The Underground Economy in the United States: Estimates and Implications*

Definition and Causes

After being ignored for many years, the underground economy has finally worked its way to the center stage of American public and official attention. The "discovery" of the underground (or subterranean) economy in the United States is attested by recent editorials and articles in leading newspapers and magazines, hearings by four congressional committees, reports by official agencies (Internal Revenue Service and General Accounting Office), coverage in popular television programs, and professional articles. In this discovery the United States has followed other countries where the underground economy has been the subject of official and unofficial attention and studies for some time.

The 1977 Report of the *Deutsche Bundesbank* reported (p. 23) that « ... [in Germany] cash payment is unquestionably gaining ground again in some fields, notably in the "grey areas" of business activity where services are rendered without taxes ... and [are] settled in cash ». Sir William Pile, then chairman of the Board of Britain's Inland Revenue Service, was quoted last year to the effect that "unrecorded income" may be as high as 7.5 per cent of that country's gross national product (GNP). That figure was recently revised downward to 3.5 per cent by the Central Statistical Office. In recognition of unrecorded activities of different origin, Italy's Istituto Centrale di Statistica in 1978 made upward revisions,

^{*} Paper presented at the 1980 Meeting of the American Economic Association held in Denver, Colorado, September 5-7, 1980. The views presented are personal and not official IMF views. The author is grateful to Ke-Young Chu, Patrick Honohan and William H. White for comments on parts of an earlier draft of this paper and to Mrs. Chris Wu for valuable statistical assistance.

of the order of 10 per cent, to the 1975-78 national accounts series. This substantial adjustment was still considered inadequate by some observers. The Economic Council of Denmark estimated the underground economy of that country at 10 per cent in 1970 and 6 per cent in 1974/75. Upward revisions of the national accounts data have also been made in recent years by France, Germany, Japan, Sweden, and several other countries. Furthemore, discussions of international tax evasion and avoidance connected with unreported offshore activities have been placed on the agenda of various international official meetings. Thus, like the wind, the underground economy may still be hidden to the eye, but its presence is, now, very much felt.

To make sure that the concept of underground economy means the same thing to different observers, it must be defined. An explicit and satisfactory definition is not readily available in the yet limited professional literature. In particular while some observers seem to think in terms of a gross concept — i.e., one based on the total expenditure on illegal and/or unreported activities — others have in mind a concept related to the net use of resources. Thus, some observers would include all expenditure on illegal drugs and gambling; others would include only the incomes that originate from these activities. Often the definitions given do not specify whether the observers are using gross or net concepts. A simple and convenient definition of the underground economy is the following: it is gross national product that, because of unreporting and/or underreporting is not measured by official statistics. This is the definition followed in this paper. Of course to the extent that this would be taxable income, it is associated with a loss in tax revenue.

The two main groups of factors that create an underground economy are taxes and restrictions. Both of these factors can bring about an underground economy.

Taxes: Even if there were no restrictions on activities, taxes alone would force some activities underground so that they would go unrecorded and would thus escape the payment of taxes. All taxes would do this but, in a given country at a given time, certain taxes are likely to be more important thant others. In the United States the discussion has almost exclusively centered on income taxes. In Europe, on the other hand, social security taxes

and value-added taxes have also been prominently mentioned. If one would extend the coverage to developing countries, foreign trade taxes would have to be added to the list. The role of taxation in bringing about underground activities will be discussed in greater detail below.

Restrictions: If there were no taxes, there would still be an underground economy because of various governmental restrictions on the activities of economic agents. The restrictions are imposed either because the activities themselves are inherently criminal or illegal or for other socioeconomic reasons. The restrictions, especially when accompanied by high penalties and/or efficient controls, may at times prevent these activities from coming into existence. More often, however, they will force those activities underground. As these activities still use resources and generate incomes, they inevitably bring into existence an underground economy. Furthermore, to the extent that these incomes are not reported to the tax authorities — while they use resources that might have gone into the generation of legal, and thus taxable, incomes — the incidence and possibly the level of total tax revenues are also affected.

Restrictions on legal activities are, for example, those that prevent individuals receiving social security benefits (for retirement or disability), or welfare payments, from earning incomes that exceed statutory limits; or those that prevent aliens without valid work permits to hold jobs; or those that prevent employers from paying wages below the legal minimum. Often, the recipients of these incomes will not report them to the tax authorities. And the payors are not likely to comply with reporting requirements for fear of being fined. Often the related payments are made in currency rather than with checks.

Many activities are inherently illegal and, therefore, forbidden. These include narcotics trafficking, illegal gambling, loansharking, prostitution, bribing of officials, fencing of stolen goods, etc. These activities are conducted almost exclusively through the use of currency, in order to remain anonymous, and contribute to an increase in currency requirements in the economy. They are not tax-induced as they would exist even in the absence of taxation. And they are rarely, if ever, recorded. If the resources that go into these activities would have gone into legal (and thus measu-

rable) activities, rather than remaining unemployed, the net result of the growth of these illegal activities will be a fall in legal GNP in relation to total GNP which would include the legal economy plus the underground economy. As tax revenues are collected from the legal economy and not from the underground, they are inevitably affected.

II. Review of Available Estimates

Up to recently there were no estimates of the size of the underground economy in the United States. Recently, however, three different estimates have become available in published papers reflecting three different approaches. Of these, one, by the Internal Revenue Service, reflects an attempt at measuring directly the underground economy by relying on various bits of information, some far less reliable than others.¹ The other two, by Gutmann and Feige reflect indirect approaches using monetary data. Space limitation precludes a full discussion of these approaches so that only the most salient features will be presented.

Internal Revenue Service: The Internal Revenue Service study analyzes separately unreported legal-source income and unreported illegal-source income. Unreported legal-source income was estimated on the basis of information from the IRS' Taxpayer Compliance Measurement Program (TCMP) covering the tax year 1973. This program subjects a probability sample of about 50,000 individual income tax returns to a thorough audit examination. Through a weighting procedure the results from this examination were adjusted to provide national figures for 1973. The national 1973 results were then inflated to provide data for 1976. The 1976 estimates obtained by use of the TCMP were once again adjusted upward by recourse to information available in the Exact Match File — which matched for 50,000 households information provided by a 1973

household survey, by the records of the Social Security Administration, and by the IRS Individual Master File — and in the Bureau of Economic Analysis of the Department of Commerce. Unreported illegal-source income was obtained from a collage of estimates, guesstimates, and just plain hunches by various sources. For the various categories of unreported income, the IRS study provided lower and higher estimates. These are shown in Table 1.

U.S. ESTIMATES OF UNREPORTED INCOME FOR 1976
BY TYPE OF INCOME
(in billions of dollars)

4/	Lower Estimate	Higher Estimate
Legal-source incomes	74.9	99. <i>7</i>
Self-employment	33.0	39.5
Wages and salaries	21.3	26,8
Interest	5.4	9.4
Dividends	2.1	4.7
Rents and royalties	3.2	5.9
Pensions, annuities, estates, and trusts	3,6	5.4
Capital gains	3.9	5.1
Other	2.3	2.9
Illegal-source incomes	25.3	35.2
Illegal drugs	16.2	23.6
Bookmaking	4.0	5.0
Numbers	2.4	3.0
Other gambling	1.6	2.0
Prostitution	1.1	1.6
Overall total	100.2	134.9

Source: See Internal Revenue Service, Estimates of Income Unreported on Individual Income Tax Returns (Washington, D.C.: Government Printing Office, September 1979).

Thus, according to the IRS study, the underground economy in 1976, as measured by unreported income, ranged from about \$ 100 billion to about \$ 135 billion. In relation to (legal) GNP it ranged from 5.9 per cent to 7.9 per cent.

Gutmann's estimate: Professor Peter Gutmann was the first to use monetary statistics as an indirect measure of the underground economy in an article published in 1977.² His method was applied

¹ See Internal Revenue Service, Estimates of Income Unreported on Individual Income Tax Returns (Washington, D.C.: Government Printing Office, September 1979).

² Peter M. Gutmann, "The Subterranean Economy", Financial Analysis Journal, November-December 1977, pp. 26-27 and 34.

to 1976 and gave an estimate of an underground economy equal to \$ 176 billion. He has recently re-estimated his results for 1979 and has concluded that, in this more recent year, the underground economy could be "conservatively" estimated at \$ 250 billion. For both years the estimates are a little over 10 per cent of (legal) GNP. Gutmann's method is simple and is based on a few key assumptions. First, he assumes that all underground activities avoid the use of checks and rely on currency for making payments. Second, these activities are the net result of high taxes and government-imposed restrictions. Third, and most importantly, the ratio of currency (C) to demand deposits (D) is influenced only by changes in taxes and other government rules and restrictions introduced since 1937-41 and by nothing else. Fourth, the C/D ratio that prevailed in 1937-41 is considered normal, implying that there was no underground economy in that period. Therefore this ratio would have prevailed in 1976 (or 1979) had it not been for the change in the level of taxation and in the rules and restrictions. As it was, the C/D ratio in 1976 was much higher than in 1937-41. Assuming that the level of demand deposits in 1976 was normal, he calculates the "extra" currency attributable to the existence of the underground economy. This extra currency is then multiplied by the ratio of (legal) GNP to legal (i.e., excluding the "extra" currency) money. The result is assumed to reflect income rather than gross expenditure and is thus taken as an estimate of the underground economy.

Feige's estimate: In a recent article Professor Edgar L. Feige has presented yet another, and somewhat more extraordinary, estimate of the underground or, as he calls it, the irregular economy. Feige's method is derived from Irving Fisher's quantity theory of money. That theory can be written as MV = PT where M is money (both checks and currency), V is the transaction (and not income) velocity of money, P is an average price level for all goods exchanged (and not just newly created ones), and T is an index of the physical volume of all transactions. Thus if M (checks plus currency) and V (the transaction velocity of currency and checks) are known, MV, and consequently PT, can be calculated. If the ratio

of PT (total dollar transactions) to nominal GNP is known and is assumed constant, nominal GNP can be estimated for any year once PT is known. In the absence of an underground economy, the nominal GNP so derived should, ceteris paribus, be equal to the GNP that is measured in the national accounts. Feige assumes that in 1939 there was no underground economy so that the ratio of total dollar transactions to nominal GNP that he derives for that year — equal to 10.3 — was normal. He then calculates PT for 1976 and 1978. Dividing the results by the 1939 ratio, he derives estimations of nominal GNP for 1976 and 1978. The difference between these indirectly obtained GNPs and the official ones in the national accounts are assumed to measure the underground economy. The underground economy is estimated to be between \$ 225.5 billion and \$369.1 billion in 1976 and between \$541.7 billion and \$ 704.4 billion in 1978. In relation to GNP these estimates range from 13.2 per cent to 21.7 per cent in 1976 and from 25.5 per cent to 33.1 per cent in 1978.

III. A New Approach

As indicated above, Peter Gutmann was the first to attempt to measure the underground economy through the use of currency statistics. Unfortunately, he relied on an assumption that surely cannot be accepted, namely that the ratio of currency to demand deposits is influenced only by changes in taxes and government restrictions and by nothing else. In this section an attempt is made to derive a more firmly based estimate of the "underground economy" by making the demand for currency explicitly a function of several variables including the level of taxes. By measuring the sensitivity of currency demand to taxes, an estimate of currency held for illegal purposes is derived. From this "illegal currency", estimates of the underground economy and of income tax evasion are also derived. Therefore, the method proposed depends basically on the derivation of a good demand-for-currency equation.

While much effort has gone into the analysis of the factors that determine the demand for money (M1 and/or M2), the literature dealing with the demand for currency is meager. Of the few

³ EDGAR L. FEIGE, "How Big Is the Irregular Economy", Challenge, November-December 1979, pp. 5-13.

studies that deal with the demand for currency, Cagan's remains. after two decades, the most comprehensive. In that study, he carefully identified various factors that might affect the ratio of currency to money, the latter defined as incorporating time deposits (M2). His comprehensive list included: (a) the opportunity cost of holding currency (b) expected real income per capita; (c) volume of retail trade; (d) volume of travel per capita; (e) degree of urbanization; and, finally, (f) the level of income taxation. He proceeded then to a statistical, time-series analysis covering the period 1919-55. In that analysis, the ratio of currency to M2 was made to depend on: (a) interest paid on time, and until the Great Depression years (1929-33), on demand deposits; (b) "expected", or "permanent", per capita real income; and (c) the ratio of personal income taxes to personal income. He eliminated retail trade, travel, and degree of urbanization from his independent variables either because of lack of data, or, more importantly, because of expected multicollinearity with income.5

This paper proposes a somewhat modified version of Cagan's original model, but the relationship between tax evasion and currency use is analyzed much more closely than by Cagan. The data used for the statistical tests are also discussed. The statistical analysis covers the period 1929-76, the longest for which the needed data are available for the United States at this time.6

The dependent variable can be taken to be either the ratio of currency to demand deposits or the ratio of currency to M2. There are reasons for preferring either one of these two ratios. For much of the period covered by the statistical analysis - 1929-76 - commercial banks did not pay interest on demand deposits; therefore. the C/D ratio allows the elimination of the rate of interest from

A PHILLIP CAGAN, The Demand for Currency Relative to Total Money Supply, National Bureau of Economic Research, Inc., Occasional Paper 62 (New York, 1958). Published also in Journal of Political Economy, August 1958.

the explanatory variables.7 As it is difficult to generate a good series on a rate of interest that is relevant to the choice between currency and M2 (the latter being made up of currency, demand deposits, and time deposits), this is an important advantage. On the other hand to the extent that, during the period, checking deposits may have been replaced by time deposits, one could get a fall in the C/D ratio caused by a decline in D rather than by an increase in C. Several writers have made this assertion. For this reason, and also because of a better statistical fit, the C/M2 currency ratio has been preferred.

It is assumed that this currency ratio is affected by what could be called legal as well as illegal factors. The legal factors are distinguished between those of a long-run, or structural, nature and those of a short-run, and/or cyclical, nature.9 The illegal factors are distinguished between those that are tax related and those that are not.

Legal factors.

Among the long-run, or structural, legal factors that might affect the currency ratio, one must include: (a) the introduction and increasing use of credit cards; (b) volume of travel per capita; (c) degree of urbanization; and (d) the spreading of branches of commercial banks throughout the country. As it would be difficult or impossible to obtain time series for specific variables that could measure each of these structural factors, I shall use, as did Cagan, per capita income (in both its measured and its permanent version) as a proxy for these developments. It is hypothesized that increases in real per capita income would bring about falls in the currency ratio.

Among the short-run, and/or cyclical, legal factors, there are (a) the composition of income and (b) the relative cost of holding currency vis-à-vis demand deposits.

⁵ A similar analysis using the same variables was carried out for Canada by George Macestch, "Demand for Currency and Taxation in Canada", The Southern Economic Journal, Vol. 29 (July 1962), pp. 33-38.

6 In a recently published paper dealing with the demand for currency for a more recent period, Gillian Garcia and Simon Pak follow a totally different approach, which gives no role to the tax factor. See their "The Ratio of Currency to Demand Deposits in the United States", The Journal of Finance, Vol. 34 (June 1979), pp. 703-15. In spite of its title, this article deals separately with the demand for currency and that for deposits.

⁷ Demand deposits received significant interest payments up to the depression years. After 1933 these payments were reduced to insignificant levels.

⁸ In an earlier version of this paper the C/D ratio was taken as the dependent

⁹ As this paper concentrates on yearly variations, seasonal factors that affect the demand for money within the year, such as Christmas shopping are not

Composition of income: Casual observation indicates that in the United States, while interest, dividends, and rents are almost always received in the form of checks, wages and salaries are paid partly by check and partly by currency. Therefore, changes in the composition of income received could induce changes in the currency ratio. More specifically, while salaried employees are overwhelmingly paid by check, nonsalaried workers are often paid in currency, especially in such industries as construction, agriculture, and mining. If the total compensation of employees could be broken down between wages and salaries, one could use, as one of the explanatory variables in the statistical analysis, the share of wages (i.e., excluding salaries) in personal income. Unfortunately, such a breakdown is not available. Thus, the ratio of total wages and salaries in personal income is used. It is hypothesized that as this ratio increases, so will the currency ratio.

Relative costs: The ratio of currency to M2 can be expected to be influenced by relative cost considerations. For much of the period analyzed, there have been some explicit costs in the form of service charges associated with the holding of demand deposits. Dividing these charges by total deposits provides an estimate of this relative cost. Unfortunately, this series suffers from shortcomings so serious as to make it worthless. First, the service charges are available only for member banks of the Federal Reserve System and, thus, are not representative of the whole commercial banking system. Second, as mentioned earlier, up to the depression years, demand deposits received significant interest payments that exceeded the service charges. Third, the series would ignore costs associated with losses of deposits that were due to failures of commercial banks. Finally, an ideal series should account not only for costs related to holding of deposits but also for costs related to the holding of currency owing to losses, fires, burglaries, robberies, etc. To the extent that these costs have changed over time, they could also affect the currency ratio. For these reasons, I shall not be able to account for these relative costs in the statistical analysis, although I do recognize that they may play a significant role.10

As the dependent variable is the ratio of currency to M2, it is necessary to add to the independent variables the rate of interest, as suggested by Cagan, as a measure of the opportunity cost of holding currency. The rate on time deposits is the one used.

Illegal factors

Even if taxes did not exist, the currency ratio would be affected by some illegal or criminal activities. For example, activities related to gambling (both legal and illegal), smuggling, narcotics distribution, moonshining, etc., are almost always carried out through the use of currency, so that an increase in these activities would increase the currency ratio. Owing to lack of suitable data, I shall not be able to take into account the effects of these activities. I shall, therefore, concentrate on the effect of tax evasion on the currency ratio. But, of course, to the extent that the incomes associated with these activities are not reported to the tax authorities they also contribute to tax evasion.

There is now an extensive literature that deals with the factors that determine tax evasion, and particularly income tax evasion. These factors include, inter alia: (1) the perceived fairness of the tax laws; (2) the attitude of taxpayers vis-à-vis their government; (3) their basic religious and cultural characteristics; (4) the severity of the penalties imposed on the tax evaders that are apprehended; (5) the facility with which taxes can be evaded; and, finally, (6) the monetary rewards to the taxpayers associated with *not* paying taxes.

The first four of these factors either are not measurable or have measurements that are available only for a short period. Either for these reasons or because it is assumed that they have remained relatively unchanged over the period covered by this study, they are ignored. On the other hand, as is argued later, the last two

¹⁰ The series obtained by dividing the service charges by total demand deposits for member banks was tried in the initial statistical analysis. Both because of its poor performance and because of its shortcomings, it was subsequently dropped.

Evasion", Journal of Public Economics, Vol. 2 (July 1973), pp. 265-70; MICHAEL G. Allingham and Agnar Sandmo, "Income Tax Evasion: A Theoretical Analysis", Journal of Public Economics, Vol. I (November 1972), pp. 323-38; RICHARD D. SCHWARTZ and Sonya Orleans, "On Legal Sanctions", The University of Chicago Law Review, Vol. 34 (1967), pp. 274-300; Internal Revenue Service, "Factors Affecting Taxpayer Compliance with Federal Tax Laws" (mimeo, January 1970).

of these factors could, in theory at least, be measured and are likely to have changed over the period. Suppose that there was a variable, T*, that provided a perfect measurement, over the period, of the taxpayers' rewards for not paying taxes, and another variable, K*, that provided an ideal measure of the facility of evading taxes. Then one could write that

$$E = f(T^*, K^*, U)$$
 [1]

where E is a macromeasure of tax evasion and U is a catch-all variable for all the random and/or nonmeasurable factors. In this expression, one would expect that increases in T* and K* would lead to increases in E. If one assumed further that tax evasion leads directly to greater use of currency, then one could introduce T* and K* among the explanatory variables of the currency ratio. However, as is shown later, the relationship between E and the currency ratio is not as simple or straightforward as it is often assumed to be, so that this relationship needs to be analyzed carefully.

Tax evasion and the demand for currency: In the United States, payors of wages and salaries, interests, and dividends have a legal obligation to report to the authorities the making of those payments. Furthermore, employers must also withhold estimated income taxes for wages and salaries and must transfer them to the tax authorities. For these particular incomes, although collusion between payors and payees aimed at reducing tax liability is of course possible, tax evasion, to the extent that it exists, is generally an activity that concerns mainly the payee and does not involve the payor. The payee simply underreports this income. This is mainly true for interests and dividends as, as already said, the payas-you-earn (Paye) system prevents this from happening for wages and salaries. Thus, tax evasion neither necessarily leads to greater

currency use nor, incidentally, does it bias the national accounts data, as these are based largely on the reporting by the payors.¹⁴

For types of income other than those mentioned in the previous paragraph, there is no reporting obligation on the part of the payors. Furthermore, for independent contractors engaged in professional and business activities, "incomes" often are not received from others but are directly created, as differences between gross receipts and allowable expenses, in the process of rendering services or selling goods. Especially for these activities, collusion between the providers and receivers of services, aimed at eyading the payment of taxes, is not only possible but, often, because of the nature of the transaction, easy. 15 Such a collusion benefits both parties: it benefits the purchaser through a reduction in the cost of the service; and it benefits the provider through a reduction in the income that he declares. The service is rendered for a lower price, provided that the payment is received in currency and is thus difficult to trace. The purchasers of these services will face two alternative supply curves: a lower one if the payment is in currency, and a steeper one if the payment is made by check. Thus, high income taxes are likely to bring about a black market for services in the same way that price controls and rationing bring about a black market for goods.16

This black market (a) will be associated with a greater use of currency, and (b) will distort the national accounts data as, for certain types of income, these will be based on the (understated)

¹² In such a case, payors may underreport actual payments to the authorities.
13 This discussion is relevant to the United States. For some other countries it has been reported that attempts to evade high social security taxes, which fall on both employers and employees, have actually resulted in collusion between them. Equally important, where lending activities do not involve the banking system but are outside organized money markets, collusion between lenders and borrowers aimed at evading taxes may lead to greater currency use and thus may bias the national accounts data.

¹⁴ From this it should not be concluded that there is no evasion vis-à-vis these incomes—only that this evasion is not necessarily associated with the use of currency. For 1976, the latest year for which this information is available, the U.S. Department of Commerce (Bureau of Economic Analysis) estimates of adjusted monetary personal interest income and of adjusted personal income dividends were, respectively, \$70.8 billion and \$30.2 billion. By contrast, Internal Revenue Service returns showed \$48.4 billion and \$24.5 billion, respectively. These large "unexplained differences" are, of course, not necessarily due to tax evasion.

¹⁵ Unpublished estimates by the Internal Revenue Service indicate that at most only 74 per cent of the total income received by independent contractors is reported. For additional evidence on tax evasion by occupation, see also Comptroller General of the United States, Who Is not Filing Income Tax Returns? Report to the Congress of the United States (Washington, General Accounting Office, July 11, 1979) Appendix III, pp. 103-12.

¹⁶ Of course, high consumption taxes (including those on imports) may also stimulate a black market for goods.

In some cases, collusion is not explicit in the sense that both parties know why the payment is made in currency but neither has specifically asked that it be made in this form.

incomes reported to the tax authorities. This means that one must be very cautious in using national accounts data to measure tax evasion. For this reason, one cannot take as a measure of K* in equation [1] — i.e., as a measure of the facility of evading taxes — the ratio of presumably more easily evadable incomes, such as rents and proprietors' incomes, to national income, as the numerator of that ratio (and, to a lesser extent, the denominator) has been distorted by tax evasion. In conclusion, although it is recognized that the facility with which taxes can be evaded is likely to affect the evasion rate, it is not clear how this necessarily leads to greater use of currency; furthermore, the data that conceivably could be used to introduce this variable would be distorted by evasion itself. For this reason, in these statistical tests I ignore this variable and concentrate instead on the variable that would measure the rewards of taxpayers associated with *not* paying taxes.

One objective measure of the benefit, or reward, of not paying income taxes is provided by the level of the legal tax liability. When that liability is low, the reward from tax evasion is also low. In such cases, the cost of being an honest taxpayer is not high. However, as the legal effective tax rates rise, so do the benefits from tax evasion. Honesty becomes a more expensive virtue. One would thus expect that, *ceteris paribus*, che temptation to evade taxes would rise with the rise in the tax level. But how does one measure such a level?

One possibility, used by Cagan and Macesich, is to take the ratio of total income tax revenue to personal income. The variable thus obtained suffers from three shortcomings: first, the numerator of the ratio uses actual tax revenue rather than potential (i.e. without evasion) revenue. In other words, the ratio may be reduced by the existence of evasion. Second, the denominator itself may have been affected by underreporting of some incomes. Third, the ratio may remain unchanged even when the rate structure is changing. This last shortcoming is possibly the most serious, as it is the marginal tax

18 This, for example, may happen when the exemptions and the rates are increased at the same time.

rate on a taxpayer's income — rather than the average rate — that is more likely to determine whether he evades the tax on the marginal dollar. If the average tax rate is also high, there could be an income effect that might reinforce the taxpayer's propensity to evade the tax. A variable of this type will be used as one of three alternative ways of measuring the monetary rewards from income tax evasion. However, as transfer payments, which are largely nontaxable, have increased tremendously over the period 1929-76, personal income net of transfer payments is taken as the denominator. This variable is referred to as T1.

A second possibility is to use the top-bracket statutory tax rate for each year as a proxy for the effective marginal rates that affect taxpayers. The major shortcoming of this variable is obvious: a relatively small and changing proportion of all taxpayers is subject to this rate. However, to the extent that often, although not always, the whole rate structure moves together in the same direction, the top-bracket rate will have informational value for the whole rate structure and can thus serve as a proxy for the entire structure. This variable is referred to as T2.

A third possibility is presented by the availability of an effective weighted-average tax rate on interest incomes that was constructed for other purposes. This series is likely to provide yearly rates that may be closer to some modal average tax rates experienced by the taxpayers than would be the previous two alternatives. Therefore, in spite of its limited nature, this series is likely to better capture changes in the level of income taxes over the period. It is indicated by T3.

Summarizing, let:

C/M2 = ratio of currency holdings to money defined as M2;

T1 = ratio of personal income taxes to personal income net of transfers;

T2 = top-bracket statutory tax rate;

T3 = weighted-average tax rate on interest income;

W = share of wages and salaries in personal income;

¹⁷ As it was done, for example, in VITO TANZI, "Income Tax Treatment of Different Kinds of Income", Ch. 5 in The Individual Income Tax and Economic Growth: An International Comparison — France, Germany, Italy, Japan, United Kingdom, United States (The Johns Hopkins Press, 1969),pp. 50-76.

¹⁸ For the period up to 1958, this series was prepared by Colin Wright for his study, "Saving and the Rate of Interest", The Taxation of Income from Capital, ed. by Arnold C. Harberger and Martin J. Bailey (The Brookings Institution, 1969), pp. 275-300. This series was extended, by Tanzi, to 1976.

R = interest rate on time deposits;

 Y^{m} per capita income from national accounts (i.e., " measured " income):

Yp = real per capita income as estimated by Friedman and Schwartz (i.e., "permanent" income).

The truncated version of the model that is subjected to empirical tests can be written as:

$$(C/M2)_t = \alpha_0 + \alpha_1 T_t + \alpha_2 W_t + \alpha_3 Y_t + \alpha_4 R_t + \upsilon_t$$

[2]

where t refers to the specific year, v is an error term, and T and Y can take, respectively, the values of T1, T2, and T3, and Y^m and Y^p. From the previous discussion, one would expect a negative sign for the Y and R variables and positive signs for the T and W variables. This model is tested for the 1929-76 period. Assuming some inertia in the adjustment of the currency ratio to changes in the independent variables, equation [2] is also tested with the addition of the lagged dependent variable. The tests are conducted in log form.

IV. Empirical Results

The regression equations for equation [2] are shown in Table 2. The results, which are remarkably good, can be summarized briefly The adjusted R² are very high exceeding 0.94 for all equations and reaching, for some, 0.97. This indicates that the model is capable of explaining most of the variance in the dependent variable over the period. In most cases, the value of the Durbin-Watson statistics and, for the equation with the lagged variable, the H statistics, are at satisfactory levels.20 However, the equations that use T2 — i.e., the top-bracket statutory tax rates — are definitely the worst. This is not surprising, as the top bracket rates were often left unchanged for several years. The tax variable is highly significant in all cases, but especially when T3 is used. Furthermore, it has the right (positive) sign, indicating that an increase in the tax rate, presumably through an evasion effect, brings about a grea-

REGRESSION EQUATIONS FOR EQUATION

+0.1977 (6.11) *** +0.2819 (4.44) *** +0.3395 +0.1339 (3.90) *** +0.1812 (5.95) *** +0.2618 (6.62) *** +0.2618 (6.62) *** +0.3466 +0.1121 (3.92) *** +0.2626 (6.35) ***		D-W or H Statistic	1.819	1.432	1.793	0.791	1.689	0.932	1.857	1.392	1.822"	.999.0	1.620	0.916°
(6.11)*** +0.2819		<u>R</u> 2	0.955	0.942	0.968	926.0	0.955	0.971	0.952	0.941	996.0	0.952	0.954	0.970
(6.11) *** +0.2819		(C/M2)-1				+0.2103	(2.40) ° +0.3176	+0.2042	**(86.2)			+0.2237	+0.3229	+0.1987 (2.85) ***
(6.11)*** T2 T3 R W +0.1977		Y2					,		-0.3278	+0.2370	-0.1280 0.1280	-0.1779	+0.13%	-0.1142 (1.50)
(6.11) *** +0.2819		īĀ	-0.3121 (2.75) **	+0.1786	-0.0849	-0.2205	+0.0786	(1.01) 0.0955	(1.64) ,					
+0.1977 (6.11) *** +0.2819 (4.44) *** +0.3395 +0.1339 (3.90) *** +0.2164 (3.91) *** +0.2618 (6.62) *** +0.2628 +0.2628 +0.2628 +0.2628 +0.2628 +0.2628 +0.2628 +0.2628	ormulation)	W	+1.9687	+1.6685	+1.7059	+1.8102	(6.06) ** +1.5135 (4.87) **	+1.5989	(6.35) ** +2.1441 (2.70) **	+1.6013	+1.7141	+1.9452	+1.5129	+1.6517
+0.1977 (6.11)*** +0.2819 (3.90)*** +0.2164 (3.91)** +0.1812 (5.95)** +0.2164 (4.52)*** +0.1121 (3.27)** +0.2176	(Log-linear 1	æ	-0.2434 (6.41) ***	$-0.2/51$ $(6.09)^{**}$	-0.2181	-0.1999	-0.1758	-0.1715	-0.2470 -0.2470	-0.2835	-0.2082 -0.2082 (6.29) **	-0.2114 (5.41) **	-0.1878 -4.1578	-0.1708 (5.10) **
+0.1977 (6.11) *** (6.11) *** +0.1339 (3.90) *** +0.1812 (5.95) ***		Т3			+0.3395	/-		+0.2618	(20.0)		+0.3466	(*)		+0.2626 (6.36)**
		T2		+0.2819			+0.2164	0.31)	·	+0.2873			+0.2176	
* * * * * * * * * * * * * * *		Ţ	+0.1977 (6.11) **			+0.1339	2.06.0		+0.1812	(in)		+0.1121	}	
Constant -5.3938 (3.86)*** -5.6312 (3.60)** -5.3751 (4.54)*** -5.3751 (4.54)** -5.3751 (4.54)** -5.4079 (4.78)** -5.4079 (4.78)** -5.4079 (4.78)** -5.4079 (4.78)** -5.4079 (4.78)** -5.4079 (4.78)** -5.4079 (4.78)**		Constant	-5.3938 (3.86) **	(3.60) **	-5.3751 (4.54) **	-5.3073	-5.4539	-5.2163	-6.1362 -6.1362 (4.51) **	-5.3963 (3.55) **	-5.4079 (4.78) **	-5.9482 (4.81) **	-5.5183 (4.58) **	-5.4231 (5.50) **
E E E E E E E E Eductions		Equations	[a]		<u>ত</u>	豆	<u>e</u>	Œ	[8]	(T)	Ξ	Ξ	[¥]	Ē

Notes: a = H statistic; b = Durbin-Watson. Regression equations have been corrected for serial correlation using a first-order Cochrane-Orcun factor. Numbers in parentheses are t statistics. One star indicates significance at 5 per cent and two stars significance at 1 per cent.

²⁰ The equations in Table 2 have been adjusted for serial correlation

FIGURE 1

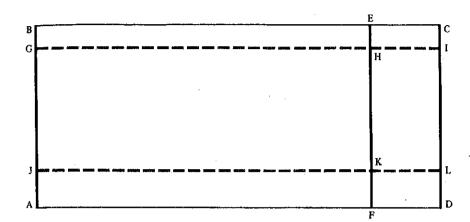
ter relative use of currency. The variable indicating the composition of income, W, is also highly significant in all cases and again has the expected (positive) sign, indicating that a larger share of wages and salaries in national income brings about a greater relative use of currency.²¹ The rate of interest variable is also highly significant and has the expected (i.e., negative) sign.

On the other hand, the per capita income variable, regardless of whether "measured" or "permanent" income is used, is in many cases not significant, casting doubts, at least for the United States, on the hypothesis that economic development reduces the currency ratio.²² However, as by 1929 the United States was already a highly developed country, this particular result might simply indicate that all the development-induced shift between C and M2 had already occurred before 1929, so that no further change should have been expected from this factor.

The table also shows the equations obtained when a stock-adjustment process is assumed. In fact, it is possible that, when the independent variables change, the dependent one does not adjust instantly but does so with some lag. Or, putting it differently, it takes some time for individuals to adjust their currency ratios to their optimal level. The addition of the lagged dependent variable among the exploratory variables has only a marginal effect on the values of the coefficients. As the analysis refers to annual data, this result is also not unexpected, as a year is a long enough period in which to make adjustments.

V. Estimates of the Underground Economy and Tax Evasion

The previous analysis has established that most of the variance in the currency ratio can be explained statistically through the use of a few variables. It has also established a connection between changes in the level of income taxes and changes in the C/M2



Total gross national product (GNP)

BEFA = Measured (or legal) GNP

ECDF = Nonmeasured (or underground) GNP

JKFA = Taxes paid (on measured GNP)

KLDF = Taxes evaded

BCIG = Total money in circulation

BEHG = Legal money used

ECIH = Money used for tax-evading activities

ratio. This connection can be attributed to the existence of a taxinduced "underground" o "subterranean" economy in which transactions are carried out mainly through the use of currency; the incomes thus generated are not reported and escape taxation.

The previous analysis can be used to attempt an estimation of tax-induced currency holdings and, subsequently, of the total values of tax-evading transactions and tax evasion. Figure 1 outlines graphically the methodology that will be followed and the key assumptions on which the estimations will be based. The total rectangle — BCDA — includes the whole economy, whether measured or not. This economy is assumed to be made up of two parts — a legal part (BEFA) equal to the measured or official GNP, and an unmeasured or underground part (ECDF) connected with tax-evading activities. The area BCIG measures total money in circulation needed for the functioning of the whole economy. As this money is needed for transactions purposes, it is identified with total currency plus demand deposits in circulation. BEHG measures the money

²¹ No significant change is observed when national income is replaced by personal income.

²² The equations using permanent income cover the years 1929-75, rather than 1929-76.

needed for transactions in the legal economy, while ECIH measures the money needed for tax-evading transactions (i.e., in the underground economy). BEHG is referred to as legal money, and ECIH as illegal money. The bottom rectangle (JLDA) measures total income taxes that would be paid if the whole economy were legal and thus taxable. JKFA measures the taxes actually paid, while KLDF measures the taxes evaded.

If one could estimate the amount of money presumed to be used for illegal activities (i.e., ECIH), that amount could be multiplied by the ratio of measured GNP to legal money to get an approximation of the underground economy. This would imply that the income velocity of money in the underground economy is assumed to be the same as in the legal economy. Once the underground economy has been measured, one could then estimate tax evasion by multiplying the underground economy by the ratio of taxes actually paid to measured (or legal) GNP.²³

For 1976 actual measurements are available for currency, demand deposits, and M2; therefore C/M2 can be calculated. From the regression equations in Table 2, one can calculate the predicted level of the currency ratio $\left(\frac{\hat{C}}{M2}\right)$ and, therefore, the predicted level of currency holdings \hat{C} , given the actual 1976 figure for M2. The procedure used to obtain \hat{C} from the equations in Table 2 is outlined below.

Let the dependent variable in the regression equations of Table 2 be represented by Z. Therefore:

$$Z_{t} = \ln \left(\frac{C}{M2}\right) = \ln C_{t} - \ln M2_{t}$$
 [3]

Rewriting this equation in terms of lnCt, we get:

$$\ln \hat{C}_t = Z_t + \ln M 2_t$$
 [4]

where the hat on the C indicates that this is the value predicted from the regression equation. Solving this equation, we get:

$$\hat{C}_t = \exp(Z_t + \ln M Z_t)$$
 [5]

This is the value of currency at time t predicted by the regression equations in Table 2. The values of C obtained from

equations [c] and [f] in Table 2 are shown in Table 3, column (3). These two equations were chosen because, statistically, they were the best. Column (6) in Table 3 shows the differences between the actual value for currency in 1976 and the predicted values. These differences are quite small reflecting the equations' substantial explanatory power.

With the aid of Table 3 we can now proceed to the estimation of the underground economy. We can follow two alternative and conceptually different ways depending on what we want to measure. First, we might want to measure the size of the underground economy brought about by the *increase* in taxes over the 1929-76 period. In other words, we might wish to answer the question: what underground activities were induced by the increase in taxes between 1929 and 1976? The second alternative does not concentrate on changes over time but attempts a measure of the total underground economy associated with the 1976 level of taxation. This alternative approach obviously gives a somewhat larger estimate.

TABLE 3
ACTUAL AND PREDICTED VALUES OF CURRENCY HOLDINGS
IN THE UNITED STATES IN 1976
(in billions of \$)

1		Curre	•				
Equation Used	Actual Predicted v		Predicted with lowest taxes over period Predicted with zero taxes		• Differences		
	С	ĉ	ĉ	â	c–ĉ	ĉ−Ĉ	Ĉ–Ĉ
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[c]	77.8	78.3	63.8	46.8	-0.5	14.5	31.5
[f]	77.8	78.0	68.1	55.5	-0.2	9.9	22.5

Source: See text.

I personally prefer the first approach because it recognizes that as long as there are taxes (and other restrictions) there will be some underground activity. As it is unrealistic to conceive of an economy without taxes (and restrictions), it does not seem very productive to attempt to measure all the underground activities but it seems preferable to concentrate on changes over relevant periods. Nevertheless, I shall provide estimates for both alternatives.

²³ This assumes that the effective average tax rate is the same in the legal and in the underground economy.

To obtain the predicted values of currency for 1976 shown in column (3) of Table 3, the regression equations estimated in Table 2 were solved for the 1976 values of the independent variables. These equations can alternatively be solved in the same way except that (a) for the first alternative mentioned above, we assume that the tax variable, instead of having the 1976 value, has the lowest value over the period; (b) for the second alternative we assume that the value of the tax variable falls to zero. In both cases it is assumed that the coefficient of the other variables do not change. For the first alternative we refer to the new predicted value of currency holdings as \hat{C} , while for the second alternative we refer to it as \hat{C} .

More specifically, let Z* indicate the value of the dependent variable in the regression equations obtained when, *ceteris paribus*, the tax variable is assumed to be at its lowest level over the period. Then Ĉ is:

$$\hat{C}_t = \exp\left((Z_t * + \ln M Z_t)\right)$$
 [6]

The predicted values for \hat{C} are shown in Table 3, column (4); the differences between \hat{C} and \hat{C} are shown in the same table in column (7).

There is a technical difficulty in trying to derive an estimate of $\hat{\mathbb{C}}$ simply by setting the tax variable equal to zero and solving the equation as it was done for $\hat{\mathbb{C}}$ above: the double-log formulation, that proved best in the specification of the regression equations, does not allow the setting of taxes equal to zero as the logarithm of zero is minus infinity. This problem can be solved by replacing, in the estimation of the regression equations, the tax variable T by a variable (T+K), where the K is a constant calculated through a search procedure. Therefore a new set of regression equations was estimated.²⁴ From these new equations $\hat{\mathbb{C}}$ was estimated using the same procedure as for $\hat{\mathbb{C}}$. The results are shown in column (5) of

Table 3. The differences between \hat{C} and $\hat{\hat{C}}$ are shown in column (8).25

Table 3 (column (7)) indicates that between \$ 10 billion and \$ 14 billion of currency holdings can be attributed to the *change* in the tax factor between its lowest level in the 1929-76 period and that reached in 1976. It also indicates (column (8)) that between \$ 22 billion and \$ 31 billion could possibly be attributed to the *existence* of taxes although, as argued above, the basis for this further assertion is somewhat weaker. In any case these are our estimates of the illegal money that can be assumed to be fueling the underground economy. If taxes had remained at their lowest level, or if there had been no taxes at all, currency holdings would be correspondingly lower.

If it is assumed that the relationship between tax-induced currency holdings (i.e., between illegal money) and the underground economy is the same as that between legal money holdings used for transactions (including currency and demand deposits) and legal or measured GNP, one can then multiply illegal money by the income velocity of legal money ²⁶ to obtain an estimate of the underground economy. ²⁷ Once the underground economy has been estimated, one can proceed to estimate tax evasion.

Table 4 provides the estimates for the underground economy and income tax evasion. Columns (1) and (2) allocate total money for 1976, equal to \$ 304.3 billion, between legal and illegal money. Column (3) calculates the income velocity of legal money—equal to GNP divided by legal money. Column (4) calculates the underground economy by multiplying illegal money by the income velocity

²⁴ These equations have not been included to economize on space. They are ratically identical to those in Table 2.

²⁵ Another difficulty with this second alternative is that it requires solving the estimated regression equations for a value of the tax variable outside the range used for the estimation.

The income velocity of legal money is obtained by dividing GNP by legal money.

The assumption that the income velocity of money is the same in the underground and in the legal economy is clearly a crucial one. It is the result of agnosticism. The author is unable to take a position between those who would argue that the velocity of money in the underground economy must be lower than in the legal economy, and those who would argue the contrary. The first alternative was backed by Nancy H. Teeters in a recent statement before a congressional committee. See Federal Reserve Bulletin (September 1979), pp. 742-43. The second alternative is backed by Professor Edoar Frigo in "How Big Is the Irregular Economy?" Challenge, November-December 1979, pp. 5-13.

Table 4
ESTIMATES OF UNDERGROUND ECONOMY AND TAX EVASION, 1976
(in billions of \$)

			Income	Underground	Economy	Tax
Equation Used	Illegal Legal Velocity Money Money of Legal Money (1) (2) (3)		Money Money of Legal \$ Money		% of GNP (5)	Evasion \$ (6)
		V	Vith lowest	taxes		
[c]	14.5	289.8	5.94	86.1	5.1	6.7
[f]	9.9	294.4	5.85	57.9	3.4	4.5
			With zero to	axes		
[c]	31.5	272.8	6.31	198.8	11.7	15.4
[f]	22.5	281.8	6.11	137.5	8.1	10.6

Source: See text.

of legal money. It must be recalled that a key assumption here is that the velocity of money for currency in the underground economy is the same as that for M1 in the legal economy. This is clearly a debatable assumption. Column (5) shows the underground economy as a proportion of GNP while column (6) shows the estimates for tax evasion. Tax evasion in calculated by assuming that the average tax liability for underground incomes would be the same as for legal incomes. This is also a key assumption.

The estimates indicate that the increase in taxes over the period brought about an underground economy of between 3.4 and 5.1 per cent of GNP and a tax evasion of \$4.5 to \$6.7 billion. On the basis of a somewhat more debatable procedure, the table indicates also that the existence of taxes (as distinguished from their increase) may have generated a total underground economy of between 8.1 and 11.7 per cent of GNP and a tax evasion of between \$10.6 and \$15.4 billion.

V. Implications and Conclusions

The existence of a substantial underground economy would have important and, at times, disturbing implications for (a) the measurement of macroeconomic variables, (b) the pursuit of economic policy, and (c) the efficient functioning of the economy.

Employment statistics are related to workers engaged in economic activities within the legal sector of the economy. To the extent that individuals are counted in the labor force but, being employed in underground activities, are not counted among the employed, the unemployment rate could be overstated. The degree of overstatement will, of course, depend on the size of the underground economy. Professors Feige and Gutmann, in independent recent articles, have argued that this overestimation of the unemployment rate is substantial.

It has been argued by some observers (e.g., Professor Feige) that the inflation rate is also overstated as prices in the underground economy are likely to be growing at a lower rate than in the legal economy. In fact part of the alleged shift from legal to underground activities, for those areas where the two economies are in competition, is attributed to this differential in price changes.

Besides the uneployment and the inflation rates, other macro-economic variables which will be distorted are the following: the true rate of growth of the economy will be higher than the measured growth, if the underground economy is growing relatively to measured GNP; the measured size of the public sector, whether measured as a ratio of taxes or expenditure to GNP, will be magnified; statistics on income distribution as well as those on tax incidence will also be distorted.

As economic policy responds to the signals provided by these macro-economic variables, the policies that may be pursued will be distorted so that at times they could do more harm than good. For example, to the extent that the unemployment rate is consistently (and, possibly, increasingly) biased upward, the policymakers will pursue policies that are too expansionary vis-à-vis actual needs and will thus promote inflation. If, as indicated above, the cost of living index is overstated, and if inflationary expectations as well as actual indexing clauses (for wages and salaries, pensions, etc.) are based on the distorted index, the inflationary effects will be compounded. And these effects could be magnified if the monetary authorities, in the pursuit of a monetary rule, based the expansion of the money supply on the rate of change of the consumer price index. In such case the monetary authorities might be induced to accommodate a higher rate of inflation by expanding the money supply at a faster pace than warranted by the true rate of inflation. TABLE 5

UNDERGROUND ECONOMY IN THE UNITED STATES: COMPARISON OF ESTIMATES (in per cent of GNP)

	Feige		Gutmann	Internal Revenue			Tanzi		
	Low	High		Low	High	Low		Hìgh	
1976	13.2	21,7	10	5.9	7.9	3.4 ^a 8.1 ^b	{	5.1 ^a 11.7 ^b	
1978	25.5	33.1	10		,	0.1	,	11,7	

Notes: a = associated with tax increase; b = associated with total taxes.

The underground economy will affect the functioning of the economy in different ways and will probably have negative effects on efficiency. For example, as the total economy (both legal and underground) expands, the need for public services will grow. However, taxes are collected only from its legal part so that the tax level on legal activities will increase. This increase will force more activities underground as the benefits from tax evasion will grow. Untaxed underground activities will compete with taxed, legal ones and will succeed in attracting resources even though these activities may be less productive in a social (rather than a private) sense. This flow of resources from legal, and taxed, activities toward underground, and untaxed, ones can be expected to continue as long as, ceteris paribus, the net-of-tax rate of return, adjusted for the risk of being caught and having to pay penalties, is higher in the underground economy. There will of course be significant welfare losses associated with this transfer. The resulting equilibrium, if reached, will imply an allocation of resources different from the optimum. An application to this process of the methodology used by Harberger in connection with his model of the shifting of the corporate income tax should prove productive.

For all of these reasons, it is important that we improve our knowledge of the underground economy. Unfortunately, the results available so far are widely different as can be seen with the aid of Table 5. The range between the lowest and highest estimates in that table is uncomfortably high. Furthermore, it is not clear whether these four estimates are actually measuring the same thing.

Feige and Gutmann both measure changes in the underground economy since the late thirties — since 1937-41 for Gutmann and

since 1939 for Feige.²⁸ Only if in 1937-41 or in 1939, there was no underground economy, as Feige and Gutmann assume, can their results be assumed to measure the *level* of, rather than the *change* in, the underground economy. But as long as the underground economy results from criminal as well as tax-evading activities surely it must have existed in the thirties. After all, the era of Prohibition extended into that decade and the New Deal had already sharply increased income taxes. What this means is that, if Feige's and Gutmann's methods were right, the present level of the underground economy in the United States would be sharply higher than their already extraordinary estimates.

Washington, D.C.

VITO TANZI

Therefore, methodologically they are measuring the same phenomenon as in our estimates associated with tax increases rather than that measured by the Internal Revenue or our estimates associated with tax level.