments whose origin or destination are the eleven non-New York Federal Reserve Districts as well as shipments whose origin and destination are countries other than those appearing in the FED data system. The FED data also excludes all inflows and outflows of currency physically transported by individuals.

In order to gauge the empirical importance of the conceptual information content excluded from the FED data system, table 10 (A and B)

Table 9(A). FED-Currency Inflow Matrix: Notation

Mode of transport	Shipped by bank			Carried by individual			Total FED
Destination district	Country of origin						
	FED countries	All other countries	Total shipped	FED countries	All other countries	Total carried	
NY district	FSI	NA	NA	NA	NA	NA	NA
All other districts	NA	NA	NA	NA	NA	NA	NA
Total FED	NA	NA	NA	NA	NA	NA	NA

Table 9(B). FED-Currency Outflow Matrix: Notation

Mode of transport	Shipped by bank			Carried by individual			Total FED
District of origin	Destination country						
	FED countries	All other countries	Total shipped	FED countries	All other countries	Total carried	
NY district	FSO	NA	NA	NA	NA	NA	NA
All other districts	NA	NA	NA	NA	NA	NA	NA
Total FED	NA	NA	NA	NA	NA	NA	NA

Table 10(A). CMIR-Currency Inflow Matrix: Percent of Total Currency Inflows (CTI) by Source

Mode of transport	Shipped by bank			Carried by individual			Total CMIR
Destination district	Country of origin						
	FED countries	All other countries	Total shipped	FED countries	All other countries	Total carried	
NY district	50.9	0.5	51.5	3.8	0.5	4.2	55.7
All other districts	15.5	1.2	16.7	23.7	3.9	27.6	44.3
Total CMIR	66.4	1.7	68.2	27.4	4.4	31.8	100.0

Table 10(B). CMIR-Currency Outflow Matrix: Percent of Total Currency Outflows (CTO) by Source

Mode of transport	Shipped by bank			Carried by individual			Total CMIR
District of origin	Destination country						
	FED countries	All other countries	Total shipped	FED countries	All other countries	Total carried	
NY district	82.9	0.9	83.8	1.7	0.1	1.8	85.6
All other districts	4.4	1.8	6.2	5.2	2.9	8.1	14.4
Total CMIR	87.3	2.7	90.0	6.9	3.1	10.0	100.0

presents a percentage breakdown of the various disaggregated inflows and outflows contained in the CMIR data for the period 1988-1994 which corresponds to the time interval covered by the Federal Reserve statistics.

A major finding is immediately apparent from table 10 (A and B) concerning the relationship between Federal Reserve data as proxied by NYNET and CMIR data. Only 50.9 percent of all recorded CMIR inflows are likely to be captured by the FED data system, whereas 82.9 percent of all CMIR outflows are likely to be recorded by the FED data system. Since the CMIR data system includes many inflows of currency which are not represented in the FED data source, the FED net outflow series will overstate the net outflows of currency abroad and therefore give the misleading impression that a larger fraction of U.S. currency is held abroad.

Table 11 presents the means and standard deviations of various direct measures of average quarterly gross currency inflows and outflows. Comparing the means of the Federal Reserve and CMIR inflow and outflow estimates reveals that the CMIR data are considerably more inclusive than the Federal Reserve estimates.²¹ Recorded average quarterly CMIR total currency inflows (CTI) exceed Federal Reserve (FSI) recorded bulk receipts of currency by some \$2.86 billion. Similarly, recorded average CMIR total currency (CTO) outflows exceed Federal Reserve (FSO) bulk currency shipments by \$1.39 billion per quarter. However, when conceptually comparable magnitudes are compared we find that CMIR estimates of bulk shipments to the New York Federal Reserve District from countries included in the FED data system (CSINYF) are, as expected, very close to the FED data (FSI). This is true to an even greater extent for comparable outflows.

Comparing conceptually comparable magnitudes reduces the inflow discrepancy to less than \$.2 billion (2,765-2,567) and the outflow discrepancy (6,015-5,869) to \$.15 billion. The average quarterly inflows of all commercially shipped currency (CSI) recorded by the CMIR data still exceed the Federal Reserve's estimated inflows (FSI) by roughly \$1.13 billion while the average CMIR shipped outflows (CSO) exceed the Federal Reserve's estimates of outflows (FSO) by \$.67 billion per quarter. These discrepancies are due to the fact that 16.7 percent of total CMIR currency shipped inflows arrive at banks outside the New York district, and 6.2 percent of CMIR outflows originate at banks outside of New York. The Federal Reserve's data appear to *understate* gross inflows to a greater degree than they understate gross outflows,

Table 11(A). Descriptive Statistics of Direct Measures of Quarterly Gross Currency Inflows, 1988-1994

\$ (Mil.)	CCI	CSI	CSINYF	CTI	FSI	NYIN
Mean	1,729	3,700	2,765	5,429	2,567	2,028
Std. Dev.	1,074	2,062	1,756	2,219	1,742	1,271

Table 11(B). Descriptive Statistics of Direct Measures of Quarterly Gross Currency Outflows, 1988-1994

\$ (Mil.)	CCO	CSO	CSONYF	CTO	FSO	NYOUT
Mean	724	6,535	6,015	7,259	5,869	6,265
Std. Dev.	583	2,473	2,532	2,539	3,250	3,045

leading to an *overstatement* of net shipped currency outflows. This is true to an even greater extent for the shipment proxy (NYNET). As displayed in table 10, individuals physically transport 31.9 percent of total gross inflows but only 10 percent of total gross outflows.²² Since the FED data do not include currency physically transported by individuals, the overstatement of net outflows is further accentuated. Any conclusions derived from Federal Reserve data or from series closely correlated with FED data (such as NYNET), are therefore likely to overstate net outflows and therefore lead to the erroneous conclusion that foreign holdings of U.S. currency are increasing at a faster rate than is in fact the case.

This conclusion is subject to one caveat. It is likely that the CMIR filing compliance rate is higher for currency physically carried into the United States than for currency physically carried out of the United States, since U.S. Customs forms are routinely collected from incoming travelers. During the period 1988-94, there were roughly nine inflow filings for every one outflow filing, and the average size of reported outflows was almost four times larger than the average size of the corresponding reported inflow. The underreporting of physically transported outflows and their large average size would induce a downward bias to the carried currency component of CMIR net outflows.

In summary, the CMIR data represent the most comprehensive direct measure of currency inflows and outflows. We have shown that the one component of the CMIR data that can be directly compared with other direct data sources is very reliable. As displayed in table 12, the CMIR measure of net shipped New York outflows (CSNYFNET) is almost identical to the conceptually comparable figure (FSNET) yielded by the FED data system. The total CMIR quarterly net outflows (CTNET = CSNET + CCNET) are approximately \$.47 billion lower than the average of quarterly net outflows estimated from the note ratio models (PCE; PDI; POP,CPI) whereas the NYNET and SEANET estimates are approximately \$2.4 billion higher. The CMIR data's lower estimate of net outflows appears to be entirely due to the component (CCNET) representing physical transport of currency by individuals. This component is negatively correlated or uncorrelated with other net outflow measures. The individual component of the CMIR either reflects a unique information signal not contained in the other series or it may be the result of a compliance bias that differen-

Table 12. Direct and Indirect Measures of Quarterly Net Currency Outflows, 1988-1994

\$ (Mil.)	CCNET	CSNET	CTNET	CSNYFNET	PCE	PDI	POPCPI	NYNET	SEANET	FSNET
Mean	-1,005	2,835	1,830	3,250	2,391	1,839	2,661	4,237	4,201	3,303

tially effects the incoming and outgoing reporting of individual transport of currency.

Unrecorded Travel and Remittance Flows

The only inflows and outflows not captured by the CMIR direct measures are those flows that fall below the reporting requirement and those flows that should legally be reported but escape detection due to noncompliance. If these flows are relatively large, and outflows substantially exceeded inflows, they could account for some of the observed disparities between the CMIR measure and other estimates of net outflows.

Three flows require further investigation. Unrecorded inflows include U.S. currency carried into the country by foreign travelers visiting the United States below the filing threshold; unrecorded outflows include U.S. currency taken abroad by U.S. travelers and net remittances of currency sent abroad. To estimate the size of these flows we examined annual data on total expenditures net of airfares made in the United States by foreign travelers to this country and total expenditures made by U.S. travelers abroad for the period 1962-1994.²³ To simulate the net outflows of currency deriving from travelers we assumed that prior to the widespread use of credit cards, foreign travelers to the U.S. made 20 percent of their purchases of goods and services with U.S. currency brought into the country from abroad, whereas U.S. travelers carried U.S. currency out of the country to cover 10 percent of their expected overseas expenditures. These proportions were reduced over time to allow for the growing use of credit cards. Net remittances are estimated by the Bureau of Economic Analysis, whose economists suggested that, at most, 10 percent of remittances involve transfers of U.S. currency. Our simulations over several sets of plausible assumptions revealed that annual net outflows of U.S. currency from travel and remittance sources rarely exceeded \$.5 billion and in most recent years, net outflows from these sources were actually negative. We conclude that the net currency flows arising from travel and remittance transactions that may be excluded from CMIR reports are too small to account for the disparities between CMIR net outflows and those estimated by indirect means. If anything, travel flow estimates tend to accentuate the observed discrepancies.

Composite Estimates

Given the diversity of indicators of the unknown net flows of currency overseas, it is desirable to combine these alternative measures in order to obtain a single composite estimate of net currency outflows based on all available information. One approach is to use a factor analysis model to estimate the common signal or latent variable (L_t) associated with several alternative indicators of net overseas currency flows (M_{it}). In the factor model,

$$M_{it} = \beta_i L_t + \epsilon_{it},$$

each of the M_i indicators of net outflows is linearly related to the latent common factor (L_t). The β_i 's represent the factor loadings and the ϵ_{it} are the temporal measurement errors in each of the i measures of net currency outflows.²⁴ Since different estimates of net outflows are available for different time periods and different frequencies, we estimated several factor models for both annual and quarterly frequencies in order to examine the stability of the results. The variables employed and the periods covered by the various estimates are as follows:

Annual Factor Model 1 (AF1) - Period 1962-1994

Six variables: Net outflows: MDM(S); NRM(PDI); NRM(PCE); NRM(POP,CPI); MDM(C/N); TRAV.²⁵

Quarterly Factor Model 1 (QF1) - Period 1961:1 - 1994:4

Four variables: Net outflows: NRM(PDI); NRM(PCE); NRM(POP,CPI); MDM(C/N).

Quarterly Factor Model 2 (QF2) - Period 1977:1 - 1994:4

Seven variables: Net Outflows: CCNET; CSNET; NYNET; NRM(PDI); NRM(PCE); NRM(POP,CPI); MDM(C/N).

Figure 12 displays the maximum likelihood estimates of annualized net outflows derived from each of the foregoing factor models. The temporal pattern of all the estimates is broadly similar, suggesting a rising level of net outflows during the decade of the 1980s and a significant upward shift in net outflows during the early 1990s, largely associated with the increased use of U.S. currency as a co-circulating

medium of exchange in Eastern Europe and the newly independent republics of the former Soviet Union. The annual model (AF1), incorporating MDM, NRM, and TRAV results produces the highest estimated net outflows during the most recent years, whereas the seven-variable quarterly model (QF2) employing direct (CMIR) and proxy estimates (NYNET) as well as MDM and NRM estimates produces the lower bound net outflow estimates. Given these alternative factor model net outflow estimates, it remains to determine what current share of overseas holdings is consistent with the net outflow evidence. To examine the plausibility of alternative assumptions concerning the current share of U.S. notes held abroad, we conducted a series of simulation experiments to determine the implied share of currency held domestically, given alternative assumptions concerning the present share of U.S. notes abroad.

**Figure 12. Estimated Factor Model Net Outflows
1962-1994 Annual Net Outflows**

Figures 13 and 14, respectively, display the implied shares of currency held domestically over time assuming that the current share of notes held overseas is 40 percent and 60 percent. For the 40 percent Figures 13 and 14 respectively, display the implied shares of currency held domestically over time assuming that the current share of notes held overseas is 40 percent and 60 percent. For the 40 percent abroad simulation, all of the factor model estimates imply a feasible range of domestic holdings between 0 and 100 percent, whereas the 60 percent abroad simulations produce infeasible estimates which suggest that prior to the mid-1970s no U.S. currency was held domestically.²⁶

Figure 13. Domestic Share of U.S. Currency
Simulation: 1994=40% of Notes Overseas

On the basis of the foregoing simulations, we conclude that the most plausible estimate of the share of U.S. notes presently held abroad is roughly 40 percent, which implies that roughly 36 percent of U.S. currency is held abroad.²⁷ Employing this current value, figure 15 displays the implied time series of the share of currency held overseas between 1973 and 1994 for each of the factor model net outflow estimates. Given the considerable range of overseas shares implied by the three factor model estimates, we conclude that precise estimates of the domestic money supply remain elusive.

Figure 14. Domestic Share of U.S. Currency
Simulation: 1994=60% of Notes Overseas

Figure 15. Foreign Holdings of U.S. Currency
Percent of Currency Held Abroad

Implications of Foreign Holdings of U.S. Currency for Estimates of the Domestic Unreported Economy

The provisional estimates of overseas holdings suggest that earlier currency ratio model estimates of the unreported economy were mis-specified insofar as they erroneously assumed that the entire stock of U.S. currency was held domestically. It is now possible to reestimate the currency ratio models employing the new alternative estimates of the domestic stock of U.S. currency.

Figures 16 and 17 display estimates of total unreported income obtained from both the simple C/D model and the GCR model employing alternative factor model estimates of the domestic U.S. currency stock. Figure 18 displays the GCR estimates of unreported income as a percentage of AGI.

Total unreported income appears to have grown secularly until 1985, declined briefly around the time of the 1986 tax reform, and then peaked in 1991. The temporal pattern of the alternative GCR estimates of unreported income as a percent of AGI tells essentially the same story. Unreported income appears to have grown rapidly from 1966 and peaked as a percent of AGI in 1980. The percent of unreported income then declined until 1987, rose again until 1991, and fell again to a lower level approximating levels last observed in the early 1970s.

Of the three factor model estimates of the domestic currency supply, the QF2 estimate may be the most reliable, being based on quarterly frequencies and the largest amount of direct and indirect information concerning net currency outflows. The QF2 estimates from both the simple C/D and GCR models suggest that total unreported income in 1993 was roughly \$700 billion, representing approximately 20 percent of AGI.

The main conclusion to be drawn from these revised estimates of unreported activity is that once account is taken of foreign holdings of U.S. currency, the range of uncertainty concerning the magnitude of unreported income is substantially reduced. The difference between the unadjusted GCR estimates of unreported income and the IRS estimates for 1992 (figure 7) amounted to more than \$400 billion. The revised estimates displayed in figure 17 reveal that the difference between the IRS estimate and the QF2 estimate is now reduced to roughly \$100 billion.

**Figure 16. Total Unreported Income
Adjusted C/D Models: 1973=IRS Base**

**Figure 17. Total Unreported Income
Adjusted GCR Models: 1973=IRS Base**

Figure 18 reveals that unreported income as a percent of AGI varies considerably over time. Whereas earlier studies of the underground economy reported its secular growth in both absolute and relative terms, the most recent data suggest that since 1980, unreported income as a percent of AGI has fluctuated quite dramatically. The two most plausible explanations of these fluctuations in unreported income are changes in average tax rates and changes in the level of dissatisfaction with government.

Figure 19 displays the relationship between the QF2 revised estimates of unreported income as a percent of AGI to the average effective federal tax rate, and figure 20 displays the relationship between unreported income and an index of dissatisfaction with government.²⁸

Figure 19 reveals that tax evasion does appear to rise in response to higher average taxes, and conversely falls when the incentives to cheat are reduced by lower average tax rates. Similarly, in figure 20 we find the expected relationship between tax evasion activity and the level of dissatisfaction with government. The dramatic fall in the level of dissatisfaction with the government between 1980 and 1984, coincides with a fall in the relative level of tax evasion. Conversely, the increases in the level of dissatisfaction with government observed between the mid-1980s and 1990 are associated with a relative increase in tax evasion activities. It seems that when taxpayers perceive their public representatives to be dishonest, and when they perceive a decline in the public benefits obtained from their tax dollars, they are more likely to engage in tax evasion.

The finding that a substantial portion of U.S. currency is held overseas provides a partial resolution of the currency enigma. It will be recalled that Federal Reserve surveys suggest that U.S. households admit to holding roughly 12 percent of the nation's currency, firms account for roughly 3 percent, the unreported economy employs roughly 4 percent, and foreign holdings amount to roughly 36 percent. These estimates suggest that we have accounted for 55 percent of the nation's currency supply, leaving 45 percent still to be explained. Porter and Judson (1995), who place considerable emphasis on their version of the MDM(S) and MDM(SR) results, suggest that a much larger fraction of U.S. currency is held abroad—between 50 percent and 70 percent. Our analysis suggests that these estimates of overseas hold-

**Figure 18. Unreported Income as a Percent of AGI
Adjusted GCR Models: 1973=IRS Base**

**Figure 19. Unreported Income and Tax Rate
1973-1994**

ings are too high. We are more inclined to believe that survey respondents may systematically underreport the amounts of currency they hold and that surveys of currency usage may understate actual domestic holdings because of self-selection biases. Whether these domestic cash hoards are derived from underground activities that we continue to underestimate, or from legitimate activities that are simply underrecorded in our NIPA accounts, remains to be resolved.

The finding that roughly 36 percent of the stock of U.S. currency is currently overseas raises another monetary puzzle. Are foreign holdings of U.S. currency being used solely as a store of value or do they function as a co-circulating medium of exchange? Preliminary evidence (Feige 1994) based on an investigation of the age and quality of

**Figure 20. Unreported Income and Dissatisfaction
1973-1994**

some 200,000 individual notes suggests that the age/quality distributions of domestically circulating notes and notes returning to the United States from abroad are quite similar. These findings suggest that the velocity of domestically held currency is on average not that differ-

ent from the velocity of currency held abroad. If foreigners' holdings of U.S. currency circulate at the same rate as that of U.S. household holdings, they would generate a flow of annual cash payments approaching the size of the GDP of the United States. Thus, the partial resolution of the currency enigma for the United States merely creates another monetary anomaly for the rest of the world. The world economy appears to subsume a U.S.-sized unrecorded economy that employs U.S. currency as its medium of exchange. The world's currency enigma deepens when one considers that our revised estimates of U.S. per capita currency holdings are still modest compared with the per capita currency holdings of other developed European and Asian nations. The problem of missing currency is not limited to the U.S. dollar but also extends to other major currencies, most importantly to the German mark and the Japanese yen.

NOTES

1. Currency in circulation refers to the amount of currency held outside of the Treasury and Federal Reserve. Except for small amounts of currency that may have been inadvertently lost or destroyed by the public (Laurent 1974), currency in circulation includes the holdings of financial intermediaries and the public. Reliable data on financial intermediary holdings of vault cash are readily available, and it is therefore possible to obtain accurate estimates of the total stock of currency outside the banking system. For a complete description of the cash payments system, see Feige (1994).

2. The taxonomy of "underground economies," which identifies the interrelationships between the illegal, unreported, unrecorded, and informal economies, is discussed in Feige (1989, 1990a).

3. On page A-101 of the IRS (1988) report, the IRS acknowledges the major limitations of its projections of unreported income: "Because we essentially hold constant rates of noncompliance through 1992, these estimates do not reflect recent trends in noncompliance. Second, we assume that tax reform has no impact on individual's behavior in terms of either their propensity for noncompliance or the types of incomes individuals will receive in future years. Third, these projections are sensitive to changes in macroeconomic model projections of incomes in future years."

4. The IRS estimates reported above are based on the recommendations of the tax examiners. Since some of these recommendations are challenged by the taxpayer, the IRS also prepared an alternative set of estimates on an assessed basis. These are reported in Appendix E (IRS 1988).

5. As described in Feige (1989), the foregoing restrictions imply that the ratio of unreported (Y_u) to reported (Y_o) income can be estimated as follows:

$$Y_u/Y_o = (C - koD)/(ko+1)D$$

where:

C = Currency

D = Checkable Deposits

$ko = C_o/D_o$.

6. The general currency ratio model (GCR) permits a relaxation of several of the assumptions employed in the simple currency ratio model. In particular, currency needs no longer to be the exclusive medium of exchange in unreported transactions, and any year for which an independent estimate of unreported income is available can serve as a benchmark.

The GCR model can be solved to obtain the equation for the ratio of unreported income which is:

$$Y_u/Y_o = (k_u + 1)(C - k_o D) / (k_o + 1)(k_u D - C)$$

where:

k_u and k_o respectively represent the currency-deposit ratios in the unreported and in the reported economies.

7. The IRS estimate is the sum of the legal source unreported income estimate displayed in figure 3 plus a 15 percent imputation for illegal source income. The imputation for illegal source income is based on the illegal source income estimates reported in the earlier (IRS 1983) study. The currency ratio models yield estimates of total unreported income from all sources.

8. The methodology for estimating the velocity (turnover) of currency is described in Feige (1990b). The estimates are based on the Federal Reserve's Survey of Currency Usage, which finds that the income velocity of household cash holdings is roughly fifty turnovers per year. Share weighted denomination-specific velocities are obtained by estimating the average lifetime of each note denomination derived from FR-160 data on currency issues (births) and redemptions (deaths).

9. The research reported here is one of several approaches undertaken as part of a broader study for the Board of Governors of the Federal Reserve System and the U.S. Treasury Department Financial Crimes Enforcement Network (FinCEN) on estimating foreign holdings of U.S. currency.

10. The application of demographic theory and methods to currency populations is developed in Feige (1990b), which includes estimates of age-specific currency mortality and survival rates. Feige (1994) presents a full demographic model describing the life cycle of the individual note and the dynamics of note populations and cohorts.

11. See the 1990 Federal Reserve Soil Level Impact Task Force (SLITF) study entitled "Comprehensive Assessment of U.S. Currency Quality, Age & Cost Relationships."

12. The estimates presented for denominations \$1, \$5, \$50, \$100 are averages obtained from Baselines 1 and 2 of the SLITF study. The Baseline 1 model for the \$10 denomination and the Baseline 2 estimates for the \$20 denomination failed to converge even after significant outliers were deleted from the samples. We therefore report a range of estimates for these two denominations. The similarity of the age distributions of overseas and domestic notes suggests that the reported results are likely to contain a wide margin of error.

13. The assumption is justified by the argument that the U.S. and Canada have identical currency denomination structures and that the Canadian dollar is rarely used overseas.

14. Porter and Judson claim that almost all currency sent to and received from abroad is processed by the New York Federal Reserve Bank. The veracity of this assumption can be tested by an examination of CMIR data disaggregated by Federal Reserve District of origin and destination. The CMIR data reveal that only 52 percent of all reported currency inflows for the period 1977 and 1994 had the New York Federal Reserve District as their point of destination. The New York District was reported as the point of origin for 85 percent of total outflows during the period.

15. The reported results include an adjustment of the coin/note ratio to take account of the introduction of a one dollar coin in Canada in July 1987. The Bank of Canada continued to issue one dollar notes until June 30, 1989, at which time there were 246 million of the one dollar coins in circulation. By the end of 1989, the number of one dollar coins in circulation had risen to 464

million. The reported results are based on a time series forecast of what the coin/note ratio would have been in the absence of the introduction of the one dollar coin.

16. Porter and Judson obtained the seasonal components employing the STL seasonal adjustment procedure applied to the currency component (coin plus notes) of the Canadian and U.S. M1 series. In our replication, we employed the X11 ARIMA procedure on the Canadian and U.S. notes in circulation series, since neither Canadian nor U.S. coins are assumed to circulate overseas. The results reported by Porter and Judson are not based on the ratio of seasonal components as specified by Equation (6), but rather on the ratio of the seasonal amplitudes of the U.S. and Canadian series, derived by taking the difference between the December and February seasonals (Porter and Judson 1995, pp. 16-17). Our replication efforts suggest that the results are relatively insensitive to the use of different seasonal adjustment procedures and the substitution of the note series for the currency component series. However, the time series estimates of the share abroad are quite sensitive to the use of the seasonal amplitude metric employed by Porter and Judson. In particular, when the MDM(S) model is estimated on a monthly or quarterly basis, and the estimated monthly or quarterly overseas share is estimated as the ratio of each of the seasonal components minus one as suggested by Equation (6), the estimated monthly and quarterly shares abroad fluctuate wildly within a year, often yielding estimates of the share abroad that exceed 100 percent.

17. Throughout the analysis we assume that all U.S. coin is held domestically.

18. During the inter-war period between 1923 and 1941, the Federal Reserve published data on net currency shipments to European countries (Banking and Monetary Statistics: 1914-1941, Board of Governors of the Federal Reserve System, 1943, pp. 417-418). Over the entire period for which data are available, cumulative net inflows from Europe amounted to 4.8 percent of the average outstanding stock of currency during the period. The average annual net inflow of currency from Europe amounted to .25 percent of the average outstanding currency stock.

19. In order to maintain strict confidentiality of individual records in the CMIR data system, aggregations were performed at the offices of the U.S. Treasury (FINCEN). Subsequent analysis was performed on the aggregated data. Since Federal Reserve Banks are not required by law to file CMIR statements, the CMIR shipment series were augmented to include direct overseas currency shipments to and by the Federal Reserve Bank of New York.

20. The remaining minor incompatibility derives from the fact that CMIR records were created by Federal Reserve District whereas the FED data includes only observations for the New York Federal Reserve Branch of the New York Federal Reserve District. CMIR observations for the New York District therefore include the negligible transactions of the Buffalo Branch.

21. Under current CMIR reporting requirements, direct Federal Reserve Bank overseas currency shipments are not reportable on CMIRs but are included in the FED shipment series. In order to make both series comparable, direct Federal Reserve Bank shipments were added to the CMIR shipment series.

22. The series on individual inflows and outflows appear to be very different from the bulk shipments undertaken by financial institutions. Two explanations are possible for this important disparity. The data on physical transport by individuals include travel transport companies such as airlines or cruise ships that may generate U.S. currency outside the United States and regularly transport it back for deposit in their domestic bank. The discrepancy may also be due to differential rates of compliance with CMIR reporting. Individuals transporting currency out of the country are not monitored as carefully by the U.S. Customs as individuals returning to the United States. It is therefore possible that there is a lower rate of reporting compliance with physically transported outflows than with physically transported inflows.

23. The data were generously provided by the United States Travel and Tourism Administration, Washington, D.C.

24. A general discussion of factor analysis models can be found in Mulaik (1972). Bollen (1989) contains a review of factor analysis in the context of latent variables.

25. TRAV is an annual estimate of the net outflow of U.S. currency resulting from travel to and from the U.S. and net cash remittances sent abroad.

26. It should be noted that the individual net outflow results are even more sensitive to initial starting overseas values. In particular, the MDM(S) net outflow results only produce feasible estimates if we assume that between 65 and 75 percent of currency is presently overseas, whereas at the other extreme, the CMIR net outflows only produce feasible estimates in the 15 to 25 percent range.

27. Throughout the analysis, we assume that all U.S. coin is held domestically.

28. The average effective federal tax rate is simply the sum of federal government tax receipts divided by AGI. The dissatisfaction with government index is constructed as an equally weighted average of three normalized indices representing answers to the University of Michigan's Institute for Social Research (ISR) surveys concerning whether government officials can be trusted, whether they are crooked, and whether the government wastes taxpayers' money. I am indebted to the ISR for providing the underlying data.

References

- Abt Associates. 1984. "Unreported Taxable Income From Selected Illegal Activities." Government Printing Office, September.
- Anderson, P. 1977. "Currency in Use and in Hoards," *New England Economic Review* (March/April): 21-30.
- Avery, R.B., G.E. Elliehausen, A.B. Kennickell, and P.A. Spindt. 1986. "The Use of Cash and Transaction Accounts by American Families," *Federal Reserve Bulletin* 72 (February): 87-108.
- _____. 1987. "Changes in the Use of Transaction Accounts and Cash from 1984 to 1986," *Federal Reserve Bulletin* 73 (March): 179-196.
- Cagan, P. 1958. "The Demand for Currency Relative to Total Money Supply," *Journal of Political Economy* 66 (August): 303-328.
- Carson, C. 1984. "The Underground Economy: An Introduction," *Survey of Current Business* 64, 5 (May and July).
- Federal Reserve System. 1990. *A Comprehensive Assessment of the U.S. Currency System: Quality, Age and Cost Relationships*. Federal Reserve System.
- Feige, E.L., ed. 1989. *The Underground Economies: Tax Evasion and Information Distortion*. New York: Cambridge University Press.
- _____. 1990a. "Defining and Estimating Underground and Informal Economies: The New Institutional Economics Approach," *World Development* 18, 7 (July): 989-1002.
- _____. 1990b. "Currency Velocity and Cash Payments in the U.S. Economy: The Currency Enigma." Mimeo, Department of Economics, University of Wisconsin.
- _____. 1994a. "The Underground Economy and the Currency Enigma," *Public Finance and Irregular Activity*, supplement to *Public Finance/Finances Publiques*, 49: pp. 119-136.
- _____. 1994. "The Cash Payment System: Monetary Demography." Report prepared for the Board of Governors of the Federal Reserve System and the Financial Crimes Enforcement Network.
- Gutmann, P. 1977. "The Subterranean Economy," *Financial Analysts* 33 (November/December): 24-27.
- Internal Revenue Service (IRS). 1979. "Estimates of Income Unreported on Individual Tax Returns." Department of the Treasury Publication No. 1104 (9-79).
- _____. 1981. "Study of the Information Returns Program." Department of the Treasury, Research and Operations Analysis Division, July.

- _____. 1983. "Income Tax Compliance Research." Department of the Treasury, July.
- _____. 1988. "Income Tax Compliance Research." Supporting Appendices to Publication 7285, Department of the Treasury, Research Division Publication 1414 (7-88).
- _____. 1990. "Income Tax Compliance Research: Net Tax Gap and Remittance Gap Estimates." Department of the Treasury, Research Division Publication 1415 (4-90).
- Laurent, Robert. 1974. "Currency in Circulation and the Real Value of Notes," *Journal of Money, Credit and Banking* 6, 2 (May): 213-226.
- Muliak, S.A. 1972. *The Foundations of Factor Analysis*. New York: McGraw-Hill.
- Office of National Drug Control Policy. 1991. "What American Users Spend on Illegal Drugs." Technical Paper, June.
- Porter, R. 1993. Appendix to "Nominal Income Targeting with the Monetary Base as Instrument: An Evaluation of McCallum's Rule." Federal Reserve Board Finance and Economics Discussion Series, Working Study No. 1.
- Porter, Richard D., and Ruth A. Judson. 1995. "The Location of U.S. Currency: How Much Is Abroad?" Reproduced by Board of Governors, Federal Reserve System.
- Sprenkle, C. 1993. "The Case of the Missing Currency," *Journal of Economic Perspectives* 7, 4 (fall): 175-84.
- Sumner, Scott B. 1990. "The Transactions and Hoarding Demand for Currency," *Quarterly Review of Economics and Business* 30, 1 (spring): 75-89.

The Mismeasurement of Illegal Drug Markets

The Implications of Its Irrelevance

Peter Reuter

University of Maryland

The largest illegal market in the United States currently is that for illegal drugs. It may be, in terms of share of Gross National Product, the largest ever. Given the poor quality of estimates of income from any other illegal market, the statement is hard to challenge but I believe that one can make a reasonable case that, compared with the other candidates (e.g., prostitution, illegal gambling, counterfeiting), illegal drugs are likely to generate much higher total revenues to sellers. Certainly numbers are presented in a variety of fora, predominantly political, suggesting that it is a major economic activity both in the United States and globally. Figures such as \$500 billion for world sales are thrown around quite glibly.¹

Even brief scrutiny of the global numbers suggests that they are grossly overstated. Though drug markets are large and involve a surprisingly large number of Americans on a part-time basis, the total value of annual sales in the United States is likely to be around \$50 billion, less than 1 percent of Gross Domestic Product and less than 2 percent of personal consumption expenditures.² The global figure is likely to be no more than twice this. One hundred billion dollars represents a large market, but in the context of total global trade flows of almost \$3 trillion, it is a very modest share indeed. That share declines to the trivial when account is taken of the fact that most of the value added is domestic, so that valuing the trade at import prices reduces it to probably no more than \$20 billion.

In recent years, the federal government has developed systematic estimates of domestic expenditures that do indeed provide a reasonable basis for scaling the size of these markets. However, these coexist with an essentially madcap series of federal figures on international production and prices that make a mockery of the whole enterprise. These

estimates and their components are so inconsistent and erratic that they demonstrate what might reasonably be called a “reckless disregard” for the truth. Moreover, though it would be surprising if the government estimates of domestic expenditures were (say) only half of the true value, the year-to-year fluctuations in these estimates may be wrong even in direction, let alone scale.

Does this mismeasurement matter? For those interested in the size of the underground economy, the answer is clearly yes; estimates of the largest illegal market are potentially of considerable significance. However, the estimates were not developed for those purposes, but to help in the development of drug policy. If policy making with respect to drugs were rational, or at least as analytically driven as, say, monetary policy, then the exaggeration would be a serious problem. I shall argue that the numbers are in fact just decorations on the policy process, rhetorical conveniences for official statements without any serious consequences. Indeed, the irrelevance of the numbers is itself a condemnation of drug policy decision making.

This paper has three sections. The first examines the official estimates of drug production, both in the United States and the rest of the world, and sales in the United States, showing how implausible they are. The second describes the process that generates the estimates and its bureaucratic imperatives. Finally, I consider the policy interpretation of the mismeasurement.

The Estimates

The U.S. devotes considerable resources to estimation of drug production by other nations.³ Those estimates, published each year in the “International Narcotics Control Strategy Report” (INCSR), are essentially without competition internationally; certainly they are regarded as more authoritative than any other reports, such as the documents of the United Nations International Drug Control Program, which themselves often cite the INCSR estimates. The failure to include estimates of U.S. domestic production, particularly of marijuana is a conspicuous omission, explained awkwardly by the specific use to which Congress

intended to put these numbers, namely, providing incentives to other countries to improve their drug control efforts.⁴

No detail has ever been published on the methodology of these estimates, beyond the fact that they are generated from estimates of growing area, crop per acre, and refining yield per ton of raw product; the information sources, even the technology used to produce them (for area estimates) are classified. But while the estimation task is clearly a difficult one, the current estimates have unnecessarily low credibility. They show inexplicable inconsistency over time and across sectors of the industry. Some numbers are simply implausible.

Consider as an example of the unprincipled variation over time, estimates of Burmese opium production, consumption, and export. Burma has been estimated to be the largest producer of opium for the illicit market since the early 1980s; since 1989 it may have accounted for over half of world production. In the 1991 INCSR, the opium available for refining (primarily into heroin) for 1988 was estimated to be 679 metric tons; the figure for 1989 was 1600 metric tons, an increase of more than 140 percent.⁵ The difference reflected two factors: (1) A rise in total production, generated by a 25 percent increase in cultivation and an unexplained increase in yield per acre; and (2) a dramatic reduction in exports of opium (as opposed to heroin). The result was that estimated heroin production increased from 68 tons to 128 tons.

Then in the following year, the 1992 INCSR revised the figure for domestic consumption in 1988 downward from 400 tons to 150 tons, reflecting a downward revision in the number of Burmese opium users from 400,000 to 34,000. Yet this did not lead to any change in the estimate of the amount available for refining in 1988. Also odd is that a 25 percent increase in the estimated number of Burmese heroin users from 12,000 in 1988 to 15,000 in 1989 led to a more than doubling of the estimate in Burmese domestic heroin consumption from 2.0 tons to 4.5 tons.⁶

These are figures that do not bear close scrutiny, either individually or collectively. The number of opium users in Burma is unlikely to fall by 90 percent in one year; clearly some analyst in 1992 decided that previous assumptions about the number of opium users in Burma were overstated. Given that the United States has had the thinnest diplomatic relations with Burma since about 1989, it is highly unlikely that this change resulted from acquisition of any substantial new data. Simi-

larly, the decision to increase the estimated average consumption of Burmese heroin addicts in 1989 was probably no more than some analyst's distant judgment. The reasonable stability of the total in recent years hides implausible variation in the components of the estimates. The totals have little credibility.

The problems are illustrated even more graphically by the published estimates for Mexican marijuana production in the late 1980s.⁷ Whereas Burma is distant and hostile, with a government that has little control in many of the major opium producing areas, Mexico is close, an ally for most purposes, and (except for Chiappas since 1994) firmly in control of its drug producing territory. Yet the Mexican estimates have relied on often inconsistent and inadequately described methodologies, leading some analysts, including one at the Bureau of International Narcotics Matters (INM), to conclude that actual production is "unknowable" and that the agency's estimates are at best rather unscientific guesswork.⁸

The preface to the 1988 annual report of the National Narcotics Intelligence Consumers Committee (NNICC, an interagency group chaired by the Drug Enforcement Administration) warns that "there is little reliable data upon which to base estimates of the quantities of drugs involved" (NNICC 1988). Yet, notwithstanding the lack of first-hand evidence of illicit activity that limits accurate production measures, it is often claimed that "the general trends portrayed can be considered reliable" (NNICC 1984). In the 1980s, this was belied by discrepancies in the trends for the two official series for these estimates, provided by NNICC and the INCSR. For example, the NNICC estimate of net production (after eradication) rose in 1986 to 5,460 metric tons from 4,125 the year before; the INCSR estimates for the same years were 2,800 in 1986 and 2,700 in 1985. Note that they differ substantially both in absolute value and in the trend; the NNICC figures were higher and rising while the INCSR figures were lower and essentially flat. In 1987 INM showed a slight decrease (about 5 percent) while the NNICC showed an increase of about one-third. INM officials explained the discrepancy in the mid-1980s as follows: "The Department of State considers its country estimates more reliable because the data were derived principally from aerial surveys. There are, however, no survey data on marijuana cultivation in Mexico; the State Department relied on random reports from Mexico that were

higher than the NNICC figure, which is an extrapolation of seizure data" (General Accounting Office 1988, p. 53). This is less an explanation than an evenhanded condemnation of both estimates!

The fundamental unsoundness of the whole series of estimates was demonstrated vividly at the end of the decade. The U.S. estimate of Mexican marijuana production was dramatically increased in the 1990 INCSR, from a total of 5,700 tons in 1988 to 47,000 tons in 1989, as the result of changes in estimation techniques.⁹ No details of those changes were provided in the published document. Yet it was possible to determine, with no great technical skill, that these figures were implausibly high and should never have been published.

Consider the various ways that Mexican marijuana might have been disposed of: seizures, domestic Mexican consumption, exports to Europe, and exports to the United States. Mexican domestic consumption is thought to be quite low, notwithstanding the important historical association of marijuana in the United States with Mexican immigrant groups in the 1930s, a major factor in the passage of the Marijuana Stamp Act in 1937. The State Department estimated the total in the late 1980s to be 100 tons. There are no reports of Mexican exports to European markets, probably because Mexico is not well located to compete with North African, Middle Eastern and domestic production in Europe, which is primarily a hashish rather than marijuana market anyway. Seizures are usually estimated at a few hundred tons. Nonetheless, assume that all these figures are major underestimates and that the total for seizures, Mexican consumption, and European exports was 12,000 tons.

This would leave 35,000 tons for consumption in the U.S. markets. How many marijuana users would have to purchase the Mexican product to dispose of this? Rhodes et al. (1995, p. 20) estimated that a joint in 1993 had an average of 0.136 ounces (0.39 grams) of marijuana. However, Mexican origin marijuana is believed to be of lower potency, so let us assume that each joint contains 1 gram; this will bias the calculation in favor of finding large figures plausible.

A very heavy user of marijuana consumes about three joints per day. Giving the user time off for colds and work-related drug tests, assume that he or she consumes this amount 333 days a year; this (conveniently) gives a total of 1 kilogram of marijuana annually. That implies that we need 35 million very heavy marijuana smokers to con-

sume 35,000 tons; that would be about half of all persons aged 15-35, the heavy user ages (Chen and Kandel 1995). And this does not take into account consumption of domestic U.S. production or what is imported from Jamaica, Colombia, etc. Yet reasonable estimates of the total number of heavy users (at least one joint per day) are only about one-tenth of the 35 million needed to dispose of the imports from Mexico. Moreover, other estimates of total U.S. consumption have been far lower; e.g., Kleiman (1989) came up with figures for 1985 of only about 5,000 tons and prevalence was still declining in the late 1980s.

Perhaps as the result of this critique, the 1991 INCSR announced a further revision in estimation methodology. The estimate of area harvested was increased, reflecting a dramatic downward revision in the estimate of acreage eradicated. However, a new distinction was introduced between “usable plant yield” and “whole plant yield”; the former, more relevant to consumption estimates, was put at only half the latter. The new 1989 estimate of usable plant available for export after domestic consumption and Mexican seizures was 29,700 tons. Though an improvement over the previous figure, it was still utterly implausible, requiring U.S. consumers to account for far more than is consistent with current estimates of prevalence. Only in 1991 did the official figures start to approach even remotely plausible levels, though still being far higher than other estimates of total U.S. consumption.

Though I will not describe all the twists and turns since then, it is worth noting the current state of estimates. The 1994 INCSR lists the following series for usable plant yield:

Table 1. Estimates of Mexican Marijuana Production
(in metric tons)

1989	1990	1991	1992	1993
30,200	19,700	7,795	7,775	6,280

SOURCE: “International Narcotics Control Strategy Report” (1994).

Only for the 1990 figure is any footnote explanation offered suggesting a change in methodology; yet no one seriously maintains that 1993 Mexican cannabis production is 20 percent of its 1989 level. The wholly implausible 1989 figure remains in the series, though the INCSR makes numerous later year revisions.

Another means of establishing the implausibility of the 1989 and 1990 figures is to consider their implications for Mexican export earnings. The farmgate price of marijuana in Mexico was estimated to be \$50 to \$100 per pound; the border price, after distribution within Mexico, was estimated at \$136 to \$455 per pound. Even taking the lower bound of the border price (which seems very low, given other price estimates for wholesale prices in the United States), the 35,000 tons of marijuana exports generates earnings of \$11 billion for 1989, compared to recorded export earnings of \$40 billion; surely this would be enough to make nonsense of analyses of Mexican currency fluctuations.

These figures should never have survived a review process. As the above analysis suggests, it is easy to establish that they are far outside the plausible range. Yet the estimation process is so detached from analysis of domestic indicators that these figures have been able to survive for many years. Colleagues of mine at RAND have recently developed elementary "mass conservation" models that impose consistency checks on estimates;¹⁰ the Mexican figures do not survive in that kind of framework. The discipline is simple enough that one may reasonably ask why it has not been done before.

International Price Series

The federal government also publishes "The Illicit Drug Wholesale/Retail Price Report," which contains price data for foreign production covering both intermediate and final product. Though prices are in principle easier to observe than quantities, these series show obvious inconsistencies. This is not the place for a lengthy description of all the anomalies of these series. Instead let me just provide two illustrations:

1. In every year for which the data are published, the price for refined cocaine is lower in Colombia than the other two major producing countries (Bolivia and Peru). Yet it is also asserted that most exports of Bolivian and Peruvian cocaine go through Colombia. It is difficult to tell an economically plausible story in which Colombian exporters prefer to buy the more expensive product in foreign countries, and then incur the cost and risk of shipping it into Colombia for re-export. If there is some constraint on expanding production in

Colombia, the price for that cocaine should be higher, representing its transportation cost advantage.¹¹

2. The relative prices of the intermediate product, base, and the final product (HCL) move erratically for both Bolivia and Peru. In 1990 the difference between HCL and base prices in Peru is reported to be \$3,480; in Bolivia in the same year the difference is \$1,525. In 1992 the Peruvian margin has increased to \$4,950, while the Bolivian margin has slipped to \$1,150. Given the low transportation cost for base and the ease of processing, the large and growing difference between processing costs in the two countries seem quite implausible.

The heroin series are no better. The price of morphine, an intermediate product in the refining of illicit heroin, is sometimes recorded as higher than the final product. The export price of heroin from Southeast Asia is stable in the late 1980s, notwithstanding large increases in estimated product and falling prices elsewhere.

Expenditure Estimates

Only recently have domestic counterparts to the foreign production estimates become available from the government. Though estimates of the number of persons using illicit drugs are produced annually¹² and attract a great deal of attention, there has never been a similar interest in quantities and expenditures; the sources and implications of this lack of interest are discussed in the final section. The Office of National Drug Control Policy has been sponsoring such estimates since 1991 but they have attracted little attention or scrutiny.¹³ The most recent available figures are given in tables 2 and 3.

In contrast to the international figures, these estimates are thoroughly documented and have their origins in very diverse types of calculations. The cocaine and marijuana estimates are derived from the household survey data, supplemented by various other surveys that cover populations (e.g., prisoners and college dormitory students) not included in the household survey, or which provide more data about the average amounts consumed by particular groups of users. Given that a small share of all users account for a large share of total cocaine consumption, particular attention is given to ethnographic studies that include consumption figures for heavy users.¹⁴

Table 2. Total U.S. Consumption of Illicit Drugs, 1988-1993
(in metric tons)

	1988	1989	1990	1991	1992	1993
Cocaine						
High	311	382	230	293	280	283
Low	244	286	215	230	224	224
Heroin						
High	9	11	13	10	9	11
Low	5	5	6	4	3	9
Marijuana	910	880	850	750	715	740

SOURCE: Rhodes et al. (1995, tables 5 and 6).

NOTE: The marijuana consumption totals are estimated from figures provided in table 6 on the number of users, number of joints per months per user, and the weight of marijuana per joint.

Table 3. Total U.S. Expenditures on Illicit Drugs, 1988-1993
(in billions of 1994 dollars)

	1988	1989	1990	1991	1992	1993
Cocaine	\$41.1	\$42.5	\$38.9	\$35.2	\$33.1	\$30.8
Heroin	11.2	11.5	10.3	8.2	7.0	7.1
Marijuana	8.9	9.0	9.6	9.0	10.1	9.0
Other drugs	3.2	2.8	2.3	2.4	2.2	1.8
Total	64.4	65.8	61.1	54.8	52.4	48.7

SOURCE: Rhodes et al. (1995, table 8).

Heroin estimates are generated by a different process, since The National Household Survey on Drug Abuse is thought to miss most of the population of frequent heroin users, who tend to be transient if not homeless and difficult to track down for interviews. The estimates are generated by a complex procedure from the number seeking admission to treatment and interviews with intravenous drug users in a large scale ethnographic study.¹⁵

Note first that the marijuana consumption estimates in table 2 are almost an order of magnitude smaller than implied by the Mexican production estimates, though most of that production is destined for

the U.S. market. Even the recent INCSR estimates for Mexico are at least eight times those for total U.S. consumption. That seven hundred tons in 1993 generated almost \$10 billion in retail sales; assuming prices are measured with reasonable accuracy, certainly much more so than quantities, then the notion that the United States imports 6,000 tons, with domestic sales value of \$70 billion seems particularly implausible.

Though the heroin and cocaine expenditure series are carefully produced and show no internal anomalies, it is useful to also examine their relationship to official price series and another international production series produced by the federal government. The price data (table 4) come from a series maintained by the Drug Enforcement Administration called STRIDE (System to Retrieve Information from Drug Evidence).¹⁶

Table 4. Retail Prices for Cocaine and Heroin, 1988-1993
(in 1994 dollars per pure gram)

	1988	1989	1990	1991	1992	1993
Cocaine						
High	\$186	\$165	\$200	\$168	\$163	\$161
Low	146	123	187	132	130	120
Heroin						
High	\$3,007	\$2,713	\$2,199	\$2,543	\$2,614	\$2,553
Low	1,612	1,343	997	1,046	968	837

SOURCE: Rhodes et al. (1995, table 4).

Note the odd relationship between the heroin price and quantity series in tables 2 and 4. Though prices fall by somewhere between one-sixth and one-half (depending on whether one uses the high or low price figures) between 1988 and 1993, total consumption (as reported in table 2) is essentially flat over this period. Moreover, the consumption totals are very low, given estimates of the total number of dependent users (ca. 500,000) and recent estimates of how much users purchase weekly, now that heroin is both purer and cheaper than it was in the early 1980s. One recent study estimated that in New York and Chicago, median weekly consumption was about 700 milligrams,

yielding an annual total of about 35 grams (Boyum and Rocheleau 1994); taken nationally that would generate a total consumption of about 18 tons.

The cocaine price and quantity series also produce some implausible relations. In 1990 it is believed that the Colombian government's crackdown on the Medellin traffickers led to a sharp, temporary, decline in total exports, perhaps partly compensated for by a run-down in inventories of traffickers. Prices rise by between 20 and 50 percent from 1989 to 1990 and consumption falls by between 25 and 40 percent, suggesting a rather higher elasticity of demand than might have been expected but otherwise not an implausible set of figures. In 1991, when the crackdown had ended, prices fell to the 1989 levels but consumption was much closer to the 1990 figure, rising only between 7 and 26 percent.

A review of the official supply side figures in table 5 points to a much more substantial problem of consistency. Separately from the cocaine consumption figures, the same research group also produces a series labeled "Cocaine available for consumption in the United States (after discounting [sic] for Federal seizures)." These are supply side estimates, produced by taking estimates of total production in the source countries (Bolivia, Colombia, and Peru) and then subtracting estimates of what is seized along the way, consumed elsewhere (including Western Europe) and seized by U.S. federal government agencies. The estimates for 1989 to 1993 are produced in table 5. Note that these figures are substantially higher than the domestic consumption estimates reported in table 2; indeed, in 1990 the high end of the range is twice the high end for domestic consumption.

Table 5. Trends in the Cocaine Supply, 1989-1993

	1989	1990	1991	1992	1993
Cocaine available for consumption in the U.S. (metric tons)	361-473	348-463	337-481	376-539	243-340
Retail value of cocaine in the U.S. (in billions of dollars)	\$52-68	\$67-90	\$51-72	\$55-79	\$33-46

SOURCE: Rhodes et al (1995, table B).

Some of the difference may be explained by seizures made by state and local agencies; in 1987 these agencies were estimated to have seized about half as much as their federal counterparts (Godshaw, Koppell, and Pancoast 1987). Over the period considered here, federal seizures were between about 100 and 120 tons. However, even subtracting 50 or 60 tons from the supply side series, the supply side estimate in most years is substantially higher than the domestic estimate.

But of more immediate concern to us is the lack of any notable change in the supply available to the United States in 1990; the drop from 1989 is about 3 percent. This of course is inconsistent with the observed increase in U.S. retail price in 1990 and the reduction in estimated consumption, as well as with the official (and plausible view) that the Colombian crackdown had a substantial effect. Reference to the European market or local seizures does not help explain this inconsistency, since these series are unlikely to exhibit year-to-year changes large enough to explain much of this. To make this lack of change in 1990 even more dramatic, note that the only year of substantial change is 1993, when the estimated supply dropped by about one-third. Yet this decline is not reflected, as one might have expected, in either the domestic consumption or retail price figures; indeed, prices dropped by about 5 percent!

The supply side series undermines the credibility of the domestic estimates. Or it may simply be that any figures generated from distant lands by unknown processes should be viewed as worthless.

Estimates of world consumption or expenditure have no known provenance. Five hundred billion dollars is the standard figure cited in official and unofficial publications. The U.S. State Department, despite its imperialistic estimation of production and (often) consumption in individual foreign countries, does not attempt a global estimate. I believe it unlikely that the global figure is much more than twice the U.S. total. Though the United States accounts for only 5 percent of world opium production (through its heroin consumption), most opium and heroin are sold in Asia at prices that are a minuscule proportion of the U.S. street price. The United States still consumes most of the world's cocaine production. For marijuana, again most of the world production is consumed in countries with much lower retail prices.

The Production Process

I will assume the reader is convinced from the above that the official numbers (prices, quantities, expenditures) are seriously flawed, though the domestic expenditure estimates may be exempt from that judgment. While I might claim to have expertise in this area, having spent ten years in empirical research related to drug policy, clearly none of the flaws exposed here require any special technical skill or insight to discover; all that is required is a willingness to subject the series, individually and as a group, to some modestly serious analytic scrutiny. This section offers some observations as to why the numbers persist in such disarray.

The current drug production estimates are generally detached from the policy process. Though it is worth noting that it would be hard under any circumstance to produce estimates with the authority and precision of the standard indicators of economic activity, one reason they are of low quality is that they simply have no consequence for any senior decision makers. This is an example of a principle called GOGI (garbage out, garbage in), the converse of the usual GIGO (garbage in, garbage out); since no one uses the estimates for any important purpose, they are produced without care or scrutiny. The State Department is required by statute to produce the international production figures annually. It cannot simply claim an incapacity to meet the statutory requirement. Nor, I suspect, is there any justification for large investments in improving the available data. There is still no excuse for producing implausible figures or inconsistent series but, given the lack of political or external analytic scrutiny, not much incentive to do otherwise.

The Mexican marijuana production figures illustrate dramatically the detachment of the estimates from the policy process. The increases and declines in the INCSR estimates have no apparent consequence for U.S. policy decisions. Congress did not feel the need to take new measures against Mexico when suddenly the State Department produced figures suggesting that the U.S. marijuana market was completely dominated by Mexico; nor did the administration make any noticeable change in policy. There was no cry for increased domestic efforts against marijuana because of the sudden discovery that American users

were consuming vastly more than had previously been estimated. The 1989 figure was initially explained simply as the result of changes in estimating techniques, not in the underlying realities. Nonetheless, the implication of the new figures was that Mexico was swamping the United States, and indeed the world, since its 30,000 tons accounted for five-sixths of the officially estimated global total. Surely this would be reason for alarm, for a call to Mexico to deal with a problem that must be much more serious than previously realized. Yet, the harshness of U.S. political rhetoric about Mexico's drug control efforts was much diminished at the end of the 1980s, precisely when these estimates were at their highest levels.

The price series are even more separated from the policy process they are supposed to inform. They serve no bureaucratic purpose that I have been able to discern and are almost never cited politically or in the media. One can reasonably ask why they are produced at all. Perhaps the need of U.S. officials to appear authoritative in public discussions is sufficient motivation to produce the numbers, without providing a good reason for doing the job well.

With no audience for these estimates, agencies are motivated to give them little attention. My impression is that no one of any prominence at the Drug Enforcement Administration or the Department of State ever spent time as an analyst or producer of these estimates. Both agencies, with very different organizational cultures, share a marked lack of interest in numeracy.

These problems are exacerbated by the fact that the National Institute on Drug Abuse (NIDA), the principal federal research agency concerned with illicit drugs, and the Substance Abuse and Mental Health Services Administration, which produces the major indicator series such as the NHSDA and Drug Abuse Warning Network (DAWN) system, play peripheral roles in the estimation process described here. They are much more data- and analysis-oriented agencies than DEA or the Department of State, but as public health bureaus have never been much interested in markets per se, seen as principally the concern of law enforcement agencies. Though the NHSDA does provide critical information for the consumption estimates sponsored by Office of National Drug Control Policy (ONDCP), NIDA itself has only recently started to give attention to questions about expenditures and consumption in the survey.

The Potential Uses of Quantity, Price, and Revenue Estimates

What is needed by way of measurement? Clearly this is entirely a policy-driven enterprise, without any scientific goals. In this respect it differs from the collection of data on domestic use and abuse, which serves a variety of scientific purposes (for example, understanding the etiology of the drug use) as well as programmatic planning. So the central question, putting aside the detail of the statutory requirements imposed on the executive branch by the boundlessly ambitious and imperialistic Congress, is what decisions these figures could usefully inform. At the moment only the certification process makes even a pretense of being connected to these estimates, and a pretense is all that it is. Certification decisions are driven by politics rather than numbers; i.e., the administration forgives U.S. friends even when they are deemed noncooperative (for example, Colombia in 1995), and condemns those with whom it is not on good terms (e.g., Burma, Iran, and Syria).¹⁷ Without a specification of policy needs, the estimates will continue to flounder.

Sensible estimates of quantities, prices, and revenues could, in principle, serve a number of important functions. For example, the efficacy of interdiction may be crudely estimated by comparing seizures with estimated shipments or by estimating the margin between import and export prices, which represents the smugglers' margin.¹⁸ Export prices from individual producing countries could be a useful measure of how stringently that nation enforces laws against growers and refiners. The revenue estimates provide a basis for estimating the impact of asset forfeitures; for example, if total revenues are \$50 billion, then \$2 billion in asset forfeitures constitutes a 4 percent tax, stochastically applied. These kinds of calculations are helpful in developing and understanding the promise and limits of these kinds of programs.

In 1971 Max Singer published an oft-cited article arguing that if one believed the official numbers with respect to the prevalence of heroin addiction, the dependence of addicts on theft, and the price of a heroin habit, then New York City did not exist anymore; it had been stolen by junkies. My first published article on drug policy was an extension of that critique, pointing out that though the measurement enterprise had become more elaborate, it still produced striking inconsistencies (Reu-

ter 1984). Another ten years on, the process has become yet more elaborate. There is coherence and good documentation for the revenue estimates, but the process is still capable of producing much nonsense. Until these series and estimates are integrated into the policy process, they are unlikely to get much better.

NOTES

1. For a scholarly claim citing \$500 billion, see Tullis (1991). For the United Nations use of this figure, see United Nations Drug Control Program (1994, p. 1). A respected journalist, Jessica Mathews (1995), recently retailed this figure, citing federal enforcement agencies.

2. The most systematic estimate, using the household survey described below (note 12), is 1.8 million persons selling in 1991 (Caulkins and McCaffrey 1993). It is likely that most of them sell on a very occasional basis (see Reuter, MacCoun, and Murphy 1990).

3. I infer this from the publications and contacts over the years with participants in the estimation process. No cost figures have ever, to my knowledge, been published.

4. If a nation is not certified as making maximum efforts to control production and export, then the United States will not provide certain aid and will vote against loans from multilateral financial institutions such as the World Bank.

5. Ten tons of opium yield one ton of heroin.

6. The implied annual heroin consumption per addict is 300 grams, 10 to 20 times the figure for the United States. For a recent estimate of weekly heroin consumption per addict of about 700 milligrams, about 35 grams per annum, see Rocheleau and Boyum (1994). No doubt heroin is vastly cheaper in Burma, but incomes are also on the order of 1 percent of U.S. levels.

7. This section draws extensively on Reuter and Ronfeldt (1992).

8. See Kleiman (1989).

9. "New analytic methodologies have enabled the U.S. government to assess more accurately the extent of marijuana cultivation during the past several years" (INCSR 1990, p. 13). The report included no revision of previous years' estimates.

10. Separate models are available for cocaine, heroin, and marijuana (Resetar and Dombey-Moore 1994; Childress 1994a; Childress 1994b).

11. In theory it is possible that the Colombian cocaine prices are not really market prices but instead internal prices for an integrated producer, set low for some intrafirm bargaining. This seems implausible, given that there are some independent cocaine producers in Colombia.

12. The National Household Survey on Drug Abuse (NHSDA), with sample sizes between 20,000 and 32,000 in recent years, produces these estimates.

13. The latest of three published versions is *What America's Users Spend on Illegal Drugs: 1988-1993* (Rhodes et al. 1995).

14. For example, Everingham and Rydell (1994) estimate that 22 percent of cocaine users account for 70 percent of total consumption.

15. Details are provided in Rhodes et al. (1995, appendix 1).

16. On using STRIDE systematically, see Caulkins (1994).

17. In 1995, much to the chagrin of Colombian leaders, the president offered only a qualified certification, invoking a "national interest" waiver clause that allows for certification even if the nation is not meeting the criterion of making maximum effort to suppress the trade.

18. For a price-oriented analysis of interdiction effectiveness, see Reuter, Crawford, and Cave (1988).

References

- Boyum, D., and A.M. Rocheleau. 1994. *Heroin Users in New York, Chicago and San Diego*. Washington, DC: Office of National Drug Control Policy.
- Caulkins, J. 1994. *Constructing a Price Series for Cocaine*. Santa Monica, CA: RAND.
- Caulkins, J., and D. McCaffrey. 1993. "Drug Sellers in the Household Population." Unpublished paper, November.
- Chen, K., and D. Kandel. 1995. "The Natural History of Drug Use from Adolescence to the Mid-Thirties in a General Population Survey," *American Journal of Public Health* 85, 1 (January): 41-47.
- Childress, M. 1994a. *A System Description of the Heroin Trade*. Santa Monica, CA: RAND.
- _____. 1994b. *A System Description of the Marijuana Trade*. Santa Monica, CA: RAND.
- Everingham, S., and P. Rydell. 1994. *Modeling the Demand for Cocaine*. Santa Monica, CA: RAND.
- General Accounting Office. 1988. "Border Control: Drug Interdiction and Related Activities Along the Southwestern U.S. Border." Fact Sheet GAO/GGD-88-124FS, Washington, DC, September.
- Godshaw, G., R. Koppell, and R. Pancoast. 1987. *Anti-Drug Law Enforcement Efforts and Their Impact*. Wharton Econometric Associates, Bala Cynwyd.
- "Illicit Drug Wholesale/Retail Price Report." Various years. Washington, DC: Drug Enforcement Administration.
- "International Narcotics Control Strategy Report" (INCSR). Various years. Washington, DC: Drug Enforcement Administration.
- Kleiman, M. 1989. *Marijuana: Cost of Abuse Costs of Control*. Boulder, CO: Westwood Press.
- Mathews, Jessica. 1995. "We Live in a Dangerous Neighborhood," *Washington Post*, April 24.
- "National Household Survey on Drug Abuse: Main Findings" (NHSDA). Various years. Washington, DC: Substance Abuse and Mental Health Services Administration.
- "National Narcotics Intelligence Consumers Committee Report" (NNICCR). Various years. Washington, DC: Drug Enforcement Administration.
- Resetar, S., and B. Dombey-Moore. 1994. *A System Description of the Cocaine Trade*. Santa Monica, CA: RAND.
- Reuter, P. 1984. "The (continued) Vitality of Mythical Numbers," *Public Interest* (Spring).

- Reuter, P., G. Crawford, and J. Cave. 1988. *Sealing the Borders: The Impact of Military Involvement in Drug Interdiction*. Santa Monica, CA: RAND.
- Reuter, P., R. MacCoun, and P. Murphy. 1990. *Money from Crime: A Study of the Economics of Drug Selling in Washington, D.C.* Santa Monica, CA: RAND.
- Reuter, P., and D. Ronfeldt. 1992. *Quest for Integrity: The Mexican-U.S. Drug Issue in the 1980s*. Santa Monica, CA: RAND.
- Rhodes, W., et al. 1995. *What America's Users Spend on Illegal Drugs, 1988-1993*. Washington, DC: Office of National Drug Control Policy.
- Rocheleau, M.A., and D. Boyum. 1994. *Measuring Heroin Availability in Three Cities*. Washington, DC: Office of National Drug Control Policy.
- Singer, M. 1971. "The Vitality of Mythical Numbers," *Public Interest* (Spring).
- Tullis, L. 1991. *Handbook of Research on the Illicit Drug Traffic: Socioeconomic Consequences*. New York: Greenwood Press.
- United Nations Drug Control Program. 1994. *Drugs and Development: Discussion prepared for the World Summit for Social Development*, June.

The Supply of Youths to Crime

Richard B. Freeman

Harvard University

and

National Bureau of Economic Research

Crime is a major problem for the United States, creating misery for its victims, costing the country substantial resources on the criminal justice system and private crime-prevention activities, and turning many inner-city neighborhoods into social disasters for residents and the rest of the society. Young men, usually out of school young men with limited skills and employment or earnings prospects, commit a disproportionate number of crimes. Inner-city black youths are the most crime-prone group in the society, with other inner-city blacks as victims.

From the 1980s through the early 1990s, the number of persons incarcerated increased massively, incapacitating many criminals and increasing the risks of being caught and penalized for crimes. These factors should have greatly reduced the crime rate. Yet the standard administrative measure of crime, the FBI's Uniform Crime Reports (UCR), stabilized in the 1980s. It fell from 1980 to 1984, then rose through 1991, then fell modestly through 1993. The standard survey measure of victimizations, the National Crime Victimization Survey, which typically shows two to four times as much crime as the UCR due to crimes not reported to the police, recorded a sizable drop in crime, but the declines was far below what could be expected on the basis of the incapacitation of so many criminals and the increased risk of apprehension and incarceration. Crimes that are best measured—murder and auto thefts—showed no sign of falling: murders stabilized, while auto thefts increased.

Why has crime remained high? If, in 1976, a political leader had announced a tough anti-crime program that would triple the number incarcerated and increase the risk of imprisonment for crime, we would all have expected drastic reductions in crime rates. But no such drastic reductions occurred. The economist naturally seeks an explanation in

terms of the labor market determinants of the supply of young men to crime. For crime to persist at high levels despite massive incarcerations, there must be offsetting increases in the returns to crime or an outward shift in the supply schedule of young men due to other factors, such as increased drug use, family breakdown, social disorder, etc.

In this paper I explore whether changes in labor market factors may explain some of the persistence of high crime in the United States. First, I show that participation in crime among American men has become so large that crime is an integral part of the lives of many men. Then I examine the argument that the 1980s-1990s job market was an important factor maintaining the crime rate. I present evidence that youths often combine crime and legitimate work and sketch out a “foraging” model of the supply of youth to crime that helps explain this pattern of behavior. There is a brief conclusion.

Dimensions of Criminal Participation

How many Americans are involved in serious crime?

A useful identity for examining criminal involvement decomposes the number of crimes per capita as follows:

$$(1) \quad \begin{aligned} \# \text{ crimes/population} &= (\# \text{ in crime-prone group/population}) \\ &\times (\# \text{ who commit crimes}/\# \text{ in crime-prone group}) \\ &\times (\# \text{ crimes}/\# \text{ who commit crimes}). \end{aligned}$$

The first term on the right-hand side of (1) measures the share of the population typically involved in crime. For simplicity, I take the crime-prone group as consisting of men, disproportionately those aged between 18 and 34 years. Despite considerable attention given to the effect of the age distribution on crime, changes in this share have had only modest effects on the crime rate (Phillips and Votey 1990), so I will not focus on demographic developments in this paper.

The remaining terms in (1) reflect the behavior of the crime-prone group. The second term is the criminal participation rate—the proportion of the group who commit crimes. The third term is the average intensity of criminal activity—the number of crimes committed per criminal.

There are four ways to estimate the criminal participation rate. First, administrative records on *the number apprehended and convicted of crime* (and thus “under the supervision of the criminal justice system”) provide one estimate of the number involved in crime in the recent past. The number is a lower bound because it excludes criminals who have not been apprehended or who have “beaten the rap.” In 1993 roughly 1 man was incarcerated in the United States for every 50 men in the workforce. For every person incarcerated, an additional 2.1 were on probation and 0.5 were on parole. This gives a figure of 1 man “under the supervision of the criminal justice system” per 15 men in the workforce. One in 15 involved in such serious crime as to be under supervision? My immediate reaction when I did this calculation was to say, I must have made a numerical mistake. The number seems outlandishly large. But here are the estimates for 1993:¹

859,400	men in state or federal prison
428,800	in jail
1,288,200	total incarcerated
2,690,400	probated after conviction for crime
600,700	on parole
3,291,100	probated or paroled
4,579,300	under supervision of criminal-justice system
69,600,000	male workforce

Since most crimes are committed by younger men (aged 18-34), the estimate of the criminal participation rate of young men is even larger: 1 out of every 9 men aged 18-34 in the United States is under supervision of the criminal justice system. The figures for blacks are: 1 black man in prison for every 11 men in the workforce; 1 black man under supervision of the criminal justice system for every 4 men in the workforce. Combine race and age, and you find the remarkable fact that 37 percent as many black men aged 18-34 are under supervision of the criminal justice system as in the labor force.

These numbers are a decimal place beyond comparable statistics in other advanced countries. Since many of the incarcerated are recidivists in crime—studies show percentage re-arrested are on the order of 50 percent to 70 percent depending on the number of years covered (Needels 1993) and have poor employment records years into the

future (Freeman 1992)—this population can be viewed as a relatively permanent part of U.S. society—our equivalent of Europe’s long-term unemployed. Leaving prison is not like leaving long-term unemployment—a step back toward a relatively permanent legitimate job. It is often simply a return to criminal life until the police apprehend the ex-offender again.

The second source of data on criminal behavior is the *self-reported criminal activity of individuals*. These data are possible contaminated by reporting bias. If people don’t admit to criminal activity, self-reported crime would understate criminal participation. If, on the other hand, young men think it “cool” to claim to commit crimes, self-reported numbers would overstate criminal participation. Criminologists have explored these biases through studies that ask people whether they had been arrested and then comparing their responses to police records. The evidence shows that young white males report criminal activity roughly accurately, but that young black males under-report criminal participation (Hindelang and Hirschi 1981), possibly because criminal involvement among blacks extends beyond “hard core” youths. The proportion of young men who admit to committing crimes on major surveys ranges from 20 percent to 40 percent (Freeman 1992).

The third source of data is the *number of arrests*. To be sure, not everyone arrested is guilty of crime, but the number of arrestees does indicate the number of persons whom police believe have committed crimes—X might be wrongfully arrested but somewhere there is Y who in fact committed that crime. The number of arrests in the United States is immense. In 1992 there were 9.9 million arrests of men (including those under 18) and 2.2 million arrests of men for the crimes judged most serious by the FBI index of crimes (U.S. Department of Justice, *Sourcebook*, 1994, table 33). Most of those arrested were between the ages of 16 and 44 (85 percent), but a surprising 13.3 percent of those arrested for serious crimes were aged 13-15. Taking as the base population the male civilian labor force in 1992 (69.2 million) gives a ratio of arrests per man in the labor force of .14 overall and of arrests for index crimes of .032 per man. Since 30 percent of the male workforce is 45 or over, the ratio of arrests per labor force participant below that age is substantially higher. Using the crime module of the National Longitudinal Survey, I calculate that there are 2.3 arrests per

young man arrested in a given year. Dividing the ratio of arrests per man by 2.3 suggests that the number arrested was about 6 percent of the male workforce in 1992, and the number arrested for index crimes was 1.4 percent of the male workforce.

The fourth source of data on criminal behavior is *number of crimes committed*. From (1), it is apparent that given an estimate of the average number of crimes per criminal, we could use the number of crimes on the UCR or victimization surveys to determine the criminal participation rate. There are two sources of data on the number of crimes per criminal. Some surveys of prisoners ask inmates for detailed histories of crimes, arrests, and so on—which can be used to estimate crimes per period of time. These surveys yield estimates on the order of 12 to 15 crimes per year (Piehl and DiIulio 1995). Some surveys of youths, including the National Longitudinal Survey of Youth (NLSY), ask similar questions of those who are not incarcerated. Using the NLSY, I calculate that the average number of crimes per young man who admitted to crime was 6.6. Given the number of crimes presumptively committed by men, this implies that 2.6 percent as many men committed crimes as were in the workforce in 1992.²

Whichever of these estimates one prefers, it is clear that a large proportion of American men, particularly young men, are involved in criminal activity.

The Trend in Incarceration and Criminal Propensity

Exhibit 1 shows that from the mid-1970s or so through 1993 the number of persons in prison or jail in the United States increased massively. The rate of increase in the 1980s averaged 8.5 percent per year, so that in 1993 there were over three times as many persons in prison or jail as in 1976! The combination of an increasing number of criminals incarcerated and unable to commit crimes and a roughly constant UCR crime rate implies that the number of crimes committed by the noninstitutional population rose; either criminal participation among nonincarcerated men went up, or the number of crimes per criminal went up to compensate for the incarceration of so many criminals. The falling rate of victimization in the victimization survey could yield a

contrary conclusion, but in fact the predicted drop in victimizations due to incapacitation exceeds the actual drop by so much as to imply a large increase in the rate of criminal participation (Freeman 1994).

Exhibit 1. Prison and Jail Population in the United States, 1947-1992

SOURCE: U.S. Department of Justice, Bureau of Justice Statistics. 1992. *Sourcebook of Criminal Justice Statistics 1991*, pp. 611, 636 and U.S. Department of Justice, Bureau of Justice Statistics Bulletin by Darrell K. Gilliard and Allen J. Beck. 1994. Prisoners in 1993, June, NCJ-147036.

NOTE: Estimates of jail population before 1983 based on prison population.

For the period 1977 (prior to the large increase in the jail and prison population) to 1992, I have estimated the trend in criminal activity by the noninstitutional population. I calculated the reduction in the number of crimes that the increased number of prisoners should have produced under hypotheses about the number of crimes the newly incarcerated would have committed on the street. Then I compared this expected number of crimes to the actual number of crimes in the UCR or victimization survey. The ratio of crimes committed to the predicted number gives an index of the *Propensity to Commit Crime*—a mixture of criminal participation and intensity of criminal activity that reflects the overall involvement in crime by the noninstitutionalized population.

The following example shows precisely how I calculated the index. Assume a population of 100, in which there are 40 crimes committed per year, giving a crime per person of .40. If each criminal commits 10 crimes, there are 4 criminals in the population. When the government apprehends and imprisons 2 criminals, the number of crimes should, all else the same, fall to 20, and crimes per person would fall roughly in half, to .204 ($= 20/98$). Any crime rate beyond .204 implies an increase in the index of crime propensity. If 30 crimes were committed, the propensity would have risen 50 percent; if 40 crimes were committed, the propensity would have risen 100 percent. If the number of crimes committed per criminal was constant, the criminal participation rate must have risen by those amounts. Incapacitation “created opportunity” for new entrants into crime.

Exhibit 2 gives my estimates of criminal propensity from 1977 to 1992. Based on UCR data, I assume 10 crimes committed per person incarcerated. The calculations indicate that the increase in the prison/jail population should have more than halved the crimes committed per male. But between 1977 and 1992 crimes per male rose, albeit modestly. Reconciling these trends, I estimate that the propensity for criminal activity by noninstitutionalized men increased by 163 percent! My suspicion is that most of this rise is due to an increase in the criminal participation rate. Also reported in exhibit 2 are similar calculations using the victimization data. Because there are more victimizations than UCR crimes, I assume the number of victimizations per criminal to be 30. In this case, I estimate that criminal propensity increased by 80 percent from 1977 to 1992. Alternative estimates of crimes per incarcerated person would affect the extent of the rise in propensity but not its direction.

Exhibit 3 uses supply-demand schedules to show alternative ways to account for the upward trend in criminal propensity. The “demand curve” in the diagram is not citizens’ demand for being mugged, robbed, or murdered (though for some crimes, such as drug purchases, it could be so interpreted) but a schedule of criminal earnings opportunities. The schedule slopes downward because more criminals presumably reduce the potential rewards from an additional criminal act. The supply curve is depicted as a response to criminal earnings; it will shift outward if legitimate earnings fall and inward if they increase. In panel A, the supply curve of crime is upward sloping. When more criminals

are imprisoned, the supply shifts inward, which raises the wages of criminals. Assuming unchanged criminal opportunities, the only way to maintain a given number of crimes is for the noninstitutionalized to commit more crimes, shifting the curve back to its original position. Panel B shows that an increase in the criminal opportunities curve—due, say, to an increased consumer demand for drugs, for instance—could have a similar effect, maintaining the number of crimes at higher rewards, despite huge incarceration. Panel C gives a qualitatively different picture: the elasticity of supply to crime in this case is infinite (presumably because crime pays off much more than legitimate work). This means that there is no incapacitation effect on crime: the police arrest Joe for dealing drugs on main street and presto! Harry takes Joe's old place on the street. Given that economists rarely find infinite elasticities, I regard this as an unlikely situation, but it highlights the point that the more elastic the labor supply curve the less effect will incarceration have on the crime rate.

Exhibit 2. Index of Propensity for Crime, 1977-1992 (1977=1.0)

SOURCE: Calculated by dividing the actual number of crimes by the expected number of crimes, where the expected number is estimated by taking the 1977 number of crimes minus the increased number of male inmates from 1977 to the given year times the postulated number of crimes that inmates would have committed: 10 in the UCR and 30 in the victimisation survey. The graph in Freeman (1994) using these data divides the crimes by the male population 16 and over.

Exhibit 3. Supply of Crime, Criminal Opportunities, and Incapacitation

In short, economics suggests that we look for an explanation of the persistent high level of crime despite mass incarceration in potential increases in the rewards to crime and a highly elastic supply curve of youths to crime.

Changes in Relative Rewards to Crime

As a first approximation, consider the crime decision as a dichotomous choice between legitimate and criminal work. The person considering crime compares the present value of earnings from crime, net the loss of earnings due to being apprehended and imprisoned, with the present value of earnings from legitimate work; weighs the riskiness of crime; and makes his decision. Assuming that the marginal criminal is risk-averse, there will be a compensating differential premium from crime. The three factors that enter the calculus are legal earnings, risks and extent of penalties, and illegal earnings.

From 1973 through the 1990s, the real earnings of the less-skilled young men who constitute the bulk of the crime-prone population fell sharply. In the 1980s, the position of these men in the earnings distribution also fell as overall income inequality skyrocketed. Moreover, despite the putative job-creating effects of pay reductions, their hours worked also fell and their employment/population rate fell relative to that of the more skilled. The magnitude of the worsened job market opportunities for less-skilled young men was sufficiently large—drops in real earnings of 20-30 percent, accelerating in the 1990s (Mishel and Bernstein 1994)—to have at least potentially raised their propensity to choose crime.³

Working in the opposite direction is the increased likelihood that an individual involved in crime would be incarcerated in the 1980s. Justice Department data suggest that the clearance rate for crimes known to the police has been relatively constant at about 20 percent (U.S. Department of Justice, *Sourcebook*, 1994, table 4.24). The combination of a rising rate of imprisonment, constant clearance rate, and stable crime rate implies that those apprehended for crimes were incarcerated more frequently or longer. As time in prison has trended downward, the data imply a greater probability of incarceration upon apprehension for

crime. Indeed, Langan (1991) estimates that the chance that someone who commits a crime would be imprisoned rose for individual crimes from 1974 to 1986: for robbery, it increased by 47 percent (from .19 in 1974 to .28 in 1986); for burglary, by 61 percent (.095 to .153); for larceny, by 59 percent (.017 to .027); for drugs by 108 percent (table 4). Mendel (1995) reports that between 1975 and 1989 the expected prison time for a violent crime nearly tripled (p. ii). The increased probability of incarceration should, all else the same, lower the present value of crime.

What is the net effect on the returns to crime versus legitimate work from the increased probability of incarceration upon arrest and the reduction in legitimate earnings for crime-prone youths? This is a difficult question, whose answer depends on the crime, whether or not it involves violence as well as property crime, and so on. My back-of-the-envelope calculations suggest that the 30 or so percent drop in legitimate earnings dominates the increased probability of incarceration. The largest increase in imprisonment rates shown by Langan is .091 for robbery. If the person imprisoned is locked up for 1.25 years (median time served before release from prison in 1986 was 15 months according to Langan 1991, table 1), the loss in lifetime earnings from increased chance of incarceration for robbery would be, roughly, 11 percent ($1.25 \times .091$). This falls short of the 30 percent drop in real earnings from legitimate work. Since time locked up will differ depending on repeat offenses, however, and since imprisonment reduces future legal employment opportunities and possibly increases future illegal opportunities, this is an exceedingly crude calculation. At the minimum, however, it suggests that the increased chance of incarceration did not “dominate” the reduction in legitimate earnings in the returns to crime calculation.

Although criminal earnings are—for various reasons—difficult to estimate, it is difficult to argue that they have fallen in real terms since the 1970s.⁴ The limited evidence that I have examined suggests the opposite, at least for youth. This information consists of responses to survey questions on *perceived* criminal and legitimate earnings and employment opportunities at the outset of the 1980s and at the end of the decade. In 1980 the NBER Inner City Youth Survey asked youths in Boston, Chicago, and Philadelphia whether they thought they could make more “on the street” than in a legitimate job. It also asked them

about their perceptions of the availability of criminal opportunities. The 1989 Boston Youth Survey, conducted at the peak of the booming "Massachusetts Miracle" job market, asked the same questions. Between these dates, the proportion of youths who reported that they could earn more on the street went up, from 31 percent in the three cities and 41 percent in Boston in 1980 to 63 percent in Boston in 1989. Similarly, the proportion who said they had "chances to make illegal income several times a day" roughly doubles over the period, to reach nearly 50 percent in 1989 (Freeman 1992).

Consistent with this, youths who made money from crime in the 1980 NBER Survey of Inner City Youth reported average annual criminal earnings of \$1,807 per year, whereas in the 1989 Boston Youth Survey youths reported average criminal earnings of \$3,008—which, deflated, implies a real increase of some 5 percent.⁵ These annual criminal earnings are, the reader will note, hardly large numbers. Even those who said they committed crimes weekly in 1989 reported earnings of \$5,376 over the year—hardly the stuff of riches. Still, transformed into "hourly pay," these figures imply hourly earnings from crime of around \$10.00 for criminal activity in Boston in 1989. This exceeds the \$7.50 youths reported from legitimate work and substantially exceeds take-home pay from legitimate work, after social security and tax deductions. Estimates of earnings for adult criminals tell a similar story. Reuter surveyed drug dealers in Washington, D.C. and found that they earned \$2,000 per month net of expenses, which he translated into \$30.00 per hour, making drug selling "much more profitable on an hourly basis than are legitimate jobs available to the same person" (Reuter, MacCoun, and Murphy 1990, p. viii). He further estimated that the illegitimate earnings of drug dealers exceeded their legitimate earnings by enough to make it financially worth their while to spend one year in jail for every two years they sold drugs. In the 1986 Inmate Survey I found that criminals who said all of their earnings came from crime made \$24,775 per year (Freeman 1993), a figure comparable to Reuter's \$2,000 a month; but so few criminals reported that all their earnings were from crime, this is an unrepresentative statistic. Wilson and Abrahamse (1992) stress that criminal earnings from burglary/theft, robbery, and swindling are not that high and may fall short of the legitimate earnings available to those criminals (though not necessarily on an hourly basis).

My bottom-line assessment is that the returns to crime increased relative to those from the job market for crime-prone less skilled men in the 1980s, and that the hourly rewards to crime exceeded the hourly rewards from work. Assuming this to be the case, the next question is whether the magnitude of supply response to the change in returns is large enough to have contributed significantly to the observed trends.

Supply Responses

There are five pieces of evidence that suggest that supply responses may be sufficiently large to play a role in the rise in criminal propensity.

1. The demographics of the criminal population

Those who commit crimes consist disproportionately of persons with low legitimate earnings prospects—the young, the less-educated, persons with low test scores, etc. As long as these people do not have commensurately lower criminal earnings prospects and as long as they respond to differential legal/illegal incentives, this distribution is qualitatively what virtually any labor supply model would predict. In fact, evidence from the NLSY suggests that greater schooling, age, and test scores pay off more in the normal job market than in crime, so that the observed demographics of the criminal population is consistent with differing rewards for characteristics and responses by individuals to those characteristics.⁶

2. The estimated effect of joblessness on crime

Literature reviews (Chiricos 1987; Freeman 1983, 1994) find that higher unemployment is associated with greater crime. Most studies comparing crime rates and unemployment rates across areas find that high unemployment areas have high crime rates, though coefficients of response are not large, and an occasional study finds little relation. Most time series studies also find that unemployment or related measures of aggregate labor market activity are associated with rises in crime, but cyclical changes in labor market tightness cannot explain secular changes in criminal propensity. Perhaps most striking, comparisons of individuals show that those who commit crimes are more likely to do so when they are unemployed. This is consistent with a

joblessness effect on crime though it could simply reflect the timing of criminal behavior. Finally, longitudinal evidence on the correlates of violent criminal behavior over time shows that persons who have engaged in "serious violent behavior" are more likely to terminate this if they are employed than if they are unemployed (Elliot 1994, table 1).

3. Estimates of the effect of inequality on crime

Some studies have explored the relation between inequality in a geographic area and the rate of crime. Given that criminals are low-skilled, greater inequality is a plausible indicator of the rewards to crime (robbing the wealthier) compared to low-skill work. Most studies find that more inequality is associated with more crime (see the reviews by Chiricos 1987; Freeman 1983, 1994). Land, McCall and Cohen (1990) even report that homicide rates are correlated with measures of inequality across cities. In the most comprehensive work to date, Lee (1993) found a substantive positive relation between inequality and crime rates across SMSAs in 1970 and 1980. His estimated effect of inequality on crime suggests that the increased inequality in the 1980s induced a 10 percent increase in the UCR, which falls far short of the observed rise in the propensity to commit crime in figure 2. When Lee compared *changes* in inequality and crime among SMSAs, however, he found no relation, possibly because of the decreased signal-to-noise ratio in changes in inequality, but also possibly because the cross-area relation reflects an omitted area variable rather than a true inequality-crime link.

4. Estimates of the crime behavior of individuals

Studies that examine the effects of incentives on the criminal behavior of individuals are potentially the most compelling. In the first such major study in economics, using the NBER Inner City Youth Survey, Viscusi (1986) found that perceptions of risk combined with earnings opportunities influence the supply of young blacks to crime. Using the same data, I found a significant positive relation between criminal participation and whether individuals perceived that they could earn more on the street than in the job market (1987). More recently, Grogger (1994) estimated an econometric model of the crime behavior of young men in the NLSY that makes extensive use of the fact (to be examined shortly) that many youths who engage in crime also work. His estimated supply elasticity to crime is roughly unity: a 10 percent decrease in the real wages of youths would increase their crime rate by nearly 10

percent. Applying this elasticity to the observed drop in real earnings of less-skilled young men, he predicts a 23 percent increase in crimes committed by these youths from the mid-1970s to the late 1980s, which he points out is of comparable magnitude to the 18 percent increase in the index arrest rate for the young over the period.

These studies, particularly Grogger's, should move priors toward the job market explanation of the rise in criminal propensity. But none of the studies are smoking guns. Viscusi and Grogger used sophisticated structural models—economists' lenses as it were—for viewing the evidence. My analysis used self-reported incentives and could simply be a consistency check: yes, people involved in crime thought it paid off better than those who did not. Is it possible to provide something more?

The NLSY asked one question in its 1980 crime module that can be used to examine supply responsiveness to the relative rewards to crime. The question asked respondents the proportion of their income that came from illegal activity. Holding fixed time worked at legitimate jobs, and the number of crimes committed, persons whose income consists disproportionately of illegal earnings will have higher criminal pay relative to legitimate pay. They should thus be more deeply involved in crime than others, and all else the same, end up incarcerated in the future.⁷

In the NLSY the proportion of income from illegal sources in 1980 does in fact help explain incarceration years into the future. Exhibit 4 documents this claim with a simple linear probability analysis in which the dependent variable is being interviewed in jail in 1983, 1986, and 1989 for a sample of young men who reported some criminal earnings in 1980. For simplicity of presentation, I include only two regression controls: the numbers of crimes committed and weeks worked in the past year. Calculations that include persons who report no criminal income or that add additional controls tell the same basic story: the higher the relative pay from crime in 1980, the greater the chance a young man is incarcerated in ensuing years. The magnitude of the relative criminal earnings effect varies among the years, seeming to rise over time: it averages around 0.10, which given the proportion who go to jail (.03) and the mean proportion of income from crime (.20), implies an elasticity of supply to relative rewards on the order of 1.5. This is in the same ballpark as Grogger's estimate, and implies that the

decline in the legitimate wages of youths might account for roughly one-third of the increased criminal propensity

Exhibit 4. Linear Probability Estimates of the Effect of Illegal Income on Future Incarceration

	Jail 1983	Jail 1986	Jail 1989
Mean of dependent variable	.03	.04	.04
Percentage of income that is illegal, 1980	.05 (.03)	.10 (.03)	.14 (.03)
Control variables			
Number of crimes committed (/100)	.02 (.10)	.04 (.01)	.07 (.01)
Weeks worked 1980 (/100)	-.03 (.03)	-.01 (.03)	.03 (.04)
R ²	.01	.04	.10
F	4.64	11.30	29.50
Number of observations	938	790	772

SOURCE: Calculated for sample of out of school youths in NLSY who reported some illegal income in 1980.

5. The labor supply behavior of men with falling real earnings Topel (1993) and Juhn, Murphy, and Topel (1991) have shown that time worked by men in the lower deciles of the earnings distribution fell in the 1980s as their real earnings fell. This relation has the flavor of a labor supply response to falling real earnings. Interpreted in this way, they estimate that the elasticity of labor supply on young men in the lower deciles of the earnings distribution is on the order of 0.20 to 0.30 (table 9). While, as I shall shortly document, many youths commit crimes while working, and while the supply of time to crime is not the simple complement of the supply of time to work, the finding that low-paid men worked less as their real wages fell is consistent with the claim that as real wages fell these men were more involved in crime.

Work and Crime: A Foraging Model

Treating the decision to engage in crime as a dichotomous choice between legal and illegal work misses an important aspect of criminal activity. Because most criminals are self-employed, and because the U.S. job market is characterized by considerable mobility and flexibility, it is easy to combine work with crime at a point in time or to move between the two activities over time. Joe holds a job, and mugs and robs someone he meets on a dark empty street, sells some drugs on the weekend, or steals from his employer. Maybe he sells drugs for a while, decides the street is too dangerous, gets a legitimate job for a while, loses that job, and goes back to selling drugs. Ethnographic research by Reuter, MacCoun, and Murphy (1990), Fagan (1991), and Hageborn (1994) shows that legal and illegal work often overlap among young drug sellers.⁸

To see how much overlap exists between legal and illegal work, I have examined the work activity of persons involved in property crime in the NLSY. My analysis supports the view that crime and work are not exclusive activities, save possibly for those sufficiently involved in crime that they end up in jail or prison in the near future. Exhibit 5 records the employment status of young men according to four measures of criminal activity: admitted committing a crime, earning illegal income, being charged with a crime, and ending up incarcerated in the following year. The sample is limited to out-of-school youths not involved in military service. There are differences in employment between those involved in crime and those not involved in crime: a 3-point difference between those who committed and those who did not commit crime; a 7-point difference between those with positive incomes from crime and those without such income; a 13-point difference between those charged with crime and those not charged. But these differences are far below the magnitudes that would support a crime-employment dichotomy. The only grouping that yields something close to that is between youths who end up incarcerated a year later and the rest of the sample—a 35-point difference in employment.

Ecology models of foraging behavior (Stephens and Krebs 1986) offer an insightful way to analyze the tendency for youths to engage in both illegal and legal work activities, either simultaneously or by moving back and forth between them. These models apply economic opti-

mizing analysis to the problems faced by animals that forage for food. The animal must make several decisions in a short period of time: whether to “prey” on a particular food source it encounters or turn that prey down to search for better prey; whether to exploit opportunities in a given patch or search for new opportunities; and so forth. The parallels with youths “foraging” for earnings, legal or illegal, are striking. Youths must decide whether to mug someone they meet on the street; take a short-term job when they encounter an offer; burgle in the local community or try some adjoining area; sell drugs to employees, if working, or to customers in a street market.

Exhibit 5. Employment in Survey Week by Criminal Behavior of Out of School Non-Military Youth

Criminal group (#of observations)	Employed at survey week 1980
Admitted committing property crime in 1980 survey	
Yes (2,369)	70.3
No (1,847)	73.3
Reported positive illegal income	
Yes (952)	66.0
No (3,265)	73.2
Charged with crime	
Yes (744)	58.6
No (3,279)	71.5
Jail in following year	
Yes (46)	30.4
No (4,223)	65.5

SOURCE: Tabulated from NLSY, with youths in school coded as missing. In these tabulations I have also excluded those in the military. Inclusion of youths in the military reduces the employment difference between those who reported crime and those who did not (strengthening the argument in the text) but does not noticeably affect the difference in employment rates for those in jail the following year. The admitted crimes figures are based on people who said they committed any of the following crimes in the past year: shoplifting, stealing, using force to obtain things, selling drugs, conning someone, stealing automobile, breaking into building, aiding gambling operation. This leaves out some violent nonproperty crimes. Their inclusion increases the numbers committing crime without changing the results.

The foraging models direct attention to differing "reservation wages" to various money-making activities and the determinants thereof. When returns fall below the reservation wage, the youth will reject an opportunity and go on to something different. According to the NBER Inner City and Boston Youth surveys, young men in inner-city poverty areas encounter many illegal and legal opportunities in a relevant time period: McDonald's may be hiring this week; Jones Construction may need a laborer; robbers may need someone to fence stolen goods; an elderly woman may wander along the wrong street; a car with an expensive stereo system may be parked in an alley. In a world where short-run legal and illegal earnings opportunities arrive more or less randomly, it is natural for individuals to move between them, commit crimes while working, or take a legitimate job if one happens to be available even when engaged in criminal activities. If this hypothesis is correct, and the behavior of crime-prone youths is similar to that of foraging animals as opposed to that of adults with permanent careers, the supply of youths to crime will be quite elastic, consistent with the observed failure of incapacitation to reduce crime.

Conclusion

In this paper I have shown that increased incarceration of criminals has failed to arrest the nation's massive crime problem because of an offsetting increase in the crime propensity of noninstitutionalized men. I presented evidence that part of the problem seems to lie with the deterioration of the job market for less-skilled young men. It would be fitting to conclude by offering a program or policy that would improve the job prospect of the less skilled, and thus deter crime. While some programs for reducing juvenile delinquency have modest beneficial effects and some crime prevention programs may work (Mendel 1995), I do not believe at this time we have a blueprint for successful job-creating or job-enhancing programs that would offset the fall in the market for the less-skilled and thus reduce crime. What we do have is evidence that incarceration, which is highly costly (a year in prison costs as much as a year at Harvard, as they say), has not reduced the rate of crime (UCR) or has reduced it less than we would have

expected (victimization survey). The expense of incarceration is such that it behooves the nation to experiment with, and study carefully, programs to enhance the legal earnings opportunities of crime-prone young men and to try other modes of crime prevention. Even a modestly successful employment program that induced some to forego crime is likely to meet any plausible benefit-cost test, from the savings in the cost of incarceration as well as in the lower crime rate.

NOTES

1. The 1993 jail figures are estimated from 1992 data; the numbers probated and paroled are estimated from 1990 data. The estimates simply assume that the ratios of the missing data to the number in prisons remained constant over time.

2. I assume that 80 percent of crimes are committed by men, since approximately 80 percent of arrests are of males. The number of crimes in 1992 was roughly 15 million. Dividing 15 million by 69 million men in the labor force and multiplying by 80 percent yields an estimated crime per man in the workforce of 0.17. Dividing this by 6.6 gives the figure in the text.

3. The exact magnitude of the decline in real/relative earnings depends on the specific measure of earnings chosen, the deflator, years picked, the age and skill group chosen, etc., but it is invariably large.

4. One reason is that most criminals are self-employed, and thus do not face a market wage but rather an earnings opportunity schedule in which hourly pay depends on the hours of work they choose. In the Boston Youth Survey, those who committed a single crime in the past year earned \$752, whereas those who reported committing crimes once a week or more earned \$5,376, or \$100 or so per week—considerably less per crime. A second reason is that self-reported criminal earnings may be inaccurately reported: Wilson and Abrahamse (1992) suggest that the incomes that inmates claim to have earned from various crimes are far higher than those crimes could plausibly yield.

5. Here, I take an average of the 1979 and 1980 deflators for the earnings in the Inner City Survey, since the survey covered both years. Using the 1979 deflator gives an estimated 3 percent drop in earnings, which is far short of the drop in legitimate earnings.

6. Since criminal earnings are poorly measured, it is not easy to document this claim. In the NLSY I regressed the share of income from illegal sources on number of crimes reported, weeks worked in the year and three human capital measures: years of schooling, age, and AFQT score. The coefficients on all three human capital measures were negative and significant, implying that schooling, age, and AFQT lowered illegal income relative to legal income.

7. Because the NLSY has never repeated the crime module, evidence on future crime behavior is limited to whether or not the respondent was interviewed in jail or prison.

8. The “doubling up” of legal work and cocaine sales in the Fagan and Reuter, MacCoun and Murphy studies indicates that for many young men, illegal work may be temporary or transitional work that supplements difficult low-wage or otherwise unsatisfactory work. For others, legal work provides options to riskier illegal work, or perhaps broadens markets for sellers of illegal goods or services.

References

- Chiricos, Theodore. 1987. "Rates of Crime and Unemployment: An Analysis of Aggregate Research Evidence," *Social Problems* 34, 2 (April): 187-211.
- Elliot, Delbert. 1994. "Longitudinal Research in Criminology: Promise and Practice." In *Cross National Longitudinal Research on Human Development and Criminal Behavior*, G.W. Westekamp and H.J. Kerner, eds. Netherlands: Kluwer Academic Publishers.
- Fagan, Jeff. 1991. "Drug Selling and Licit Income in Distressed Neighborhoods: The Economic Lives of Street-Level Drug Users and Dealers." In *Drugs, Crime and Social Isolation*, Adele Harrell and George Peterson, eds. Washington, DC: Urban Institute Press.
- Freeman, Richard B. 1983. "Crime and the Labor Market." In *Crime and Public Policy*, J.Q. Wilson, ed. 89-106. San Francisco: Institute for Contemporary Studies.
- _____. 1987. "The Relation of Criminal Activity to Black Youth Employment," *Review of Black Political Economy* (Summer-Fall): 99-107.
- _____. 1992. "Crime and the Employment of Disadvantaged Youth." In *Urban Labor Markets and Job Opportunity*, George Peterson and Wayne Vroman, eds. Washington, DC: Urban Institute Press.
- _____. 1994. "The Labor Market." In *Crime*, J.Q. Wilson and Joan Petersilia, eds. 171-192. San Francisco: Institute for Contemporary Studies.
- Grogger, J. 1994. "Criminal Opportunities, Youth Crime, and Young Men's Labor Supply," Department of Economics, University of California, Santa Barbara, February.
- Hageborn, J. 1994. "Homeboys, Dope Fiends, Legits, and New Jacks," *Criminology* 32, 2: 197-219.
- Hindelang, M.J., T. Hirschi, and J. Weis. 1981. *Measuring Delinquency*. Beverly Hills, CA: Sage.
- Juhn, Chinui, Kevin Murphy, and Robert Topel. 1991. "Why Has the Natural Rate of Unemployment Increased Over Time?" *Brookings Papers on Economic Activity* 2: 75-142.
- Land, K.C., Patricia McCall, and Lawrence Cohen. 1990. "Structural Covariates of Homicide Rates: Are There Any Invariances Across Time and Social Space?" *American Journal of Sociology* 95, 4 (January): 922-63.
- Langan, Patrick. 1991. "America's Soaring Prison Population," *Science* 251 (March 29): 1568-73.
- Lattimore, P., A. Witte, and J. Baker. 1989. "Experimental Assessment of the Effect of Vocational Training of Youthful Property Offenders." NBER Working Paper 2952, May.

- Lee, David Sang-Yoon. 1993. "An Empirical Investigation of the Economic Incentives for Criminal Behavior." B.A. Thesis in Economics, Harvard University, March.
- Mendel, Richard A. 1995. "Prevention or Port? A Hard-Headed Look at Youth-Oriented Anti-Crime Programs." Washington, DC: American Youth Policy Forum.
- Mishel, Lawrence, and Aaron Bernstein. 1994. *The State of Working America 1994-95*. Armonk, NY: Economic Policy Institute., M.E. Sharpe.
- Needels, Karen. 1993. "Go Directly to Jail and Do Not Collect? A Long-Term Study of Recidivism and Employment Patterns Among Prison Releases." Department of Economics, Princeton University, November.
- Phillips, L., and H. Votey. 1990. "Demographics and Trends in Crime: A Failed Hypothesis." University of California, Santa Barbara, May.
- Piehl, A., and J. DiIulio. 1995. "Does Prison Pay? Revisited." *The Brookings Review* (Winter): 21-25.
- Reuter, P., R. MacCoun, and P. Murphy. 1990. *Money From Crime: A Study of the Economics of Drug Dealing in Washington, DC*. Santa Monica: RAND.
- Stephens, David W., and John R. Krebs. 1986. *Foraging Theory*. Princeton, NJ: Princeton University Press.
- Topel, Robert. 1993. "What Have We Learned from Empirical Studies of Unemployment and Turnover?" *American Economic Association, Papers and Proceedings*, May, pp. 110-115.
- U.S. Bureau of the Census *Statistical Abstract*, various years.
- U.S. Department of Justice, Bureau of Justice Statistics:
Crime in the United States, various editions.
Sourcebook of Criminal Justice Statistics, various editions.
Recidivism of Prisoners Released in 1983.
Profile of State Prison Inmates, 1986.
- Viscusi, K. 1986. "Market Incentives for Criminal Behavior." In *The Black Youth Employment Crisis*, R. Freeman and H. Holzer, eds. Chicago: University of Chicago Press.
- Wilson, J., and Allan Abrahamse. 1992. "Does Crime Pay?" *Justice Quarterly* 9, 3 (September): 359-77.

Explaining Tax Compliance

James Alm

University of Colorado at Boulder

“Always try to be honest. It will gratify some, and astonish all the rest.”

Mark Twain

The Puzzle of Tax Compliance

Tax evasion is an economic crime, perhaps the most common economic crime, and it appears to be a large and growing problem, both in the United States and elsewhere. Despite obvious difficulties in measurement, the Internal Revenue Service (1990) estimates for the United States that the tax gap, or the amount of unpaid federal income taxes, was between \$83 billion and \$93 billion in 1987 and had grown at an average annual rate of over 10 percent over the last two decades; more recent estimates by the Internal Revenue Service (1993) put the tax gap in 1991 at over \$111 billion. Similar work for other countries suggests that tax evasion is an even more severe problem elsewhere (Tanzi 1982; Feige 1989). Such underreporting has a variety of harmful effects: it reduces the tax revenues of the government, it affects public provision of goods and services, it creates misallocations in resource use, it alters the distribution of income in unpredictable ways, it increases feelings of unfair treatment by government, and it generates disrespect for the law.

Still the puzzle of tax compliance (at least for economists) is not so much “Why is there so *much* cheating?” Instead, the real puzzle is “Why is there so *little* cheating?” This may seem surprising. However, most people pay most of their taxes most of the time, even though the chances of detection are quite small and the penalties on evasion are also extremely light. For example, in the United States in recent years less than 1 percent of all individual income tax returns are audited by the Internal Revenue Service, and the penalty on even fraudulent evasion is only 75 percent of unpaid taxes. Most economic models of taxpayer behavior conclude that there should be much more tax evasion than is

actually observed. The puzzle of tax compliance—and the challenge facing people working in this area—is to explain why people pay taxes.

The study of tax compliance has grown enormously in the last twenty years. Still, there is much about compliance that is not understood. My purpose in this lecture is to discuss some work my colleagues and I have done that attempts to explain the factors underlying tax compliance behavior. My basic conclusion is that the explanation of compliance requires us to recognize the myriad range of factors that motivate individual behavior, factors that go much beyond the standard economics-of-crime approach that economists typically invoke, to include theories of behavior suggested by psychologists, sociologists, and other social scientists. Admittedly, such a broadening in the scope of analysis is a difficult one for economists to make. However, unless this approach is taken—and experimental economics can help here—we will not be able to explain the levels of compliance actually observed or to devise policies to increase compliance.

The next section discusses the major elements that I believe must be included in any theory of tax compliance. It is followed by a section explaining one of the methods that many people have found useful in explaining compliance: experimental economics. The section after that presents some of the results of experimental studies performed by my colleagues and me. The last section discusses some conclusions and observations.

Theories of Tax Compliance

This section outlines the major elements that, I believe, enter the tax compliance decision of individuals. These factors include the standard elements of audit and penalty rates. However, they also include several factors suggested by alternative theories of behavior under uncertainty. They are summarized in table 1.

Table 1. Main Factors in Tax Compliance Behavior

- Detection and Punishment
- Overweighting of Low Probabilities
- Burden of Taxation
- Government Services
- Social Norms

Detection and Punishment

The dominant economics approach to the analysis of tax compliance follows the economics-of-crime methodology pioneered by Becker (1968) and first applied to tax compliance by Allingham and Sandmo (1972). In its simplest form, this approach assumes that an individual receives a fixed amount of income I , and must choose how much of this income to declare to the tax authorities and how much to underreport. The individual pays taxes at rate t on every dollar D of income that is declared, but pays no taxes on underreported income. However, the individual may be audited with some fixed probability p ; if audited, then all underreported income is discovered, and the individual must pay a penalty at rate f on each dollar of deficient taxes, where f includes the unpaid taxes. If underreporting is detected, the individual's disposable income equals

$$I_C = I - tD - f(I - D), \quad (1)$$

while if underreporting is not detected income is

$$I_N = I - tD. \quad (2)$$

Expected utility theory then suggests that the individual will choose declared income to maximize the expected utility $EU(I)$ of the evasion gamble

$$EU(I) = pU(I_C) + (1 - p)U(I_N), \quad (3)$$

where E is the expectation operator and utility $U(I)$ is assumed to be a function only of income.

It is straightforward to show that an increase in the probability of detection p and the penalty rate f unambiguously increase declared income D . Surprisingly, an increase in the tax rate t generally has an ambiguous effect on declared income; however, under standard assumptions about an individual's attitudes toward risk, a higher tax rate actually increases declared income.

This basic model has been extended in a variety of dimensions (Cowell 1990). In particular, the assumption that the probability of

detection is fixed for an individual (a random audit strategy) can be relaxed by allowing the audit agency to use information from the taxpayers' returns in determining whom to select for audit and by examining the interaction of the taxpayers and the government collection agency in a game theory setting. Such "endogenous audit selection rules" are a central part of the enforcement strategies of many countries. Nevertheless, the essential features of this economics-of-crime model have largely remained the same.

Now this approach gives the sensible result that compliance depends upon enforcement. It is essential to recognize, however, that this approach also concludes that an individual pays taxes because—and *only* because—of the fear of detection and punishment. Again, this is a plausible and productive insight, with the obvious policy implication that the government can encourage greater tax compliance by increasing the audit and penalty rates. However, I know of no serious student of tax compliance who believes that tax compliance can be explained entirely by the level of enforcement. As noted earlier, the levels of audit and penalty rates are set at such low levels in most all countries that a purely economic analysis of the tax evasion gamble implies that most individuals would evade if they are "rational" because it is unlikely that cheaters will be caught and penalized. However, such behavior is simply not observed, even in the most evasion-ridden economy.

Put differently, the standard economics approach to the analysis of tax compliance has some serious flaws as an explanation for observed compliance choices of individuals because it concludes that individuals should pay far less in taxes than they in fact do. It is clear that the individual compliance decision either must be affected by other factors not mentioned by expected utility theory or must be affected in ways not captured by the theory.

Overweighting of Low Probabilities

Another factor is suggested by recent theoretical work by Kahneman and Tversky (1979) and others, who argue that many individuals can overweight low probabilities, such as those relevant for tax compliance. Suppose, for example, that the true probability of an audit is 1 percent. In making their decision, however, many individuals, even when fully informed, will systematically behave as if they think the

probability exceeds 1 percent, at least when their behavior is viewed from an expected utility perspective. Overweighting of low probabilities may therefore provide an additional explanation for tax compliance. If taxpayers give more weight to the probability of an audit than they ought to relative to an expected utility model, then compliance will be greater than the level suggested by the standard economics approach.

In fact, there is overwhelming empirical and experimental evidence that many (though not all) individuals overweight low-probability events: in their purchase of flood and earthquake insurance, in their willingness to pay to avoid exposure to hazardous substances, in their purchase of lottery tickets, and so on (Machina 1987; Davis and Holt 1993). It can be shown that such overweighting leads to greater compliance than predicted by expected utility theory.

The Burden of Taxation

A standard explanation for the rise of the underground economy is the general increase in the burden of taxation that has characterized most modern economies (Tanzi 1982; Feige 1989). In the face of higher burdens on reported income, it is argued that individuals will respond by reporting less income.

Now it must be remembered that the economics-of-crime approach does in fact conclude that tax compliance is affected by the level of tax rates. However, the theoretical and empirical strands of literature give different answers about the response of declared income to tax rates. As noted earlier, the typical *theoretical* result is that compliance actually rises as the tax rate rises. On the other hand, most *empirical* work finds that a higher tax rate discourages tax compliance (Clotfelter 1983; Alm, Bahl, and Murray 1990). Although there is no doubt that compliance depends in some way on the burden of taxation, the precise way in which compliance responds to changes in the burden of taxation needs further analysis.

Government Services

Another factor in the compliance decision is the use of the taxes. However, the role of government expenditures in the tax compliance decision has until recently been neglected. As emphasized by Cowell

and Gordon (1988, p. 305), “this seems a curious oversight, since while the government taketh away, it also giveth back, and the latter activity surely exerts some influence on evasion.” The compliance decision of an individual therefore seems likely to depend in some way on the individual’s receipt of government expenditures.

There is some work that is relevant here. A standard argument in public economics is that voluntary private provision of public goods will be inefficiently low because each individual will have an incentive to “free ride” on the private purchase of others (Samuelson 1954). In the context of tax compliance, this result suggests that most people will cheat. However, casual observation suggests that the likelihood of complete free riding is greatly overstated because instances of voluntary provision of public goods are widespread. Perhaps based upon these examples, there is now a large and growing literature (Axelrod 1984; Bagnoli and McKee 1991) that argues that voluntary provision of public goods may not always play as a “prisoners’ dilemma” game; that is, in many instances, individuals will in fact voluntarily contribute to a public good, or pay their taxes. This generally occurs when provision is both repeated and interdependent. In such a setting, *one* individual’s decision to contribute—or to comply—depends upon his or her perception of what *others* will contribute, both now and in the future. If the individual believes that his or her contribution is in some sense essential (or “pivotal”) to the provision of the public good, then free riding is no longer the unique dominant strategy for the individual. Instead, cooperation may become optimal.

This approach suggests that individuals may voluntarily pay taxes in part because they recognize that payment is necessary to get others to contribute and so to get government services that are valued.

Social Norms

A last factor is “social norms,” which I believe may well be the most important factor. It is clearly difficult to be very precise on the exact meaning of social norms (Roth, Scholz, and Witte 1989). However, there is overwhelming evidence that many countries with roughly the same fiscal system also have far different compliance experiences. The only possible explanation that I can suggest is that these countries have different notions of what is socially acceptable behavior; that is, they exhibit different social norms.

To illustrate, there is much survey evidence (Westat 1980; Yankelovich, Skelly, and White 1984) that suggests that compliance is strongly affected by the strength of and commitment to the social norm of compliance. These surveys conclude that:

- those who comply tend to view tax evasion as “immoral”
- compliance is higher if a “moral appeal” to taxpayers is made
- the low social standing of tax evaders may be a more effective deterrent than formal sanctions
- individuals with tax evaders as friends are more likely to be evaders themselves
- compliance decreases with perceptions of unfair treatment
- evasion is associated with feelings of distrust and alienation
- compliance is greater in societies with a strong sense of social cohesion.

Further, anecdotal evidence suggests that some people won’t pay their taxes if they dislike the way their taxes are spent, if they feel they have no say in the decision process, or if they feel that government is unresponsive to their wishes. Some quotes from taxpayers illustrate these feelings (Yankelovich, Skelly, and White 1984):

- “I wouldn’t mind it so much if I could designate where my tax dollars went to. I resent having to find out why frogs in South America croak and things like that.”
- “When we pay taxes, we like to know what it’s going for.”
- “Allow people to earmark a portion of their tax payments. Give them choices.”

It seems clear to me that such sentiments play an important, perhaps a dominant, role in tax compliance.

Social norms can be affected by a variety of government institutions and policies. There is much behavioral science evidence that implies that greater individual participation in the decision process will foster an increased level of compliance, in part because participation implies some commitment to the institution and such commitment in turn

requires behavior that is consistent with words and actions. This notion implies that one dimension by which social norms can be affected is via individual participation in the decision process, say, by voting. Compliance seems likely to be higher when the use of tax revenues is decided by majority rule than when the (same) use is imposed on the group; further, knowing the outcome of the vote reveals information to each taxpayer about the level of group support for the collective decision, and this information may be useful to individuals in projecting the group behavior. Government decisions that are imposed are unlikely to generate such feelings of participation or to provide such information. Consequently, if taxpayers feel that they have voice in the way their taxes will be spent, then they are likely to feel more inclined to pay their taxes.

Another dimension by which social norms may be affected is the level of popular support for the government program. Widespread support tends to legitimize the public sector, and so imposes some social norm to pay taxes. This support may be obviously revealed through the voting process. However, the level of support seems likely to affect compliance even when the choice of the public good is imposed on members of the group. Consequently, it seems likely that there will be more tax compliance when the public good imposed on a community is popular, even if individuals are unable to articulate directly their support via voting.

Still another dimension by which social norms can be changed is the community commitment to enforcing the tax laws. If the perception becomes widespread that the government is not willing to detect and penalize evaders, then such a perception legitimizes tax evasion. The rejection of sanctions sends a signal to each individual that others do not wish to enforce the tax laws and that tax evasion is in some sense socially acceptable. The social norm of compliance disappears. Such an outcome is common in many countries, such as the Philippines and Italy where it seems to be accepted that tax evasion is the norm.

Summary

These factors are clearly only some of the elements in the individual compliance decision, and there are numerous other factors that affect this decision: uncertainty about the fiscal system, the use of paid pre-

parers and advisors, the withholding of taxes, rewards for honest declarations, the potential for tax amnesties, the joint choice of tax avoidance and tax evasion, and so on. Nevertheless, I believe that these factors together play dominant roles in tax compliance. Methods to investigate the importance of their roles are discussed next.

Experimental Methods and Experimental Design

There are essentially two broad methodological alternatives to the use of experimental methods in the study of tax compliance: theoretical and empirical methods. Each has generated insights, but each is also subject to some serious problems. Before discussing the experimental approach to compliance, it is useful to begin by outlining the strengths and, more important, the limitations of theoretical and empirical analyses as a way of justifying the use of experimental methods as an additional tool in the study of tax compliance.

Theoretical Approaches

Virtually all theoretical work on tax compliance relies in some form on the expected utility model. This approach has generated many insights, especially regarding how an individual responds to greater enforcement activities and how government can optimally choose its enforcement strategy. However, this literature is in a sense too complex. It is only in the simpler models that clear-cut analytical results can be generated on the compliance impact of basic policy parameters. When more complex dimensions of individual behavior are introduced, the theoretical results generally become ambiguous. It is doubtful that theoretical analysis will yield more meaningful results in the future.

Paradoxically, the theoretical models of individual choice are also too simple. There are numerous factors that affect the reporting decisions of individuals, but theoretical models are capable of including only a few.

Most important, and as emphasized above, the limited ability to incorporate many relevant factors or to incorporate them in a meaningful way has meant that theories based upon expected utility theory are unable to explain the *level* of tax reporting. In particular, these models

generally imply that rational individuals should pay far less in taxes than they actually do. This is not a mere quibble. It goes to the heart of the standard approach to explaining compliance.

Empirical Literature

The obvious difficulty in applied work is the absence of reliable information on individual reporting behavior. This information is hard to come by, either for the United States or for other countries: it is difficult to measure something that by its very nature people want to conceal. This difficulty has not stopped researchers. However, there are obvious problems with the data that make much of this empirical work somewhat suspect.

For example, most empirical work for the United States has utilized data provided by the IRS through its Taxpayer Compliance Measurement Program (TCMP), which contains a detailed line-by-line audit of a stratified random sample of roughly 50,000 individual tax returns conducted on a three-year cycle. These audits yield an IRS estimate of the taxpayer's "true" income so that a measure of individual tax evasion can be calculated. However, until recently most researchers have not had access to the individual data, and instead have been forced to use TCMP data aggregated to the three-digit zip code level, an aggregate measure likely to comprise disparate elements of underreporting that reflect very different motivational factors. TCMP data also have some well-recognized deficiencies: the audits do not detect all underreported income, nonfilers are not captured, honest errors are not identified, final audit adjustments are not included, and there are few noneconomic factors to which the data can be linked. The use of TCMP data for empirical estimation of the determinants of compliance behavior is therefore problematic. Data for other countries are even more flawed.

To avoid the problems with the TCMP data, some researchers have used aggregate measures of evasion, such as the gap between income reported on tax returns and income in the national income accounts. By necessity, these studies focus on the aggregate, not the individual, response. Other researchers have used surveys of taxpayers, in part to assess factors such as perceptions of the probability of detection, the fairness of taxation, and the responsiveness of government in the

respondent's reporting decision. Unfortunately, these surveys are also subject to a number of methodological problems. Individuals may not remember their reporting decisions, they may not respond at all, or they may not respond truthfully. Surveys are also unable to control for many relevant determinants of compliance, and, given their response rates, surveys may not be representative of the population at large. Finally, they cannot determine the direction of causality between compliance and its determinants; that is, statements regarding the unfairness of a tax may result from a rationalization of noncompliance rather than be the cause of noncompliance.

Experimental Economics

Difficulties with the existing theoretical and empirical literatures have led to the use of experimental economics in compliance research, not so much as the only approach but as an additional approach. The use of laboratory experiments in economics began in the early 1960s with the work of Smith (1962, 1964) on resource allocation under alternative forms of market organization. Growth in its applications came with the establishment of a well-defined framework for experimental work by Smith (1976, 1982) and Wilde (1980), and it is now widely accepted as a methodological approach in the analysis of theory and policy. Davis and Holt (1993) survey much of the experimental literature.

As discussed by Alm (1991), laboratory experiments seem particularly well suited for the study of some aspects of the taxpayer reporting decision. Unlike theoretical work, experiments are not as constrained by the same degree of simplification required in analytical studies of reporting, which allows the impact of numerous factors not amenable to theoretical work to be examined. Unlike empirical work, experiments generate data under different settings in which there is control over extraneous influences. As discussed below, there are some obvious limitations of experimental methods. However, given the weaknesses of other methodologies, there are compelling reasons for the use of experiments.

Creating a microeconomic system: induced value theory

Experimental economics involves the creation of a real microeconomic system in the laboratory, one that parallels the naturally occur-

ring world that is the subject of investigation. The essence of such a system is control over the environment, the institutions, and the preferences that subjects face. Of these, control over preferences is particularly crucial. As stated by Smith (1976, p. 275), "such control can be achieved by using a reward structure to induce prescribed monetary value on actions."

Smith (1982) identifies several (sufficient) conditions that must be satisfied for control over preferences to be established: (1) "nonsatiation"—subjects must prefer more to less; (2) "saliency"—the rewards received by subjects must be related to their decisions, so that subjects recognize that their actions affect their outcomes; (3) "reward dominance"—rewards must be large enough to offset any subjective costs or benefits that subjects place on participation in the experiment, which requires the payment to subjects of an amount comparable to what they could earn outside the laboratory; and (4) "privacy"—each subject must know only his or her own payoffs so that they do not receive any subjective value from the payoffs of other subjects.

Several procedures should also be followed in experimental economics. For example, the experiment should be administered in a uniform and consistent manner to allow replicability. The experiment should not be excessively long or complicated, since subjects may become bored or confused. Subjects must believe that the procedures described to them are the procedures actually followed. The instructions provided to subjects should be understandable, should avoid the use of examples that lead subjects to anchor on certain choices that are the subject of the experiment, and should be phrased in "neutral" rather than "loaded" terms, to mask the context of the experiment and to avoid direct reference to the real-world phenomena under investigation. Neutrality increases the experimenter's control over subject preferences and avoids leading subjects to invoke different "mental scripts," which may enable them to fill in (potentially) missing information in the instructions but which also may unpredictably influence their choices. It is sometimes claimed that the use of neutral instructions limits the ability to generalize from the experimental to the naturally occurring setting. In fact, however, it is not possible to generalize beyond the laboratory unless one uses neutral instructions, since the experimenter cannot control (or induce) the values that subjects associate with loaded terms.

Experimental work on taxpayer compliance

The basic design of most compliance experiments is similar, and is summarized in table 2. Human subjects in a controlled laboratory are told that they should feel free to make as much income as possible. At the beginning of each round of the experiment, each subject is given income and must decide how much income to report. Taxes are paid at some rate on all reported, but not on underreported, income. However, underreporting is discovered with some probability, and the subject must then pay a fine on unpaid taxes. This process is repeated for a given number of rounds. At the completion of the experiment, each subject is paid an amount (the accumulated earnings) that depends on his or her performance during the experiment. Into this microeconomic system, various policy changes can be introduced: changes in audit probabilities or audit rules, in penalty rates, in tax rates, in public good provision, and in institutions that affect social norms. Some results from experiments run by my colleagues and me are discussed in the next section.

Table 2. Basic Design of Tax Compliance Experiments

- Student subjects are used.
- Subjects are fully informed.
- Subjects are organized into groups that stay together throughout the experiment.
- The known number of rounds is predetermined but unannounced.
- Subjects receive income in each round.
- Subjects must choose in each round how much income to declare.
- Declared income is taxed; in those experiments with a public good, all taxes finance the public good.
- Undeclared income is not taxed, but subjects face some chance of audit and penalty.
- Subjects are paid their accumulated earnings at the end of the experiment, generally \$10 to \$30 depending upon their performance.
- Parameters are set at their “real-world” levels.
- Experiments are fully computerized, and last one hour.
- Variations include changes in audit rates and rules, in fine rates, in tax rates, in public good provision, and in subject participation via voting on the fiscal system.

Limitations of experimental economics

There are sound reasons for caution in interpreting and generalizing experimental results. Some early compliance experiments did not follow now widely accepted procedures of the experimental paradigm, such as the use of repeated experiments and neutral instructions. Much early work also lacked realism because values of the various policy parameters did not approximate real-world values.

Although more recent experimental research has generally addressed these problems, some concerns remain, some of which are more real than others. A common criticism of experimental economics is that the student subjects typically used may not be representative of taxpayers. However, there is now much evidence that the experimental responses of students are no different from the responses of other subject pools (Plott 1987); there is also no reason to believe that the cognitive processes of students are different from “real” people. Another common criticism is that it is not possible to control for many relevant factors in the laboratory. However, if one cannot control for such factors in the laboratory where the experimenter establishes the institutions, the rules, and the reward structure, then one cannot hope to control for these factors in the “naturally occurring world.”

Of more legitimate concern, the results may well be sensitive to the specific experimental design, so that replication is crucial. It is also possible that subjects may modify their behavior simply because they know that they are participating in an experiment. Most important, there is a certain artificiality in a laboratory setting. A decision to report \$2 in an experiment is clearly different from a decision to report actual income on an annual tax return, even if the laboratory incentives are salient. In particular, the laboratory setting cannot capture a catastrophic loss such as jail, and it cannot capture the social stigma that some surveys suggest is an important factor in taxpayer reporting. In short, one must use the results from laboratory experiments with some care. However, such use depends largely upon the purpose of the experiment. According to Roth (1987), experiments can be classified into three broad categories that depend upon the dialogue in which they are meant to participate. “Speaking to Theorists” includes those experiments designed to test well-articulated theories. “Searching for Facts” involves experiments that examine the effects of variable about

which existing theory has little to say. “Whispering in the Ears of Princes” identifies those experiments motivated by specific policy issues. To date, most experiments on taxpayer reporting fall into the first two categories. Although this now seems to be changing somewhat, it is likely to be some time before a serious and ongoing dialogue with the princes of the IRS is established.

Experimental Evidence on Tax Compliance

There are a number of excellent experimental analyses of tax compliance, such as Beck, Davis, and Jung (1991), Collins and Plumlee (1991), and Webley et al. (1991). However, I will limit my discussion to some experimental results from papers by my colleagues and myself. It is somewhat difficult to compare these results because the specific experimental design often differs for the papers. For example, the number of periods subject to audit varies over the studies, the penalty rates also varies, and tax payments are sometimes used to provide a public good. Nevertheless, in their entirety these papers give, I believe, a good indication of the importance of the various factors discussed above in the tax compliance behavior of individuals. The experimental results are summarized in table 3.

The Impact of Detection and Punishment

Audit rates and rules

Several papers have varied the probability of detection p , in which the probability is assumed to be random and independent of any taxpayer decisions (Alm, McClelland, and Schulze 1992; Alm, Jackson, and McKee 1992b; Alm, Cronshaw, and McKee 1993). In general, the results for these “random audit rules” clearly indicate that compliance increases with a greater audit rate, with an estimated declared income-audit rate elasticity of 0.17. However, the increase in compliance appears to be small and nonlinear. This suggests the important result that there are limits to how much government can increase compliance by increasing the probability of detection.

Table 3. Experimental Results: Average Compliance Rate**Audit rates and rules**

<u>$p = 0$</u>	<u>$p = .02$</u>	<u>$p = .10$</u>	<u>Paper-design</u>
20.0%	50.3%	67.5%	A: Public good/single period audits/ $f = 15/t = .40$
<u>$p = .02$</u>	<u>$p = .04$</u>	<u>$p = .06$</u>	<u>Paper-design</u>
31.7%	33.2%	37.6%	B: No public good/back audits/ $f = 2/t = .30$
<u>$p = .05$</u>	<u>$p = .30$</u>	<u>$p = .50$</u>	<u>Paper-design</u>
27.7%	34.3%	49.2%	C: No public good/single period audits/ $f = 2/t = .30$
<u>CQ</u>	<u>CFA</u>	<u>CBA</u>	<u>Paper-design</u>
80.8%	51.6%	55.9%	C: No public good/ $f = 2/t = .30$

Penalty rates

<u>$f = 1$</u>	<u>$f = 2$</u>	<u>$f = 3$</u>	<u>Paper-design</u>
32.1%	33.2%	36.5%	B: No public good/back audits/ $p = .04/t = .30$

Overweighting

There is substantially greater compliance at low probabilities than is predicted by risk-neutral behavior in papers A, B, C, D, E, and F.

Tax rates

<u>$t = .10$</u>	<u>$t = .30$</u>	<u>$t = .50$</u>	<u>Paper-design</u>
37.6%	33.2%	20.0%	D: No public good/back audits/ $f = 2/p = .04$

Government services

<u>$m = 0$</u>	<u>$m = 2$</u>	<u>$m = 6$</u>	<u>Paper-design</u>	
43.5%	53.7%	59.2%	A: Single period audits/ $f = 15/p = .02/t = .40$	
<u>$m = 0$</u>	<u>$m = 2$</u>	<u>Lottery</u>	<u>Reward</u>	<u>Paper-design</u>
33.2%	37.4%	51.3%	44.8%	B: Back audits/ $f = 2/p = .04/t = .03$

Social norms

<u>MRD</u>	<u>MRS</u>	<u>IFC</u>	<u>INC</u>	<u>NPG</u>	<u>Paper-design</u>
45.2	41.1	33.7	8.1	33.7	E/back audits/ $f = 2/p = .04/t = .30$

Voter rejection of greater enforcement via majority rule is followed by near-zero compliance, while voter acceptance of greater enforcement via majority rule with “cheap talk” is followed by near-complete compliance, in paper F.

Paper A:	Alm, McClelland, and Schulze (1992).
Paper B:	Alm, Jackson, and McKee (1992b).
Paper C:	Alm, Cronshaw, and McKee (1993).
Paper D:	Alm, Jackson, and McKee (1992a).
Paper E:	Alm, Jackson, and McKee (1993).
Paper F:	Alm, McClelland, and Schulze (1993).

Another paper has varied the nature of the audit rule (Alm, Cronshaw, and McKee 1993). A central and obvious feature of the compliance process in most countries is that the government tax agency uses information from the tax returns to determine strategically whom to audit. Such a policy means that the probability of audit is not fixed and random but instead is variable and endogenous, dependent in part on the behavior of the taxpayer (and the tax agency).

Different “endogenous audit rules” are examined in experiments in this paper. One rule assumes that an audited individual found to be noncompliant in the current period will be audited with certainty for a number of future periods (“Conditional Future Audit” rule, or CFA). Another rule requires that an audited individual will face some back audits if found to be noncompliant in the current period (“Conditional Back Audit” rule, or CBA). A third rule imposes a cutoff (“Cutoff” rule, or CO); a taxpayer who reports less than some cutoff level of income will be audited with certainty.

Experimental results indicate that endogenous audit rules are able to generate compliance significantly greater than random audit rules, even when the random audit rate is 30 or 50 percent. The cutoff rule CO is the most effective in increasing compliance, although it requires a large number of audits. Making back audits conditional on current declarations (or CBA) is also able to increase compliance significantly, and the audit rate is far lower than the cutoff rule. A conditional future audit rule CFA appears to be the least effective of the endogenous rules, although compliance still exceeds that under all random audit rules.

These results suggest that:

In the relevant range of audit rate changes, compliance increases marginally and nonlinearly with increases in the probability of detection. Also, compliance is significantly greater with endogenous audit selection rules than with random audit rules.

Work by other experimentalists generally gives similar results.

Penalty rates

Variations in the penalty rate f are also examined in Alm, Jackson, and McKee (1992b). Experimental results indicate that individuals pay

slightly more in taxes when the penalty on detected evasion increases. Therefore:

In the relevant range of penalty rate changes, compliance increases marginally with increases in the fine rate on unpaid taxes.

However, the impact is quite small, with a reported income-fine rate elasticity of only 0.04.

The Impact of Overweighting

Several experiments across all studies are designed explicitly to test for the presence of overweighting of low probabilities. In these experiments the parameters—especially the probability of detection—are sometimes chosen such that the optimal strategy for a risk-neutral individual is to report zero income. Nevertheless, the experimental results clearly indicate that there is far more compliance than is predicted by expected utility theory, a result consistent with overweighting.

These results suggest that:

Many individuals substantially overweight low probabilities.

Although greater compliance may well be implied by other factors as well, these results are similar to much other experimental work on behavior at low probability events.

The Impact of Taxation

Another set of experiments varies the tax rate t on declared income, from 10 to 30 to 50 percent, and the level of compliance falls significantly with tax rate increases (Alm, Jackson, and McKee 1992a). The declared income-tax rate elasticity is estimated to equal -0.52. Therefore:

Compliance decreases with increases in the tax rate.

These results are consistent with most empirical evidence, even though much theoretical work concludes that compliance should rise with greater tax rates.

The Impact of Government Services

In several papers a public good is present whose magnitude depends upon the tax payments of all subjects (Alm, McClelland, and Schulze 1992; Alm, Jackson, and McKee 1992b). These tax payments are summed across the subjects, this sum is increased by some multiple m to reflect the (potential) consumers' surplus that individuals derive from government provision of a public good, and the resulting amount is then divided equally among the subjects. Note that a multiplier greater than 1 implies that individuals as a group receive more than they pay in taxes.

All experiments clearly indicate that compliance is greater in the presence of the public good than in its absence; also, compliance increases in a nonlinear way with the multiplier m .

Some additional experiments also vary the nature of the positive inducement given to taxpayers: by making audited and fully compliant taxpayers eligible for a "Lottery" whose expected value equals the average subject per round income, or by giving audited and fully compliant taxpayers an immediate "Reward" of comparable value. In both cases the presence of a positive inducement leads to greater compliance.

These results suggest that:

Compliance increases when individuals receive something for their tax payments.

There is some evidence that tax agencies around the world are starting to pursue such a "kinder, gentler" strategy.

The Impact of Social Norms

Manipulating social norms is perhaps the most difficult task facing experimentalists, and there are few studies in which such control has been achieved. Nevertheless, some of my work has, I believe for the first time, been able to induce predictable changes in social norms, with resulting impacts on tax compliance.

One study examines the effects on compliance of the uses of tax revenues and the decision process by which these uses are chosen (Alm, Jackson, and McKee 1993). In some experiments subjects must choose between using their tax payments on one of two alternative public

goods; the level of support for the public good alternatives also varies between strong and weak support for the public good. These experiments are denoted MRD (for "Majority Rule over Diverse" choices, in which the level of support for the preferred outcome is known to be quite strong) and MRS (for "Majority Rule over Similar" choices, in which the level of support for the preferred outcome is known to be quite weak). In other experiments subjects are not allowed to vote on public good provision, and a public good of variable popularity is instead imposed on the group. These experiments are denoted IFC (for "Imposed Favored Choice," in which the imposed public good is known to be popular) and INC (for "Imposed Nonfavored Choice," in which the imposed public good is known to be unpopular). Finally, in one experiment there is no public good (or NPG, for "No Public Good"). The experiments are structured so that the same use of tax revenues occurs in all experiments except INC and NPG.

These experimental results indicate that compliance is significantly greater when subjects choose via voting the use of their taxes than when the identical use is imposed upon them; compare MRD versus IFC and MRS versus IFC. Further, compliance is somewhat greater when the vote is decisive than when the vote is close (MRD versus MRS). Finally, compliance is significantly lowered by the imposition of an unpopular program (IFC versus INC); in fact, compliance is lower with an imposed, unpopular public good than in the complete absence of any public good (INC versus NPG). These results clearly show that government can change the social norm of compliance by ensuring that individuals have a say in the decision process and by spending taxes in ways consistent with citizen preferences. Such policies have seldom been thought to be part of general strategy for tax compliance, but nonetheless they appear to be effective tools.

A second study allows subjects to vote via majority rule on the enforcement regime that they face, such as the tax, audit, and penalty rate (Alm, McClelland, and Schulze 1993). These results are still somewhat tentative. However, I believe that this voting allows the subjects to alter the social norm of compliance. In particular, in all cases (except one, as discussed in a moment) subjects vote against greater enforcement such as higher audit or penalty rates, even when it is individually and socially optimal to increase enforcement. Following these votes tax compliance falls virtually to zero, even though compliance

prior to the vote is quite substantial. Rejection by the group of greater enforcement sends a clear signal to each individual that tax evasion will be tolerated and accepted; that is, rejection changes the social norm of compliance, and individuals respond accordingly.

Social norms can also be affected by communication among members of the group, however. In an additional experiment in this paper, subjects are allowed to talk with one another before the vote on enforcement is taken (or “cheap talk”). In these discussions subjects quickly discover that it is in their interests to impose strict sanctions on free-riders, in order to increase their share of the public good. They then vote overwhelmingly in favor of greater enforcement, and tax compliance following the vote approaches 100 percent. Again, I believe that the cheap talk in combination with the vote allows individuals to change the social norm, in this case to demonstrate that evasion will not be accepted.

These conclusions are striking:

Social norms can be changed by fiscal institutions. Compliance is increased when individuals participate via voting in the process by which the use of tax revenues is decided, when the outcome of the vote reveals widespread support for the program, and when the use of tax revenues is popular even if imposed. Also, compliance is decreased when there is a social expression via voting of a willingness to tolerate tax evasion, and compliance is increased when there is a social expression via voting of an unwillingness to tolerate tax evasion.

There is also some experimental evidence that tax amnesties may change social norms (Alm, McKee, and Beck 1990).

These conclusions suggest a variety of alternative government policies toward tax compliance that are only now beginning to be used.

Summary and Conclusions

People exhibit much diversity in their behavior, and they are motivated by a variety of factors. There are individuals who always cheat and those who always comply, some who maximize the expected utility of the tax evasion gamble, others who seem to overweight low

probabilities, individuals who respond in different ways to changes in their tax burden, some who are at times cooperative and at other times free-riders, and many who are guided by social norms.

These findings in total suggest that a government compliance strategy based only on detection and punishment may well be a reasonable *starting* point but not a good *ending* point. Instead, what is needed is a multifaceted approach that emphasizes enforcement, but that also emphasizes such things as positive rewards from greater tax compliance, the wise use of taxpayer dollars, and individual participation in the decision process. What is also needed, however, is a theory of tax compliance that incorporates the remarkable diversity in individual behavior exhibited by these experimental analyses of taxpayer compliance. Whether any such theory can be developed that explains the behavior of all individuals at all times, or even one that explains the actions of the same person at all times, is hard to determine. However, until this effort is made, I think it unlikely that we will come much closer to explaining the puzzle of tax compliance.

References

- Allingham, Michael G., and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics* 1, 4: 323-338.
- Alm, James. 1991. "A Perspective on the Experimental Analysis of Taxpayer Reporting," *Accounting Review* 66, 3: 577-593.
- Alm, James, Roy Bahl, and Matthew N. Murray. 1990. "Tax Structure and Tax Compliance," *Review of Economics and Statistics* 72, 4: 603-613.
- Alm, James, Mark Cronshaw, and Michael McKee. 1993. "Tax Compliance with Endogenous Audit Selection Rules," *Kyklos* 46, 1: 27-45.
- Alm, James, Betty Jackson, and Michael McKee. 1992a. "Estimating the Determinants of Taxpayer Compliance with Experimental Data," *National Tax Journal* 45, 1: 107-114.
- _____. 1992b. "Deterrence and Beyond: Toward a Kinder, Gentler IRS." In *Why People Pay Taxes: Tax Compliance and Enforcement*, Joel Slemrod, ed. Ann Arbor, MI: University of Michigan Press.
- _____. 1993. "Fiscal Exchange, Collective Decision Institutions, and Tax Compliance," *Journal of Economic Behavior and Organization* 22, 2 (December): 285-303.
- Alm, James, Gary H. McClelland, and William D. Schulze. 1992. "Why Do People Pay Taxes?" *Journal of Public Economics* 48, 1: 21-38.
- _____. 1995. "Changing the Social Norm of Tax Compliance by Voting." Working Paper, University of Colorado at Boulder.
- Alm, James, Michael McKee, and William Beck. 1990. "Amazing Grace: Tax Amnesties and Compliance," *National Tax Journal* 43, 1: 23-37.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Bagnoli, Mark, and Michael McKee. 1991. "Voluntary Contribution Games: Efficient Private Provision of Public Goods," *Economic Inquiry* 29, 2: 351-366.
- Beck, Paul J., Jon S. Davis, and Woo-Oh Jung. 1991. "Experimental Evidence on Taxpayer Reporting under Uncertainty," *Accounting Review* 66, 3: 535-558.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76, 2: 169-217.
- Clotfelter, Charles T. 1983. "Tax Evasion and Tax Rates: An Analysis of Individual Returns," *Review of Economics and Statistics* 65, 3: 363-373.
- Collins, Julie H., and R. David Plumlee. 1991. "The Taxpayer's Labor and Reporting Decision: The Effect of Audit Schemes," *Accounting Review* 66, 3: 559-576.

- Cowell, Frank A. 1990. *Cheating the Government: The Economics of Tax Evasion*. Cambridge, MA: MIT Press.
- Cowell, Frank A., and James P.F. Gordon. 1988. "Unwillingness to Pay: Tax Evasion and Public Good Provision," *Journal of Public Economics* 36, 3: 305-321.
- Davis, Douglas D., and Charles A. Holt. 1993. *Experimental Economics*. Princeton, NJ: Princeton University Press.
- Feige, Edgar L., ed. 1989. *The Underground Economies*. Cambridge, England: Cambridge University Press.
- Internal Revenue Service. 1990. *Income Tax Compliance Research: Net Tax Gap and Remittance Gap Estimates*, Supplement to Publication 7285, Publication 1415. Washington, DC: United States Department of the Treasury.
- _____. 1993. "Tax Gap Estimates for 1991," *Wall Street Journal*, May 12, 1993.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Uncertainty," *Econometrica* 47, 2: 263-291.
- Machina, Mark J. 1987. "Choice under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives* 1, 1: 121-154.
- Plott, Charles R. 1987. "Dimensions of Parallelism: Some Policy Implications of Experimental Methods." In *Laboratory Experimentation in Economics: Six Points of View*, Alvin E. Roth, ed. New York: Cambridge University Press.
- Roth, Alvin E. 1987. "Introduction and Overview." In Alvin E. Roth, ed., *Laboratory Experimentation in Economics: Six Points of View*, Alvin E. Roth, ed. New York: Cambridge University Press.
- Roth, Jeffrey A., John T. Scholz, and Ann Dryden Witte, eds. 1989. *Taxpayer Compliance, Vol. 1: An Agenda for Research*. Philadelphia: University of Pennsylvania Press.
- Samuelson, Paul. 1954. "The Pure Theory of Public Expenditure," *Review of Economics and Statistics* 36, 4: 387-390.
- Smith, Vernon L. 1962. "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy* 70, 2: 111-37.
- _____. 1964. "Effect of Market Organization on Competitive Equilibrium," *Quarterly Journal of Economics* 78, 2: 181-201.
- _____. 1976. "Experimental Economics: Induced Value Theory." *American Economic Review Papers and Proceedings* 66, 2: 274-79.
- _____. 1982. "Microeconomic Systems as an Experimental Science, *American Economic Review* 72, 4: 923-55.
- Tanzi, Vito, ed. 1982. *The Underground Economy in the United States and Abroad*. Lexington, MA: D.C. Heath.

- Webley, Paul, Henry Robben, Henk Elffers, and Dick Hessing. 1991. *Tax Evasion: An Experimental Approach*. Cambridge, England: Maison des Sciences de l'Homme and Cambridge University Press, European Monographs in Social Psychology.
- Westat, Inc. 1980. *Individual Tax Compliance Factors Study, Qualitative Research Results*. Rockville, MD, Westat.
- Wilde, Louis L. 1980. "On the Use of Laboratory Experiments in Economics." In *The Philosophy of Economics*, J. Pitt, ed. Dordrecht, The Netherlands: Reidel.
- Yankelovich, Skelly, and White, Inc. 1984. *Taxpayer Attitudes Survey: Final Report. Public Opinion Survey Prepared for the Public Affairs Division, Internal Revenue Service*. New York: Yankelovich, Skelly, and White.

Beating the System?

Ann Dryden Witte
Florida International University
and
Wellesley College

What is the Underground Economy?

The underground economy is a pervasive entity. It will always be with us. The formal or recorded economy is the new kid on the block. The formal economy only came into existence as we began to try to measure, broadly tax, and regulate economic activity. It is a creature by and large of the post-World War II period. We began to have broad-based taxes and benefits systems (during the 1930s and World War II), extensive regulations (mainly a post-World War II phenomenon) and national statistical systems (beginning in the late 1920s and picking up speed after World War II) mainly during the 1930s and 1940s.

It is the formal economy, the recorded, taxed, and regulated economy, that is the recent arrival—not the unrecorded or underground economy. This said, why are we concerned about this economy?

When and Why Did We Become Concerned with the Underground Economy?

There have been two periods of concern about the underground economy. The first period of concern occurred during and shortly after World War II and was largely confined to the United States and Canada. The large increases in taxes and regulations needed for the war effort drove much activity underground.

The second period of concern began in the late 1970s and spread from the United States to most of the developed world. By and large the underground economy has been of little concern in Third World countries (other than to the taxing authorities). Many developing countries (e.g., Brazil) even make government loans to enterprises that operate largely underground. As a source of employment and entrepre-

neurial activity, the underground economy is more often cultivated than criticized in Third World countries.

The most recent period of concern about the underground economy in the developed world began with a series of articles in the popular press during the late 1970s claiming that there was a large and rapidly growing underground economy. Politicians quickly picked up on this because it appeared to provide a “painless” method of dealing with the deficit (annual deficits in the United States averaging over \$50 billion a year began in 1975 and deficits of over \$100 billion began in 1982). The existence and claimed rapid growth of the underground economy provided support for deregulation and lowering of marginal tax rates—politically popular policies during the late 1970s and early 1980s.

My own saga with the underground economy illustrates the source and nature of concern. In 1978, I was sitting quietly in my office at the University of North Carolina when I received a call from the Chief Economist of the Joint Economic Committee. He requested that I write a monograph on the “underground economy” for the Committee’s Special Study on Economic Change. I replied that I would have to know what this economy was before I could respond. His description suggested a diverse and complicated set of activities that appeared to have little in common other than that they were somehow hidden from governmental agencies. The task seemed overwhelming and he wanted a first draft in six months. It was a hot topic and the monograph needed to be completed yesterday if not sooner. I asked for time to consider his request. I was strongly inclined to say no, since it sounded like a topic that would require a couple of years to deal with adequately.

As it happened, I was meeting with a mathematical economist, Carl Simon (Department of Economics, University of Michigan), for lunch to discuss a saddle point problem. Since I was late, I provided a detailed description of the phone call by way of apology. To my amazement, Carl indicated that he would love to undertake such a project with me. He solved my saddle point problem and I agreed to do the monograph jointly with him.

Substantial Interest by Other Agencies

At the same time that the Joint Economic Committee was commissioning our study, other congressional committees (e.g., the House

Committee on Ways and Means) asked for a study of nonfiling by the General Accounting Office (GAO) and pressured the Internal Revenue Service (IRS) into providing an estimate of unreported taxable income. The IRS did not wish to undertake estimates but was required to do so by Congress. These reports resulted in numerous congressional hearings during the early 1980s which provided valuable support for the deregulation and detaxation movements that accompanied the Reagan presidency.

A Dispassionate Look at the Underground Economy Today

It has now been over fifteen years since the rediscovery of the underground economy. There has been much research on the topic. The political passions that originally inspired renewed interest in the underground economy have cooled, at least somewhat. We are now at a point to step back and consider the nature and importance of this economy. In 1982, Carl Simon and I published a book, an expanded and updated version of our study for the Joint Economic Committee, on the underground economy (Simon and Witte 1982). The editor gave the book the title *Beating the System*. With the advantage of hind sight and an additional fifteen years of research, I would like to consider whether those in the underground economy are beating the system. I will discuss the impact of the underground economy on those who participate little, if at all, in its activities, as well as on those who are members of this underground.

What Kind of Activity Occurs in the Underground Economy?

Activities in the underground economy can usefully be divided into three categories: (1) "pure" tax evasion, (2) the irregular economy, and (3) illegal activities. The nature of these activities and the reasons for our concern are quite different for the three types of activity, and so I will discuss each type of activity separately before trying to reach general conclusions concerning the underground economy.

Pure Tax Evasion

Pure tax evasion, which is far from pure, occurs when individuals and businesses fail to fully report earnings from perfectly legal activities carried out in businesses that are properly registered and recorded in our national statistics. It also occurs when individuals and businesses overstate the deductions to which they are entitled.

Perhaps half of all Americans engage in "pure" tax evasion. However, most evasion is rather small scale. We are a nation of tax-evasion nibblers, although some individuals do take large gulps.

The level of evasion is very unequally spread throughout the population. The amount varies with *type of income*, *amount of income*, and with *attitudes toward laws and government*.

Type of income

Wage and salary income is reported most accurately. We don't know whether this stems from the law-abidingness of the middle class or from the extensive system of withholding and information reporting. "Paper trailed" income and income from which taxes are withheld are far more fully reported than income that generates no information for the IRS. Small businesses report a smaller proportion of their income than do large corporations. Rent and royalty income is less fully reported than dividend and interest income. (See U.S. Department of the Treasury 1990 or U.S. General Accounting Office 1990 for details.)

Amount of income

The relationship between the proportion of true tax liability reported and income is U-shaped. Tax compliance is highest for the middle class and lower for both poorer and richer individuals. We do not know the reason for the U-shaped relationship. People with high and low incomes have more opportunities to underreport income than do the vast middle class of wage and salary workers. For example, the poor are often involved in the irregular economy which keeps few written records. The better-off often have complex transactions which make overstating deductions and understating income more difficult to detect. The poor and the rich are often less accepting of mainstream morals and ethics. This may make tax evasion more acceptable, at least for some members of these groups.

Attitudes toward laws and the government

There are large geographic differences in tax compliance, with the Southwest and Southeast being least compliant and the Midwest and Northeast most compliant. It is not clear why these geographic differences exist. However, audit rates have traditionally been lower in the Southwest, particularly southern California, so the tax gamble has a larger expected return in this area. Research indicates that areas of rapid growth are more noncompliant than stable areas. For example, see Beron, Tauchen, and Witte (1992).

A consistent finding in the tax compliance literature is that women are more compliant than men. We do not know whether this stems from greater socialization of women, their greater aversion to risk, or the nature of their incomes.

Why are we concerned with pure tax evasion?

The most obvious answer is that we want everyone to “pay their fair share.” When some members of society pay less than their full tax liability, the rest of us suffer. Government can respond to decreased tax revenue in one of three ways. First, it can increase taxes so the rest of us pay more. Second, it can cut expenditures so that everyone has fewer publicly provided goods and services. Finally, it can finance government expenditures by running large deficits. All of these potential responses to tax evasion hurt those of us who are struggling to survive in the formal economy.

The Irregular Economy

The irregular economy generally encompasses the production of legal goods and services in unregistered and, hence, largely untaxed and unrecorded, small businesses. Activities in the irregular economy are an important form of underground activity.

What goods and services are produced in the irregular economy?

Production in the irregular economy consists mostly of personal and household services. Typical personal services produced in the irregular economy are child care, nursing services, legal, and accounting services. Household services such as gardening, remodeling, and auto

repair also are frequently provided in the irregular economy rather than by registered and licensed individuals.

Some goods and services are also produced in the irregular economy. Examples include food, clothing, jewelry, recreation, and tourism. Evasion of taxes on petroleum products is quite large. Organized crime, particularly the "Russian mafia," is heavily involved in tax evasion on gasoline and other fuels.

Firms producing in the irregular economy tend to be small. Hence, production tends to occur in areas where economies of scale are limited. The irregular economy also provides a gateway for some new immigrants, legal and illegal, and for some new entrepreneurs.

While tales of sweatshops employing hundreds of illegal aliens appear in the press, the more typical employment setting is an individual moonlighting in the irregular economy.¹ The labor force for the irregular economy comes largely from people who work full time in the formal economy.

Costs and benefits of the irregular economy

To those in the formal economy. The confirmation hearings of early Clinton nominees point up the prevalence of purchases from the irregular economy by all strata of U.S. society. We would be hard-pressed to run our households and provide for our personal needs without the irregular economy. This economy provides us with more convenient and often less expensive (both in terms of time and money) goods than the formal economy. The quality of goods and services provided by the irregular economy tends to be highly variable, since the only quality controls are word-of-mouth reports. Regulation and government information provide important quality control that is missing from the irregular economy.

Irregular economic activity pays little if any tax. This means that governments must tax the rest of us more, provide fewer services or run large deficits—a privilege only allowed the federal government and selected state and local governments.

To those working in the irregular economy. Jobs in the irregular economy have advantages and disadvantages. For those moonlighting in the irregular economy (the majority of this economy's labor force), the economy has primarily benefits. Income is free of taxation, and

work methods are left to the individual's discretion. Work may be more informal and better integrated with the family and communal elements of the individual's life than is possible with much formal employment.

Individuals working exclusively in the irregular economy are of two basic types—those there by choice (e.g., retirees, welfare recipients, mothers with small children, “free spirits”) and those there for lack of alternatives (e.g., illegal aliens, the unemployed without unemployment benefits, new entrants to the labor force).

For individuals working exclusively in the irregular economy by choice, the irregular economy provides an entrepot or a more agreeable combination of work conditions and wages. Take for example the individual with small children. She may choose to work at home because of an ability to combine child care with her work activities. Similarly, the handicapped individual may find the home environment with its familiarity or convenience more congenial than the more rigid and structured work environments typical of the formal economy.

For those working in the irregular economy for lack of known alternative, the irregular economy provides sustenance, but at a cost. The cost is the foregone benefits available in the formal economy such as health benefits. Work conditions can be harsh and benefits nonexistent.

Illegal Activities

To this point, we have considered mainly activities that run afoul of our regulatory and tax laws, but not our criminal law. To close on what might be considered a low note, I will now turn to activities that violate our criminal laws. In spite of declining crime rates, crime is still one of the topic of most serious current concern to the U.S. public. It also provides the titillation of the forbidden.

In terms of economic activity, crime comes in two forms—illegal transfers of goods and income, and production of illegal goods and services.

Illegal transfers of goods and services

The illegal sector has developed complex ways of transferring income and goods among individuals just as our legal economy has. Lacking the power to tax, the illegal economy transfers by means of

force (e.g., “your money or your life”), or subterfuge (e.g., fraud, counterfeiting).

Like the irregular economy, the illegal transfer economy seems to provide mainly part-time employment. Further, illegal transfers of the “blue-collar” variety provide surprisingly low returns (see Reuter, MacCoun, and Murphy 1990). Illegal transfers by subterfuge, such as the notorious financial dealings of the 1980s, provide far larger returns but to very few individuals. Economists are split on whether or not activities such as “insider trading” increase or decrease market efficiency.

Production and distribution of illegal goods and services

The major goods produced and distributed illegally in our country are illegal drugs. Illegal drugs run the gamut from substances believed to have large-scale debilitating effects (e.g., heroin, crack) to those considered by many to have relatively minor effects (e.g., marijuana). As Reuter, MacCoun, and Murphy (1990) have recently documented, illegal drug sales provide good income for mainly part-time work. However, drug dealing is a very risky endeavor with substantial chances of arrests (over 1 in 5 per year) and death or injury (1 in 10 per year).

As Reuter and his colleagues have described it, selling illegal drugs is the preferred illegal form of “moonlighting” for many urban males, particularly minority males. Drug dealing provides sporadic (mainly nights and weekends), high-paying, part-time employment. Individuals selling drugs daily in Washington, D.C. in the late 1980s typically earned about \$2,000 per month while individuals selling one day a week or less had income supplements of \$50 or less. The typical dealer in Washington, D.C. is a black male between 18 and 29 years of age. (See Pozo (1996) for a further discussion of drug markets.)

Our attitude toward drugs tends to be both emotional and schizophrenic. We have a spectrum of rules governing the sale and use of various drugs. Some drugs (e.g., generic drugs) are sold in mainly unregulated spot markets (i.e., you can buy them freely at your local supermarket or drugstore). Others, such as tobacco and alcohol, are singled out for special taxation (i.e., excise taxation) and purchase is restricted to individuals who have reached a specified age. Prescription drugs require that a medical doctor authorize use. Finally, there are a

whole series of drugs (e.g., marijuana, cocaine, heroin) that we generally make illegal either to sell or use.

Consider the two spectrums—one spectrum of regulation possibilities and one spectrum of taxation possibilities. Generic drugs would fall at one end of the regulation spectrum and illegal drugs at the other. Oddly, generic and illegal drugs both fall on the untaxed end of the tax spectrum and alcohol and tobacco at the high end of this spectrum. If one considers the placement of various drugs on these two spectrums, it is very difficult to see a consistent set of priorities.

The debate about illegal drugs has mainly centered on the two extremes of the regulation spectrum—legalization versus criminalization. This seems to narrow our options unnecessarily. For various drugs that we currently make illegal, we may want to consider a number of different positionings on the regulatory and tax spectrums. For example, one could imagine making marijuana a prescription drug and placing an excise tax on use. For a compilation of recent papers on drug laws, see Krauss and Lazear (1991).

Perhaps we need to think more carefully about the use of drugs and the benefits that they provide users. Some drugs are designed purely to treat disease. It is hard to imagine someone taking an antibiotic for the fun of it. Others serve both treatment and sensory needs (e.g., pain relief, relaxation). Valium is used by some as a tranquilizer, by some as a muscle relaxant, and by some to produce a euphoric state. Still other drugs seem to have mainly recreational or psychological benefits (e.g., alcohol, tobacco).

We also need to consider the external effect that drug use entails. Consumption of some drugs imposes costs on society as a whole in the form of higher medical expenditures, increased crime, or decreased productivity. Consumption of others provides external benefits by lowering medical costs, making people more psychologically healthy (e.g., less depressed, less aggressive), and more productive. Economists would suggest that you encourage use of the drugs with positive external effects and discourage use of those with negative external effects. Encouragement can come in the form of subsidies (e.g., government-financed medical research or subsidized prescriptions such as under Medicaid or Medicare). Discouragement can come in the form of taxation, information campaigns (very effective for tobacco), restrictive

access (e.g., age restrictions on purchase, requiring prescriptions), or making the substance illegal either under the civil or criminal code.

If we placed all drugs on the regulatory and tax spectrums described above and considered how the net benefits of each drug relates to its position on the spectrums, we might find some surprising locations. Such an exercise might also suggest innovative policies in an area where thinking has been impacted for some time.

The major services produced in the illegal economy are gambling (increasingly preempted by legal forms), prostitution, and loan sharking. While these industries are still quite large, all are believed to have declined in recent years. Prostitution has been hard hit by AIDS and more liberal social mores regarding premarital sex. The estimated number of individuals working as prostitutes is believed to have declined since 1980. Legalization of gambling and low interest rates have taken a similar toll on the gambling and loan sharking industries. (See Carlson 1990 for details.)

Costs and Benefits of the Illegal Economy

It is not hard to make the argument that the illegal economy as a whole has costs that substantially outweigh the benefits that flow from this economy. There is even a debate in economics regarding whether or not we should count any benefits that result from infringement of our criminal law. The issue is whether, when breaking the social contract, the individual has lost his or her right to be counted in the "social welfare function." For a discussion of whose benefits should be included in the social welfare function, see Trumbull (1990).

Even if we assume that the gains of individuals who break the social contract are to be counted, it is hard to justify the illegal economy on economic grounds. The transfer payments of the illegal economy are quixotic at best and worsen the income distribution in many cases. Further, these transfers, on balance, lower productivity in the formal economy.

The costs and benefits of the production of illegal goods and services are more difficult to assess. If we accept the proposition that our criminal laws reflect strongly held moral beliefs, then the costs to society of illegal markets are substantial. However, it may be that at least some current laws reflect the moral beliefs of previous generations. If

so, these laws may no longer be justified on efficiency grounds. For example, it may be that many citizens no longer support the sodomy laws or adultery laws that remain on the books in many states.

From an economist's perspective, the relevant question might be: Do the total costs of the activity outweigh the total benefits? This is, of course, the traditional criterion of benefit-cost analysis. In some cases the benefit/cost ratio seems clear. Most citizens would agree that child prostitution imposes more costs than benefits on society and that it has a high moral cost as well. Transactions between consenting adults are more difficult to evaluate. When an adult prostitute makes a deal with an adult client, the deal is clearly utility-enhancing for the parties to the transaction. Most negative externalities from this transaction flow from the fact that prostitution is illegal, not from the transaction per se.

There are clearly transactions between adults that yield sufficient negative externalities that we wish to outlaw them. For example, it is hard to justify legalizing without severely taxing or requiring prescriptions for such drugs as heroin. However, it seems wise to carefully consider the full spectrum of options (i.e., regulations and taxation) before deciding that making something illegal is a reasonable thing to do. It is also wise to consider the degree to which criminal prohibitions can be enforced. Activities between consenting adults are notoriously difficult to enforce. Making laws that are unenforceable may make us feel better. However, they have substantial hidden costs. They tend to decrease respect for law and, hence, make other laws more difficult to enforce.

Beating the System?

Returning to the question with which we began the paper: Is the underground economy beating the system?

Benefits and Costs to the Nibblers and Nonparticipants

Those of us who are only peripherally touched by the underground economy are, I would contend, harmed more than we are benefited by it.

Pure Tax Evasion

We all bitch and moan about taxes, but pay a substantial amount of what we owe. Some tax revenue is wasted or spent on things we would rather not have, but much tax revenue is spent for things we value very highly—e.g., law enforcement, water and sewers, education. Tax evaders lower the amount of tax revenue that our tax system would otherwise be able to produce. The public sector can respond in one of three ways: (1) lower the level of goods and services provided, (2) raise taxes on the rest of us, (3) in the case of the federal government, increase the deficit.

Tax increases enacted to raise revenue lost to evasion tend to fall disproportionately on the wage and salary earnings of the middle class. Tax evasion tends to redistribute income to low-income and high-income groups willing to falsify tax returns. This is an odd and rather unpalatable form of redistribution. Tax evasion imposes an implicit tax on honesty, a characteristic we would be better off subsidizing than taxing.

Many tax evaders do truly beat the system, but their gain is at our cost. While audit rates have fallen dramatically during the last two decades from 10 percent in the late 1960s to under 1 percent today, the amount of information-matching has gone up considerably and will increase further with the IRS systems-modernization effort. The level of penalties for evasion and the frequency of their imposition has gone up dramatically as well.

Decreasing the current level of tax evasion appears to be a socially desirable goal. How can we most cost-effectively achieve this goal? Tax simplification and increased information-reporting coupled with systems modernization seem most promising at this point. The Tax Reform Act of 1986 (TRA86) simplified our tax laws at least in some ways. Tax laws passed since 1986 have, on balance, made our tax code more complicated. They have also raised *nominal* tax rates on better-off individuals. The last time this happened, both evasion and avoidance increased. Nominal tax rates were progressive, but the tax burden did not fall disproportionately on the rich.

We need to have realistic expectations regarding the amount of tax evasion that can be cost-beneficially combated, given our present tax code. The amount is probably less than 20 percent. It would never be

cost beneficial to uncover and punish most tax evasion. As with taxes and death, we need to learn to live with a certain level of tax evasion.

The Irregular Economy

The irregular economy appears to be the least virulent or relevantly most beneficial sector of the underground economy. This economy provides goods and services as well as jobs that might not otherwise be available. However, both the goods and services and the jobs tend to be of highly variable quality.

This economy may be most successfully attacked by providing alternative entrepot for new immigrants and new labor market entrants. It may also be combated by providing lower levels of regulation and increased benefits for participating in the formal economy. For example, the extensive reference and referral system of Massachusetts, which provides customers, training, and services (e.g., accounting services) to child care providers, has led most child care providers, including providers of child care in their home, to obtain state licenses as required by state law. Only approximately 5 percent of family day care providers are unlicensed in Massachusetts. Other states that require family day care providers to be licensed but don't provide benefits to licensees have as much as 50 percent of providers failing to comply with the law.

Some workers in the irregular economy beat the system. Others are beaten by it. The general public may, on balance, benefit from the existence of this economy, but the balance is a delicate one.

The Illegal Economy

This is largely an economy employing young males. It provides an exciting alternative to the low-wage jobs currently available to young males with high school educations or less. However, these jobs come at high cost to the rest of us. The transfers of the illegal economy are very costly in terms of lifestyle, antitheft devices, and fear. Goods generally lose, not gain, value as a result of illegal transfers. This is reflected in the low prices available from fences.

While illegal goods and services often provide net benefits to those who purchase, the benefits may be short-lived and may come at sub-

stantial cost to the general public. Effective policy to combat the illegal economy may combine legalization of some goods and services (e.g., gambling, prostitution, marijuana) with stricter enforcement of others.

Conclusions

Everyone wants a fast buck. It is the American way. We are a nation that began its existence with a revolt against government. We didn't like taxation when the British imposed it upon us and we don't like it much better now that we impose it upon ourselves.

We are a society of individualists who resent government interference in either our work or consumption choices. However, it has been a long while since we have been a frontier society. We have come to rely on government to provide goods and services, to regulate behavior, and to enforce a commonly agreed upon social code. The underground economy points up gaps in government enforcement and shows us the limits of government.

Through our representatives, we make decisions regarding the goods and services we wish our federal, state, and local governments to provide. To finance these expenditures, we agree to a set of user charges and taxes. Evasion of taxes is often in the individual's narrow self-interest. He or she can enjoy publicly provided goods and services without paying for them. In the jargon of economics, the individual can be a "free rider." The tax evader is beating the system, but at the expense of the rest of us.

As citizens, we require licenses to undertake certain types of businesses (e.g., medicine, plumbing) and we regulate the way in which businesses are allowed to operate (e.g., through environmental and health and safety regulations). We do this to enforce minimum standards. Individuals and firms operating in the irregular economy avoid both licensing and regulatory requirements. In some cases this evasion of regulatory requirements has no external costs. For some goods and services, consumers may obtain adequate information regarding the quality of products informally. In other cases, irregular activity has high external costs and information regarding quality may be both hard to obtain and unreliable. Unregulated firms may dispose of hazardous

materials in ways that cause substantial social harm. Sweatshops may provide hazardous and unhealthy work environments. Some unregistered child care providers may damage young children in ways that have serious social consequences.

Some producers and consumers of the goods and services of the irregular economy are beating the system and at little or no cost to the rest of us. Others are beating the system, but imposing substantial costs on the rest of us. We need to think carefully about what we license and what we regulate. We should impose mandatory licensing only in areas where informal information regarding quality is either likely to be unreliable or costly to obtain. In other areas we may be better served by allowing consumers to choose whether they purchase from licensed or unlicensed providers. We should only regulate in areas where external effects are large or where there are substantial asymmetries in power (e.g., the sweatshop situation) or information (e.g., pharmaceuticals).

Our criminal laws reflect our society's generally held beliefs about what behaviors are acceptable. We condemn certain types of transfers (e.g., theft, fraud) because they offend our sense of fairness and because they lower the productivity of our formal economy. We outlaw the sale and consumption of certain goods and services because they offend our moral sensibilities and/or because they have substantial negative externalities.

The mores of our society have changed quite rapidly since the 1960s, with each generation developing its own sense of what is socially acceptable. Our criminal laws change much more slowly. It is hard to say if the rigidity of what Bagehot called the "cake of custom" is, on balance, good or bad. Most violations of most of our criminal laws, on balance, hurt us as a society. It is not even clear that many of the participants in illegal activities (e.g., the young drug dealer or prostitute) are in the long run beating the system.

Are those who participate in the underground economy beating the system? From a short-run, narrowly individual perspective, they by and large are. However, the system they are beating is a system that we set up because we believed that it was by and large beneficial to society. The tax laws, regulations, and criminal code evolve slowly and can be affected by special interests. The underground economy will grow in those areas where existing laws and regulations are inefficient or most at odds with current mores. The underground provides an escape valve.

It can also provide valuable information regarding rules and regulations that need to be reassessed.

Tax evasion peaked in the early 1980s because our tax laws had rates and provisions that made evasion both relatively easy and very beneficial. Since TRA86, levels of evasion have declined. TRA86 broadened the base of taxation by limiting loopholes. It also lowered tax rates. The growth of the tax evasion component of the underground economy during the late 1970s and early 1980s warned us that something was wrong with our tax laws. Our response was to alter the tax code, not to increase enforcement.

As a society, we would be wise to look closely at areas of rapid growth in the underground economy. These growth areas may need to be attacked by increased enforcement. For example, the growth of crack use during the 1980s was effectively dealt with by a combination of enforcement and education. However, rapid growth may warn us that it is time to revise outdated and inefficient rules. For example, the failure of many upper middle income nominees for federal posts to pay nanny taxes did not lead to greater enforcement. It led to changes in tax laws and regulations.

NOTE

1. A Michigan study found that 50 percent of illegal vendors of goods and services had regular jobs, and that only 5 percent were working off-the-books full time. It is estimated that over half of all illegal aliens had income tax withheld and, hence, were working in the regular as opposed to the irregular economy.

References

- Beron, K., H. Tauchen, and A. Witte. 1992. "The Effect of Audits and Socio-Economic Variables on Tax Compliance." In *Why People Pay Taxes*, Joel Slemrod, ed. Ann Arbor, MI: University of Michigan Press, 67-89.
- Carlson, Kenneth. 1990. "Unreported Illegal Source Income: 1983-1995." Report, Abt Associates, Cambridge, MA.
- Krauss, M., and E. Lazear. 1991. *Searching for Alternatives: Drug-Control Policy in the United States*. Stanford, CA: Hoover Institution Press.
- Pozo, Susan. 1996, forthcoming. *Price Behavior in Illegal Markets*. Aldershot: Avebury.
- Reuter, P., R. MacCoun, and P. Murphy. 1990. *Money from Crime: A Study of the Economics of Drug Dealing in Washington, DC*. Santa Monica: Rand Corporation.
- Simon, Carl, and Ann D. Witte. 1982. *Beating the System: The Underground Economy*. Boston: Auburn House.
- Trumbull, William N. 1990. "Who Has Standing in Cost-Benefit Analysis?" *Journal of Policy Analysis and Management* 9, 2: 201-218.
- U.S. Department of the Treasury, Internal Revenue Service, Research Division. 1990. *Income Tax Compliance Research*. Publication 1415 (4-90).
- U.S. General Accounting Office. 1990. *Tax Administration: Profiles of Major Components of the Tax Gap*. Washington, DC: General Accounting Office, Publication GAO/GGD-90-53BR.

The Informal Economy

Perspectives from Latin America

Alejandro Portes

Johns Hopkins University

The topic of this chapter is the informal economy, a diversified set of economic activities that are vast and growing in many settings.¹ In the industrialized countries, there appears to be growing consensus among researchers who describe the informal sector as encompassing “those actions of economic agents that fail to adhere to the established institutional rules or are denied their protection” (Feige 1990, p. 990). Or, alternatively, the informal sector includes “all income-earning activities that are not regulated by the state in social environments where similar activities are regulated” (Castells and Portes 1989, p. 12). These definitions do not advance an *a priori* judgment of whether such activities are good or bad, leaving the matter to the investigator.

In the case of the developing economies, however, the “informal economy” concept is less precise. This imprecision may originate from the more diverse nature of these economies, and from the various theories that purport to explain the existence and proliferation of these alternative “centers” of production in developing economies.

This paper reviews and explores the various concepts of informal production used to describe what many are observing in the less-developed world. Alternative measures that have been employed to quantify and characterize these sectors for the Latin American economies will be presented and discussed.

Defining the Informal Economy in the Less-Developed World

The concept of the informal economy was born in the Third World, out of a series of studies on urban labor markets in Africa by Keith Hart. Popular entrepreneurship in Accra and other African capitals was

clearly at odds with received wisdom from “the western discourse on economic development” (Hart 1990, p. 158).

In a report to the International Labour Office (ILO), Hart outlined a dualist model of income opportunities of the urban labor force by distinguishing between wage employment and self-employment. Hart emphasized the notable dynamism and diversity of the self-employed and their activities which went well beyond “shoeshine boys and sellers of matches” (Hart 1973, p. 68). However, Hart’s dynamic characterization of the informal sector was subsequently lost as the concept became redefined and institutionalized within the ILO bureaucracy as synonymous with poverty. The informal economy was taken to refer to an urban “way of doing things” characterized by:

- low entry barriers in terms of skill, capital, and organization
- family ownership of enterprises
- small scale of operations
- labor-intensive production with outdated technology
- unregulated and competitive markets (Peattie 1980)
- low levels of productivity and a low capacity for accumulation (Tokman 1982).

In later publications of the ILO’s Regional Employment Programme for Latin America (PREALC), employment in the informal sector was consistently referred to as “underemployment” affecting those workers who could not gain entry into the modern economy (PREALC 1985; Garcia 1991). Eventually the informal sector became known as the “excluded” sector in less-developed economies in numerous ILO, PREALC, and World Bank studies of urban poverty and labor markets (Sethuraman 1981; Gerry 1978; Tokman 1978).

The ILO’s negative characterization of the informal sector has been challenged by other students of the subject who see the informal sector in precisely the opposite light. The alternative view is that informal activities are a sign of the popular entrepreneurial dynamism, described by Hart (1990, p. 158) as “people taking back in their own hands some of the economic power that centralized agents sought to deny them.” The Peruvian economist Hernando De Soto reformulated Hart’s original theme in *The Other Path* (1989). De Soto defines infor-

mality as the popular response to the rigid "mercantilist" state, found dominant in Peru and other Latin American countries, that survive by granting the privilege of legal participation in the economy to a small elite. Hence, unlike its portrayal by ILO and PREALC as a survival mechanism in response to insufficient modern job creation, informal enterprise represents the irruption of "real" market forces in an economy straitjacketed by favoritism and state regulation.

The massive population shifts from the countryside to the cities in the decades before 1980 provided the demographic base for De Soto's informal economy. The urban elites who were already in place were hostile to the migrants because "each person who migrates to the capital is in some way a potential competitor and it is a natural inclination to try to avoid competition" (1989, p. 11). Rural migrants to Peru's cities became informal workers due to the legal barriers that prevented them from participating in the mainstream economy. Informal economic activity was originally a survival mechanism, but gradually it expanded in response to the rigidities and limitations of the economy. The provision of goods and services by the informal economy proved to be cheap and efficient, leading to the conversion of unregulated enterprise into the real economic core of many industrial and service activities.

De Soto's call to dismantle state regulatory barriers so that popular entrepreneurship could flourish found a receptive audience in a number of international development agencies, some of which made his views their own (Bromley 1990). However, this optimistic portrayal also gave rise to empirical and theoretical contradictions. Whereas PREALC's analysis overemphasized the marginal character of urban informal activities and the associated poverty, De Soto's analysis leaned in the opposite direction.

A third perspective, which we will refer to as structural articulation, evolved in dialogue with the earlier work by Keith Hart and ILO experts, as well as subsequent analysis of the phenomenon by PREALC. In agreement with definitions of informality in other world regions, structural articulation characterizes the phenomenon as income-earning activities unregulated by the state in contexts where similar activities are so regulated (Castells and Portes 1989; Feige 1990). This approach is similar to that of De Soto in that it emphasizes the role of the state in the emergence and growth of the informal econ-

omy. However, it departs from De Soto's view by viewing this sector as linked to the formal economy, and it does not subscribe to the notion that this sector is composed exclusively of survival activities. In place of the dualistic images of Latin American urban economies, this third approach describes unified systems that involve many connections between formal and informal enterprises (Roberts 1993; Beneria 1989; Portes and Benton 1984).

The specific nature of these markets is not uniform across countries, but varies with the degree and specifics of state regulations, the requirements of modern firms, and the size and characteristics of the labor force. This approach is labeled "structuralist" because it focuses on the structure of formal-informal relationship in the economic system. The approach recognizes that the condition of excess labor supply created by rural-urban migration has had more complex consequences than the simple survival of the poor at the margins of the urban economy. Two such consequences are particularly important: the functions that informal enterprise plays in support of modern capitalist accumulation; and the creation of new niches in the labor market, corresponding to new positions in the class structure.

A first linkage of the informal and formal sectors is the supply of low-cost goods and services for workers in formal enterprises. Unregulated small artisans and merchants engage in the provision of everything from cheap clothing and footwear to auto and residential repairs. Foodstands, urban transport, gardening and landscaping, house cleaning, and the sale of second-hand appliances are services supplied by informal entrepreneurs. The consumption of basic goods by workers in the formal sector seldom occurs through regulated market channels but is generally supplied by informal sources. Cheaper informal goods and services increase the consumption "yield" of wages earned in the formal sector of the economy (Peattie 1974; Roberts 1976).

From the standpoint of the economic system as a whole, the informal market represents a subsidy to formal enterprises as it makes labor costs lower than they would be if these sources of supply did not exist. Informal enterprises increase the profitability of their formal counterparts by allowing the latter to maintain lower wage levels. The urban upper and middle class also avail themselves of informally provided services like cheap domestic help, gardening, and personal services to maintain enviable lifestyles (Portes and Walton 1981, pp. 91-94).

A still closer interface between the formal and informal sectors is the common practice in large firms of using the informal labor pool to allocate a variety of production and marketing tasks. Firms may hire workers off-the-books or subcontract production, input supplies, or final sales to informal entrepreneurs. The firm can thereby avoid legal regulations that increase labor costs and decrease managerial flexibility. To the extent that the costly workforce of formal enterprises can be reduced by off-the-books hiring and subcontracting, final profits can be significantly improved.

Because these arrangements are not in compliance with legal regulations, managers of firms are careful to conceal them from the authorities. For this reason, they do not appear in the official statistics, leading users of official data to the erroneous conclusion that these informal-formal linkages do not exist.

The structuralist view also stresses that the informal sector is internally heterogeneous. The coexistence and interactions of the regulated and unregulated activities opens opportunities for a number of people to insert themselves as intermediaries. Informal entrepreneurs who perform these intermediary roles commonly earn more than workers in regulated firms. This results in an informal economy that is more than the simple translation of surplus labor into survival activities.

Demographic Trends in Latin America

The demographic events that help give rise to the informal economy in Latin America took place in the second half of the twentieth century. In the 1950s, Latin America was a continent of rural dwellers with more than half the population living in the countryside and the majority of the workforce consisting of peasants and farm workers. By 1990, three out of four Latin Americans lived in cities, with the majority in large urban centers. This extraordinary transformation is evident in table 1, where the Latin American urban population distribution is compared with that of all less-developed regions and that of the developed countries. Overall, the table suggests that Latin America has achieved levels of urbanization equal to those of the developed world despite much lower per capita incomes.

**Table 1. Share of the Urban Population Within Total Population
(in percent)**

	1955	1965	1975	1985	1990
Latin America	45.3	53.3	61.2	68.5	71.1
Less-developed regions	19.0	23.4	26.4	32.8	41.2
More-developed regions	57.0	63.6	68.8	71.6	73.6

SOURCE: United Nations (1991b, tables A.1, A.5).

NOTE: Less-developed regions comprise all regions of Africa, Latin America, Asia (excluding Japan), Melonesia, Micronesia, and Polynesia. More-developed regions comprise North America, Japan, all regions of Europe, Australia, New Zealand, and the former USSR.

Significant country-by-country differences exist around these regional averages, however. Table 2 illustrates these differences by showing average annual rates of growth for the total population and the urban population of 14 countries in Latin America in the 1950s and in the 1980s. Also shown are the average annual rates of growth for the potential economically active population (EAP) defined as those aged 15-64. At the beginning of this period, Argentina, Chile, and Uruguay were already highly urbanized and their population growth rates were relatively low. Urban growth continued in these countries at a rate similar to that of the developed world and, by 1990, more than 85 percent of their inhabitants were living in cities. Demographic transformation in Brazil, Colombia, Mexico, and Peru were dramatic by contrast, with the share of the urban population growing from one-half or less in 1950 to about three-fourths by 1990.

The migrants who gave rise to such rapid rates of urban growth throughout the post-World War II period did not encounter in the cities anything commensurate with their economic aspirations. New industries in that period did generate significant labor demand (Garcia 1982), but the size of that demand was swamped by the number of new migrants. In 1950, modern nonagricultural employment absorbed 26.3 percent of the Latin American labor force. In the next thirty years, employment in that sector grew by an average of 4.1 percent per year, about the same rate as the total nonagricultural EAP. By the end of the period, therefore, the proportion employed in the modern urban sector was the same as it had been three decades earlier. Estimates of the

excess labor supply in cities during this period ranged from 30 to 60 percent of the urban EAP (Garcia and Tokman 1981; Portes 1985).

Table 2. Demographic Trends in Latin America and Selected Indicators

Country	Period	Average annual rate of growth (in percent)			Percentage urban within total population ^a
		Total population	Population aged 15-64	Urban population	
Argentina	1950-60	1.9	1.6	3.0	73.6
	1980-90	1.3	1.3	1.8	86.2
Bolivia	1950-60	2.2	2.0	2.5	39.3
	1980-90	2.6	2.7	4.2	51.4
Brazil	1950-60	3.1	2.8	5.3	44.9
	1980-90	2.2	2.5	3.4	76.9
Chile	1950-60	2.2	2.0	3.7	69.6
	1980-90	1.7	2.1	2.2	85.6
Colombia	1950-60	3.0	2.6	5.5	48.2
	1980-90	2.2	2.9	2.9	70.3
Costa Rica	1950-60	3.7	3.0	4.5	36.6
	1980-90	2.8	3.2	3.7	53.6
Ecuador	1950-60	2.9	2.4	4.6	34.4
	1980-90	2.7	3.4	4.4	56.9
El Salvador	1950-60	2.9	2.3	3.3	38.3
	1980-90	1.5	1.7	2.2	44.4
Guatemala	1950-60	3.0	2.6	3.8	35.7
	1980-90	2.9	2.9	2.2	42.0
Mexico	1950-60	3.2	2.6	4.8	50.7
	1980-90	2.3	3.6	3.2	72.6
Panama	1950-60	2.9	2.8	3.9	41.2
	1980-90	2.2	3.1	2.8	54.8

(continued)

Table 2. Demographic Trends in Latin America and Selected Indicators (continued)

Peru	1950-60	2.5	2.1	5.3	46.3
	1980-90	2.2	3.0	3.0	70.2
Uruguay	1950-60	1.3	1.3	1.5	80.1
	1980-90	0.6	0.6	0.8	85.5
Venezuela	1950-60	3.9	3.5	6.3	66.6
	1980-90	2.7	3.2	3.6	90.5
Latin America	1950-60	2.8	2.5	4.5	49.3
	1980-90	2.1	2.8 ^b	3.1	71.5

SOURCE: ECLA (1991, tables 1,2,3,5,7); United Nations (1991a, 1991b).

a. Definition of urban varies across countries. Figures are for the year ending the decade.

b. Figure is for the period 1980-85.

Measuring the Informal Economy in Latin America

The informal economy is difficult to measure, since it consists of activities generally unrecorded in official statistics. All measurement strategies attempted so far suffer from limitations. PREALC has relied on proxies available in national censuses and in household surveys in their attempts to determine the magnitude of the informal economy. The approach consists of designating entire occupational categories as informal on the basis of their presumed correspondence with the conceptual definition.

PREALC classifies the self-employed and their unremunerated family workers as "informal" because it is assumed that these workers engage in low-productivity, low-pay activities. While domestic servants were similarly so classified in earlier estimates, they have been excluded from more recent figures for reasons that are not clear. On the other hand, owners and salaried workers in small enterprises were absent from earlier estimates of the informal sector, but they are included in recent ones. How "small enterprises" are defined also var-

ies across countries and years, ranging from businesses employing "less than twenty" workers to "less than five."

Despite these changes and inconsistencies, PREALC is the only agency that has prepared estimates of the informal sector for every country in Latin America. Summary figures of these estimates for 1960 and 1989 are presented in table 3. According to PREALC, during this period approximately 31 percent of the Latin American urban EAP were employed informally. That this estimate is virtually unchanged from 1960 to 1989 is noteworthy, given both the important changes in Latin American economies and the shifting occupational categories included in PREALC's definition of the informal sector.

Table 3. Estimates of Urban Informal Employment as a Percent of Urban Economically Active Population

Country	1960	1989
Argentina	21.1	28.7
Bolivia	62.3	27.0
Brazil	27.3	28.6
Chile	35.1	30.0
Colombia	39.0	27.3
Costa Rica	29.3	22.0
Ecuador	35.2	n.a.
El Salvador	42.6	n.a.
Guatemala	51.6	n.a.
Mexico	37.4	34.8
Panama	25.3	n.a.
Peru	46.9	39.0
Uruguay	18.6	19.0
Venezuela	32.3	23.3
Latin America	30.8	31.0

SOURCE: PREALC estimates given in Garcia and Tokman (1981, table 1); Garcia (1991, tables 5, 7, 9); and Infante and Klein (1990, tables 2, 4).

Variations of the magnitude of the informal sector by country are rather large. The estimated share of persons in the informal sector within the total urban labor force ranged during the period from a low of approximately 20 percent in the more-developed countries such as Costa Rica, Uruguay, and Venezuela to about 40 percent in Peru and El Salvador, to approximately 60 percent in Bolivia and Ecuador.

The organization created by Hernando De Soto, the Institute for Liberty and Democracy (ILD), has not produced similar region-wide estimates, but has concentrated on documenting the extent of informality in metropolitan Lima, focusing on three sectors: housing, transport, and petty commerce. ILD estimates that, in 1982, some 43 percent of all housing in Lima was built informally, providing shelter to 47 percent of the city's population. The replacement cost of this housing was calculated at US \$8.3 billion. In commerce, over 91,000 street vendors supported an estimated 314,000 people and generated gross sales of US \$322 million per year. In addition, according to ILD, 39,000 other informal merchants built 274 street markets with an estimated value of US \$41 million, supporting 125,000 people. In transport, informal entrepreneurs controlled over 90 percent of urban public buses. The 1984 replacement value of their fleet was estimated at US \$620 million. The value of the related infrastructure (gas pumps, repair shops, etc.) was estimated at US \$400 million.

Other estimates made by De Soto and his associates include the claim that in the mid-1980s, 61.2 percent of total work hours in Peru were dedicated to informal activities; that 48 percent of the economically active population was engaged in informal activities; and that the latter accounted for 38.9 percent of gross domestic product (GDP), a share projected to rise to 61.3 percent by the end of the century (De Soto 1989: 12).²

It is worth noting that De Soto's definition of the informal sector coincides roughly with the definition of the "underground" economy in developed societies, although methods for measuring the latter have not been applied systematically in Latin America. Underground activities in developed economies have been defined as those activities that take place outside the existing legal framework. Their aggregate value, in terms of relative proportion of the GDP, have typically been estimated by macroeconomic methods (Feige 1990).

It is possible that failure to apply macroeconomic methods to the Latin American economies, despite the similarity of definitions, stems from the lack of suitable data. For the more-advanced Latin American countries, however, this difficulty is not insurmountable. The Center for Economic Research of the Private Sector (CEESP) estimated the size of Mexico's underground economy between 1970 and 1985 on the basis of two macroeconomic methods. The first, adapted from Tanzi (1980, 1982), estimates the currency required for the operation of the legal (aboveground) economy and then subtracts this figure from actual currency in circulation. The difference multiplied by the velocity of money provides an estimate of the magnitude of the underground economy. The ratio of that figure to observed GDP then gives the proportion of the national economy represented by these activities. The method depends on the identification of a base period in which the informal economy is assumed to be insignificant.

Estimates of the Mexican underground economy based on Tanzi's monetary method are reproduced in table 4, column 1. The Mexican underground economy is estimated to be 25 to 30 percent of official GDP in 1985, the last year for which these figures are available. Table 4 also presents estimates of the Mexican underground economy based on the "physical input" method. With this method, the ratio of some physical input of wide use (electricity consumption in this case) is calculated for the base period and then extrapolated to the present. Assuming a relative constant ratio of electricity consumption to GDP, it is possible to calculate the expected GDP for each year following the base period. The difference between observed and expected GDP is attributed to the underground economy. As seen in table 4, these methods yield estimates of the informal sector ranging from 20 to 40 percent of GDP in the 1980s. CEESP notes that both methods show the magnitude of underground activities as increasing steadily.

These macroeconomic estimates have been criticized on various grounds. They do not differentiate between criminal activities and informal activities proper. The latter involve goods and services that are otherwise licit, but whose production and sale are irregular (Castells and Portes 1989). Hence, the huge estimates of informality that are sometimes reached through these methods may be inflated by the inclusion of large organizations that provide "illegal" goods and services (e.g., cocaine, prostitution). These underground operations are of

a nature and size different from those of microentrepreneurs and artisans in the informal sector proper. Macroeconomic estimates have also been criticized for being sensitive to the choice of benchmark (base) year. Despite these limitations, macroeconomic methods provide the best available approximations of the relative weight of unregulated activities in national economies. Their application in Latin America is still rare.

Table 4. Estimates of "Subterranean" Economy in Mexico as a Percent of Official Gross Domestic Product, Calculated by Two Methods, 1970-85

Year	Monetary method	Physical input method
1970	13.5	8.0
1971	13.8	13.5
1972	15.4	15.4
1973	15.2	14.7
1974	21.1	18.7
1975	27.3	19.4
1976	25.6	22.3
1977	27.4	28.5
1978	28.0	30.9
1979	24.9	28.6
1980	33.2	23.6
1981	29.1	25.1
1982	39.3	20.6
1983	29.4	30.1
1984	28.0	33.5
1985	25.7	38.4

SOURCE: CEESP (1987, table 10).

Empirical work undertaken from a structuralist perspective has focused mainly on documenting the relationships between the formal and informal sectors of urban economies. For national estimates, this school has relied on secondary figures on the proportion of the economically active population that is not covered by the social security

system. Absence of coverage serves as a proxy for unregulated work. Most estimates of social security coverage or exclusion are reported for the total, not urban, economically active population. They also do not differentiate between workers and microentrepreneurs. Further, coverage by the social security system does not exclude the possibility that workers engage in unregulated activities on the side; hence it underestimates the actual magnitude of informality. Lastly, as Mesa-Lago (1991) notes in his analysis of Latin American social security systems, the quality of the data is generally poor and furthermore does not differentiate between types and levels of coverage.

Despite these limitations, a comparison of the figures on the EAP excluded from legal labor protection and figures on "underemployed" workers, provided by PREALC, is instructive. This comparison is presented in table 5 for the Latin American economies. Except for Brazil, where the extension of some assistance to the entire population led to an official claim of near universal coverage, the uncovered percentage of the EAP shown here consistently exceeds the PREALC estimates. The range of the gaps is from a few percentage points to over 40 percent of national EAP. These gaps can be tentatively interpreted as an approximation to the proportion of wage workers who labor under irregular conditions, a category implicitly defined by PREALC as formal and by the structuralists as informal. For Latin America as a whole, the difference is approximately 14 percent of the EAP in 1980.

One of the few large data sets that contain reliable information on both employment category and social security coverage is the Mexican government's Urban Employment Survey of 1989. The survey collected data on a representative sample of about 30,000 adult Mexican workers in seven metropolitan areas, including the four largest cities. As analyzed by Roberts (1993), this survey yields estimates of the proportion of the employed labor force working informally according to the PREALC and structuralist definitions, as well as an internal differentiation of the informal sector by employment categories.

Table 6 presents results of a reanalysis of these data using the four largest urban centers. Mexico City, Guadalajara, Monterrey, and Ciudad Juarez. The empirical definition of informality commonly employed by PREALC yields estimates that are approximately half in magnitude (19.6 percent) relative to those produced when the lack of social security protection is the basis of categorization (46.8 percent).

Inclusion in the PREALC estimates of owners of and workers in microenterprises significantly reduces this difference (37.6 percent), but the figures are still lower than those based on lack of coverage (46.8 percent). The remaining gap is directly attributable to workers in large firms who do not receive legal protection ($46.8 - 37.6 = 9.2$ percent).

Table 5. Estimates of Informal Sector in Latin America, 1980

Country	Economically active population not covered by social security (percent)	Underemployed ^a within economically active population (percent)
Argentina	30.9	25.7
Bolivia	81.5	74.1
Brazil	13.0 ^b	44.5
Chile	32.7	28.9
Colombia	80.3	41.0
Costa Rica	51.6	27.2
Dominican Republic	88.7 ^c	40.6
Ecuador	78.7	63.3
El Salvador	88.4	49.0
Guatemala	66.9	50.9
Honduras	85.6	49.7
Mexico	59.5	40.4
Nicaragua	81.1	52.1
Panama	47.7	45.5
Peru	62.7	55.8
Uruguay	34.2 ^d	27.1
Venezuela	55.8	31.5
Latin America	56.3 ^e	42.2

SOURCE: Mesa-Lago (1985, p. 342, 1991, p. 50); Portes, Blitzer, and Curtis (1986, p. 731); PREALC (1982, pp. 34-81); Wilkie and Perkal (1985, table 1308); Wilson (1985, p. 254).

a. Defined by PREALC as the sum of self-employed workers minus professionals, unremunerated family workers, domestic servants, and "traditional" rural workers.

b. Coverage in 1980 is based on selected assistance programs according to universalistic criteria rather than tied to employment.

c. 1985 figure.

d. Based on a probability survey of Montevideo's working-class population in 1985. Official figures report near-universal coverage based on selected programs extended to all citizens.

e. Weighted average.

Table 6. Estimates of Urban Informal Economy in Mexico Based on Occupational Classification and Lack of Social Coverage, 1989**Share of informal employment within the urban economically active population (percent)****By occupation**

Unremunerated family workers plus the self-employed	19.6
Above two categories plus workers in micro-enterprises and unregistered microentrepreneurs	37.6

By lack of coverage^a

Entrepreneurs	7.2
Self-employed	14.0
Workers	25.6
Total	46.8
N	22,019

SOURCE: Roberts (1993, table 3).

a. Entrepreneurs, workers, and the self-employed without social security protection.

The breakdown of the informal sector by employment reveals that microentrepreneurs, the best-paid class within the sector, represent about 7 percent of the employed labor force in the largest cities in Mexico. The self-employed comprise an average of 14 percent of the urban EAP and 30 percent of those working informally. Unprotected wage workers represent one-fourth of all urban workers and the majority (55 percent) of the informal labor force.

Summary

There exists a rich literature on the underground economy for the industrialized economies. In contrast, understanding the informal economy is more challenging for students of the less-developed economies because researchers have been less consistent when defining informal activities for this set of nations. The lack of consistency in

describing the informal economy for the developing world may be a result of the greater diversity found in the economic organization and in the level of development of these economies.

Broadly speaking, three different descriptions of the informal economy in the developing world can be found in the literature. The dualist model adapted by the ILO characterizes the informal sector as a poor, noncompetitive sector, employing those who fail to gain entry into the modern “productive” sector. Using this approach it has been estimated that about 30 percent of workers in Latin America are employed in the informal sector.

In contrast, De Soto characterizes the informal sector as competition emerging through and in spite of confining state regulations and an oligopolistic market structure. In this view, informal activities introduce dynamism and strength into the economy. Using information gathered from Lima, Peru, De Soto estimated that about 40 percent of Lima’s GDP is produced in informal markets.

Structural articulation is a third approach to the informal economy in developing areas. As in the De Soto view, the avoidance of regulations gives impetus to this sector. However, unlike the ILO approach, the formal and informal sectors are closely linked, and to a degree, complementary to each other. The profitability of the formal sector is enhanced due to the availability of low-cost inputs from the informal sector and due to the possibility of subcontracting in the informal market, where costly regulations are easily skirted. Using information that attempts to differentiate workers who are employed in jobs that are regulated versus jobs that are not, it is suggested that, on average, more than half the economically active population in Latin America are informal workers.

NOTES

1. Much of the material in this chapter can be found in Portes and Schauffler (1993) and Portes (1994).

2. The validity of these estimates has been questioned by Rossini and Thomas (1987), who argued that they are inaccurate due to: the inappropriate use of monetary measures of transactions counted in the GDP; poorly specified models; and mishandling of econometric methods.

References

- Beneria, Lourdes. 1989. "Subcontracting and Employment Dynamics in Mexico City." In *The Informal Economy: Studies in Advanced and Less Developed Countries*, Alejandro Portes, Manuel Castells, and Lauren A. Benton, eds. Baltimore: Johns Hopkins University Press, pp. 173-188.
- Bromley, Ray. 1990. "A New Path to Development? The Significance and Impact of Hernando De Soto's Ideas on Underdevelopment, Production, and Reproduction," *Economic Geography* 66: 328-348.
- Castells, Manuel, and Alejandro Portes. 1989. "World Underneath: The Origins, Dynamics, and Effects of the Informal Economy." In *The Informal Economy: Studies in Advanced and Less Developed Countries*, Alejandro Portes, Manuel Castells, and Lauren A. Benton, eds. Baltimore: Johns Hopkins University Press, pp. 11-37.
- CEESP. 1987. *La Economía Subterránea en México*. Mexico City: Editorial Diana.
- De Soto, Hernando. 1989. *The Other Path*. Translated by June Abbott. New York: Harper and Row.
- Economic Commission for Latin America (ECLA). 1981. *Statistical Yearbook for Latin America*. New York: United Nations.
- Economic Commission for Latin America and the Caribbean (ECLAC). 1991. *Statistical Yearbook for Latin America*. New York: United Nations.
- Feige, Edgar L. 1990. "Defining and Estimating Underground and Informal Economies: The New Institutional Economics Approach," *World Development* 18, 7: 989-1002.
- Garcia, Norberto E. 1982. "Growing Labor Absorption with Persistent Unemployment," *CEPAL Review* 18: 45-64.
- . 1991. *Reestructuración, Ahorro, y Mercado de Trabajo*. Santiago de Chile: PREALC.
- Garcia, Norberto E., and Victor E. Tokman. 1981. "Dinámica del Subempleo en América Latina." In *Estudios e Informes de la CEPAL*. Santiago de Chile: CEPAL.
- Gerry, Chris. 1978. "Petty Production and Capitalist Production in Dakar: The Crisis of the Self-Employed," *World Development* 6 (September-October): 1187-1198.
- Hart, Keith. 1973. "Informal Income Opportunities and Urban Employment in Ghana." Paper presented at the Conference on Urban Unemployment in Africa. Institute of Development Studies, University of Sussex, September 1971. Subsequently published in 1973 in revised form in *Journal of Modern African Studies* 11: 61-89.

- _____. 1990. "The Idea of Economy: Six Modern Dissenters." In *Beyond the Marketplace, Rethinking Economy and Society*, R. Friedland and A.F. Robertson, eds. New York: Aldine de Gruyter, pp. 137-160.
- Infante, Ricardo, and Emilio Klein. 1990. *Empleo y Equidad: Desafío de los 90*. Documento de Trabajo, No. 435. Santiago de Chile: PREALC.
- Mesa-Lago, Carmelo. 1985. "Alternative Strategies to the Social Security Crisis: Socialist, Market, and Mixed Approaches." In *The Crisis of Social Security and Health Care*, C. Mesa-Lago, ed. Pittsburgh: Center for Latin American Studies, University of Pittsburgh.
- _____. 1991. "Social Security and Prospects for Equality in Latin America." Discussion Paper No. 140, World Bank, Washington, DC.
- Peattie, Lisa R. 1974. "Living Poor: A View from the Bottom." In *Proceedings of the Colloquium on Urban Poverty: A Comparison of the Latin American and the U.S. Experience*. Los Angeles: UCLA School of Architecture and Urban Planning.
- _____. 1980. "Anthropological Perspectives on the Concept of Dualism, the Informal Sector, and Marginality in Developing Urban Economies," *International Regional Science* 5: 1-31.
- Portes, Alejandro. 1985. "Latin American Class Structures: Their Composition and Change During the Last Decades," *Latin American Research Review* 20: 7-39.
- _____. 1994. "The Informal Economy and Its Paradoxes: The Social Basis of Unregulated Entrepreneurship." In *Handbook of Economic Sociology*, Neil J. Smelser and Richard Swedberg, eds. Princeton, NJ: Princeton University Press.
- Portes, Alejandro, and Lauren Benton. 1984. "Industrial Development and Labor Absorption: A Reinterpretation," *Population and Development Review* 10 (December): 589-611.
- Portes, Alejandro, Silvia Blitzer, and John Curtis. 1986. "The Urban Informal Sector in Uruguay: Its Internal Structure, Characteristics, and Effects," *World Development* 14 (June): 727-741.
- Portes, Alejandro, and Richard Schauffler. 1993. "Competing Perspectives on the Latin American Informal Sector," *Population and Development Review* 19 (March): 33-60.
- Portes, Alejandro, and John Walton. 1981. *Labor, Class, and the International System*. New York: Academic Press.
- PREALC. 1982. *Mercado de Trabajo en Cifras: 1950-1980*. Santiago de Chile: International Labour Office.
- PREALC. 1985. *Mas Alla de la Crisis*. Santiago de Chile: International Labour Office.

- Roberts, Bryan R. 1976. "The Provincial Urban System and the Process of Dependency." In *Current Perspectives in Latin American Urban Research*, Alejandro Portes and Harley L. Browning, eds. Austin, TX: Institute of Latin American Studies Publications Series, University of Texas, pp. 99-132.
- _____. 1993. "The Dynamics of Informal Employment in Mexico." In *Work Without Protections, Case Studies of the Informal Sector in Developing Countries*, Gregory K. Schoepfle and Jorge F. Perez-Lopez, eds. Washington, DC: U.S. Department of Labor, Bureau of International Labor Affairs.
- Rossini, R. G., and J. J. Thomas. 1987. "Comentarios al Equipo Económico del ILD." In *Los Fundamentos Estadísticos de El Otro Sendero: Debate Sobre el Sector Informal en el Perú. Taller de Investigación*, R. G. Rossini et al., eds. Lima: Friedrich Ebert Foundation.
- Sethuraman, S. V. 1981. *The Urban Informal Sector in Developing Countries*. Geneva: International Labour Office.
- Tanzi, Vito. 1980. "The Hidden Economy: A Cause of Increasing Concern," *FMI Bulletin* 9:34-37.
- _____. 1982. *The Underground Economy in the United States and Abroad*. Lexington, MA: D.C. Heath.
- Tokman, Victor E. 1978. "An Exploration into the Nature of Informal-Formal Sector Relationships: The Case of Santiago," *World Development* 6 (September-October): 1065-1075.
- _____. 1982. "Unequal Development and the Absorption of Labour: Latin America 1950-1980," *CEPAL Review* 17: 121-133.
- United Nations. 1991a. *World Population Prospects*. New York: United Nations.
- United Nations. 1991b. *World Urbanization Prospects*. New York: United Nations.
- Wilkie, James W., and Adam Perkal (eds.). 1985. *Statistical Abstract of Latin America*. Vol. 24. Los Angeles: University of California, Latin America Center.
- Wilson, Richard R. 1985. "The Impact of Social Security on Employment." In *The Crisis of Social Security and Health Care*, C. Mesa-Lago, ed. Pittsburgh: Center for Latin American Studies, University of Pittsburgh, pp. 247-278.

INDEX

- Abrahamse, Allan, 92
Abt Associates, 8
Adjusted gross income (AGI)
 gap between IRS reports and NIPA estimates, 11-12
 as indicator of underground activities, 11
Allingham, Michael G., 105
Alm, James, 107, 113, 117, 119, 120, 121, 122, 123
Anderson, P., 19
Avery, R. B., 6, 17
Axelrod, Robert, 108

Bagnoli, Mark, 108
Bahl, Roy, 107
Bank Secrecy Act (1970). *See* Currency and Foreign Transactions Reporting Act (1970)
Beck, Paul J., 117, 123
Becker, Gary S., 105
Beneria, Lourdes, 150
Benton, Lauren, 150
Bernstein, Aaron, 90
Beron, K., 133
Boston Youth Survey (1989), 92
Boyum, D., 73
Bromley, Ray, 149
Bulk currency shipments, 35-36

Cagan, P., 16
Carlson, Kenneth, 138
Carson, C., 11
Castells, Manuel, 147, 149, 157
CCNET, 36
Center for Economic Research of the Private Sector (CEESP), 157
Chen, K., 68
Chiricos, Theodore, 93, 94
Clotfelter, Charles, 107

CMIRs. *See* Report of International Transportation of Currency or Monetary Instruments (CMIRs)
Cohen, Lawrence, 94
Collins, Julie H., 117
Cowell, Frank A., 105, 107-8
Crime
 criminal behavior data sources, 83-85
 demographics of criminal population, 93
 earnings from, 90-96
 factors in decision to commit, 90-91
 foraging model of work and, 97-99
 returns to crime, 90-93
 tax evasion as, 103
 See also Earnings from crime; Illegal economy; Propensity to commit crime; Returns to crime
Criminal participation rate estimations, 83
Cronshaw, Mark, 117, 119
CSNET, 36
CTNET, 36
Currency, U.S.
 amount outstanding, 6
 direct estimates of net outflows, 35-47
 estimates of currency held abroad from MD model, 21-30
 estimates of currency held from Note Ratio Model, 30-35
 estimates of foreign holdings of, 20
 implications of foreign holdings, 52-57
 indirect estimates of foreign holdings of, 20-35
 in M1, 6
 missing stock of, 6-7
 new counterfeit protection, 20
 role in foreign countries, 2

- See also* Currency flows; Hoarding hypothesis; Monetary demography model (MDM); Note ratio model (NRM); World dollarization hypothesis
- Currency, U.S.
 bulk currency shipments, 35-36
- Currency and Foreign Transactions Reporting Act (1970), 36
- Currency census system (CCS), proposed, 20-21
- Currency flows
 bulk currency shipments, 35-36
 composite estimates of net, 48-51
 conceptual comparison of direct measures of, 37-38
 estimates of net outflows and inflows of U.S. currency, 35-47
 estimates of New York Federal Reserve District, 35
 estimates of U.S. Customs Service, 35-36
 net outflows derived from CMIRs and NYNET, 36-37
 not captured by CMIRs, 47
- Currency ratio model
 to estimate unreported economy, 15
 general specification, 16
 restrictive specification, 15
See also General currency ratio model (GCR)
- Data sources
 CMIR currency flow information system, 35
 criminal participation rate, 83-84
 foreign drug production prices, 69-70
 number of arrests, 84-85
 number of crimes committed, 85
 self-reported criminal behavior, 84
- Davis, Douglas D., 107, 113
- Davis, Jon S., 117
- De Soto, Hernando, 148-50, 156, 162
- DiIulio, J., 85
- Drug Abuse Warning Network (DAWN), 76
- Drug consumption
 estimates of, 65-68, 70
 spending for, 70-74
- Drug production
 Burma, 65
 estimates of foreign country, 64
 estimates of U.S., 64-65
 INCSR estimates, 64-67
 INM estimates, 66
 Mexico, 66-69, 75-76
 NNICC estimates, 66 potential use of estimates for, 77-78
 price data for foreign, 69-70
 use of estimates, 75-76
- Drugs
 expenditure estimates, 70-74
 production and sale in illegal economy, 136-38
- EAP. *See* Economically active population (EAP)
- Earnings from crime
 decisions related to, 90-93
 incentives to take, 93-96
- Economically active population (EAP)
 employed in Latin American informal sector, 155, 159-60
 employment in Latin America, 152-53
 growth in Latin America, 152
- Economies. *See* Formal economy; Illegal economy; Informal economy; Irregular economy; Underground economy; Unrecorded economy; Unreported economy
- Elliot, Delbert, 94
- Employment
 estimates of informal employment, Latin America, 154-61
 irregular economy, 134-35

- overlap of legal and illegal, 97-99
 - relation to earnings from crime, 93-96
 - in Third World informal economies, 129-30
 - See also* Self-employed individuals
- Entrepreneurship in underground economy, 129-30, 147-48
- Expected utility theory
 - in decision related to tax compliance, 105-6
 - model in tax compliance research, 111-12
- Fagan, Jeff, 97
- Federal Reserve system
 - currency flow estimates, 35
 - Surveys of Currency and Transaction Account Usage (SCTAU), 17
- Feige, E. L., 6, 11, 14, 16, 19, 35, 56, 103, 107, 147, 149, 156
- Foraging model, 97-99
- Formal economy
 - costs and benefits of irregular economy to, 134
 - defined, 129
 - link to informal economy, 150-51
- Freeman, R. B., 84, 86, 92, 93, 94
- Garcia, Norberto E., 148, 152, 153
- GCR. *See* General currency ratio model (GCR)
- General currency ratio model (GCR), 17
- Gerry, Chris, 148
- Godshaw, G., 74
- Gordon, James P. F., 107-8
- Grogger, J., 94-95
- Gutmann, P., 16
- Hageborn, J., 97
- Hart, Keith, 147-48
- Hindelang, M. J., 84
- Hirschi, T., 84
- Hoarding hypothesis, 20
- Holt, Charles A., 107, 113
- Illegal economy
 - as category of underground economy, 135-39
 - components of, 8
 - costs and benefits of, 138-39, 141-42
 - drug market, 63-64
- Incarceration
 - rate in United States, 90-91
 - relation to earnings from crime, 91
 - trends in United States, 85-90
- Income
 - adjusted gross, 11-12
 - declared taxable, 105-6
 - from illegal activities, 8
 - Taxpayer Compliance Measurement Program (TCMP), 12-13, 112
 - of unrecorded economy, 10
 - unreported types and amounts of, 132
- Income, unreported
 - from currency ratio model, 16
 - estimates in United States, 10-11
 - as indicator of underground activities, 11
 - measurement using currency ratio method, 1-2, 16-17
 - total U.S., 52-57
 - in unreported economy, 8
- INCSR. *See* International Narcotics Control Strategy Report (INCSR)
- Index of propensity to commit crime, 86-88
- Induced value theory, 113
- Informal economy
 - with demographic trends in Latin America, 151-54
 - De Soto's definition, 148-49, 156
 - in developing countries, 147-49
 - development of Latin American, 151-54

- economically active population in Latin American, 155, 159-60
 estimates of Mexican, 160-61
 ILO and PREALC characterizations of, 148-49
 in industrialized countries, 147
 link to formal economy, 149-51
 measurement in Latin America, 154-61
 PREALC estimates for Latin America, 155
 as underground economic activity, 9-10
- Institute for Liberty and Democracy (ILD), 156
- Internal Revenue Service (IRS)
 estimates of legal and illegal source income, 15-16
 GCR as benchmark for unreported income, 17
 Taxpayer Compliance Measurement Program (TCMP), 12-14, 112
 yield from enforcement activities, 15-16
- International Labor Organization (ILO)
 interpretation of informal economy, 148
 Regional Employment Programme for Latin America (PREALC), 148
- International Narcotics Control Strategy Report (INCSR), 64-66, 75
 estimates of drug consumption, 72
 estimation methodology, 68
- International Narcotics Matters Bureau (INM), 66
- Irregular economy
 as category of underground economy, 133-35
 costs and benefits to nonparticipants, 141
- Jackson, Betty, 117, 119, 120, 121
 Judson, Ruth A., 24-29, 35, 54
 Juhn, Chinui, 96
- Jung, Woo-Oh, 117
- Kahneman, Daniel, 106
 Kandel, D., 68
 Kleiman, M., 68
 Koppell, R., 74
 Krauss, M., 137
 Krebs, John R., 97
- Land, K. C., 94
 Langan, Patrick, 91
 Lazear, E., 137
 Lee, David, 94
- McCall, Patricia, 94
 McClelland, Gary, 117, 121, 122
 MacCoun, R., 92, 97, 136
 Machina, Mark J., 107
 McKee, Michael, 108, 117, 119, 120, 121, 123
- MDM. *See* Monetary demography model (MDM)
- Mendel, Richard A., 91, 98
 Mesa-Lago, Carmelo, 159, 160t
- Mexico
 drug production estimates, 66-69, 75-76
 estimates of informal economy, 160-61
 estimates of underground employment, 157-58
 Urban Employment Survey (1989), 159
- Mishel, Lawrence, 90
- Monetary demography model (MDM)
 age and quality characteristics, 22-24
 development of, 21-22
 seasonality, series, and coin/bill ratio characteristics, 24-30
- Money supply, M1, 6
- Murphy, Kevin, 96
 Murphy, P., 92, 97, 136
 Murray, Matthew, 107

- National Crime Victimization Survey, 81
- National Income and Product Accounts (NIPA)
estimate of AGI, 11-12
revised estimates of unreported and unrecorded income, 14
- National Institute on Drug Abuse (NIDA), 76
- National Longitudinal Survey of Youth (NYSL), 84-85, 94-95
- NBER Inner City Youth Survey (1980), 91-92, 94
- Needels, Karen, 83
- Net value of \$100 denomination notes (NYNET), 35-36
- NNICC (National Narcotics Intelligence Consumers Committee) report, 66
- Note Ratio Model (NRM)
for estimates of currency held abroad, 30-33
- NYNET. See Net value of \$100 denomination notes (NYNET)
- NYSL. See National Longitudinal Survey of Youth (NYSL)
- Office of National Drug Control Policy, 8, 70-74, 76
- The Other Path* (De Soto), 148
- Pancoast, R., 74
- Peattie, Lisa R., 150
- Phillips, L., 82
- Piehl, A., 85
- Plott, Charles R., 116
- Plumlee, R. David, 117
- Porter, Richard D., 24-29, 35, 54
- Portes, Alejandro, 147, 149, 150, 153, 157
- Pozo, Susan, 136
- PREALC (Regional Employment Programme for Latin America), 148
- Price data, foreign drug production, 69-70
- Propensity to commit crime
alternative ways to account for trend, 87-89
estimates of, 87
index of, 86-88
role of supply responses in, 93-96
- Report of International Transportation of Currency or Monetary Instruments (CMIRs), 36
- Returns to crime, 90-93
- Reuter, P., 77-78, 92, 97, 136
- Rhodes, W., 67, 71t, 72t, 73t
- Roberts, Bryan R., 150, 159, 161t
- Rocheleau, A. M., 73
- Roth, Alvin E., 116
- Roth, Jeffrey A., 108
- Samuelson, Paul, 108
- Sandmo, Agnar, 105
- Scholz, John T., 108
- Schulze, William D., 117, 121, 122
- Self-employed individuals
criminals as, 97
as informal workers, 154
- Sethuraman, S. V., 148
- Simon, Carl, 131
- Singer, Max, 77
- Smith, Vernon L., 113, 114
- Social norms
evidence for impact on tax compliance, 121-23
as factor influencing tax compliance, 108-10
related to underground economy, 142-44
- Sprenkle, C., 6
- Stephens, David W., 97
- Sumner, Scott B., 19
- Tanzi, Vito, 103, 107, 157
- Tauchen, H., 133

Tax compliance

- basic design of experiments, 115
- data and information sources, 112-13
- economics-of-crime methodology, 105-7
- evidence for factors influencing, 117-23
- expected utility model in research in, 111-12
- experimental evidence of, 117-23
- experimental research, 113-17
- factors influencing decisions related to, 105-11
- interpreting experimental results, 116-17
- IRS yields from enforcement, 15-16

Tax evasion

- benefits and costs to nonparticipants, 140-41
- as category of underground economy, 132-33

Taxpayer Compliance Measurement

Program (TCMP), 13-14, 112

Tokman, Victor E., 148, 153

Topel, Robert, 96

Trumbull, William N., 138

Tversky, Amos, 106

Twain, Mark, 103

UCR. See Uniform Crime Reports

(UCR), Federal Bureau of Investigation

Underground economy

- benefits and costs to nonparticipants, 139-42
- categories of activities in, 131-39
- CEESP estimates of Mexican, 157-58

- entrepreneurship in, 129-30, 147-48
- foreign holdings of U.S. currency, 5-7
- illegal, 8-9
- informal, 9-10
- unrecorded, 9-10
- unreported, 8-10

Uniform Crime Reports (UCR), Federal Bureau of Investigation, 81

U.S. Department of the Treasury, 132

U.S. Drug Enforcement Administration, System to Retrieve Information from Drug Evidence (STRIDE), 72

U. S. General Accounting Office (GAO), 67, 132

Unrecorded economy, 9-10

Unreported economy

- currency ratio model to estimate income in, 16-20
- estimates of unreported income, 10-11
- as underground economic activity, 8-10

Urban Employment Survey (1989), Mexico, 159

Viscusi, W. Kip, 94

Votey, H., 82

Walton, John, 150

Webley, Paul, 117

Weis, J., 84

Westat, Inc., 109

Wilde, Louis L., 113

Wilson, J., 92

Witte, Ann D., 108, 131, 133

World dollarization hypothesis, 20

Yanklovich, Skelly, and White, Inc., 109