Feige, Edgar L.

Working Paper

Sweden's 'Underground Economy'

IUI Working Paper, No. 161

Provided in Cooperation with:
Research Institute of Industrial Economics (IFN), Stockholm


This Version is available at:
http://hdl.handle.net/10419/94848

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
A list of Working Papers on the last pages

No. 161, 1986

SWEDEN'S "UNDERGROUND ECONOMY"

by

Edgar L. Feige

This is a preliminary paper. It is intended for private circulation and should not be quoted or referred to in publications without permission of the author. Comments are welcome.

Sept, 1986
INTRODUCTION

This paper summarizes the results of a research project investigating the size, growth and implications of the "underground economy" in Sweden. The popular term "underground economy" is often used to describe two different, albeit related concepts. When the substantive issue is defined to be non-compliance with tax codes, the term unreported income is used to refer to the difference between the amount of income that ought to be reported to the tax authority under full compliance and the amount actually reported. On the other hand, when the substantive issue is taken to be the quality of aggregate measures of economic activity such as Gross National Product, the term unrecorded income is used to denote the amount of income that is inadvertently omitted from National Income and Product Account (NIPA) measures of aggregate economic activity.1

In the United States, recent discoveries of large amounts of unreported income [IRS; 1983] have necessitated a comprehensive revision of NIPA accounts [BEA; 1985] to more adequately reflect previously unrecorded income. Official Internal Revenue Service estimates of unreported income, derived from intensive audits of tax returns, suggest that unreported income in the U.S. was $283 billion in 1981. This figure amounts to 16% of the reported tax base in that year. Independent analysis of the Internal Revenue Service estimates [Feige; 1986a] suggests that unreported income was closer to 25% of reported income. The most recent NIPA revision incorporates an adjustment to Personal Income for previously unrecorded income amounting to $100 billion, reflecting only official estimates of misreporting in tax source information.

---

1 The exact relationship between unreported and unrecorded income is highly complex. It depends upon legal definitions of tax liabilities; accounting conventions defining the precise components of measures of aggregate economic activity and the particular manner in which NIPA accounts are empirically constructed. To the extent that NIPA measures depend upon data sources that are administratively collected for tax purposes or could be biased as a result of efforts by firms or individuals to falsify information reported to government agencies, an increase in unreported income will also increase unrecorded income.
The ability to revise NIPA estimates to reflect estimates of unreported tax income depends upon the availability of detailed reconciliations of the amount of income reported for tax purposes and the amount of income recorded in NIPA accounts. The Bureau of Economic Analysis [BEA] has developed a historical reconciliation between the tax base (Adjusted Gross Income) and the NIPA measure of Personal Income [Park; 1981]. No similar reconciliation is regularly published for Sweden. As such, it is difficult to specify the exact empirical relationship between unreported and unrecorded income in Sweden. We shall however maintain the distinction between the two income concepts in order to avoid making false comparisons between the results of different methods employed to estimate different income concepts.

The suspicion that the "underground economy" may be an important phenomenon in Sweden is based on the fact that Sweden's high marginal and average tax rates provide strong incentives for non compliance with tax laws. On the other hand, the Swedish authorities have a well developed system for apprehending tax evaders. Moreover, the generous level of social services provided by the government may serve as an incentive to comply with an admittedly burdensome system of taxation and regulation.

The question of the size and growth of non compliance, and its consequent implications for the reliability of the nation's official information system is thus an empirical question that can only be resolved by appeal to relevant evidence. Given the inherent complexity of measuring a phenomenon whose raison d'être is to defy detection, one can at best hope to determine a plausible range of estimates of underground economic activity.

Section I briefly reviews and evaluates the evidence that has been cited as having a direct bearing on the size and growth of Sweden's unreported and unrecorded income. Section II develops a general currency ratio model (GCR) which is employed to obtain time series estimates of unrecorded income. Section III reviews a more extensive attempt to employ a Fisherian transactions framework to estimate the volume of unrecorded transactions and income in Sweden. The final section discusses some of the implications of the results for the Swedish Laffer curve and for fiscal policy.

I. A REVIEW OF EXISTING EVIDENCE

a.) Survey Data
Several Swedish studies of tax evasion and unreported income rely on evidence collected by survey methods. These studies suggest that between 10 - 30% of the population admitted to being involved in some form of tax evasion, however, given the intertemporal inconsistency of the questions asked, it is difficult to establish any evidence of the growth of admitted tax evasion over time. Hansson (1982) employs the survey evidence to make the weak inference that tax evasion does not exceed 1% of GDP. This inference however, takes no account of self selection and reporting biases inherent in survey responses to sensitive questions nor is it based on a representative sample that can support aggregative inferences.

b) National Accounts Discrepancies

A second source of evidence cited by (Hansson; 1981; 1982) for tax evasion is the NIPA discrepancy between the expenditure side estimates of household and personal firm income and the corresponding income side estimates. Figure 1 displays three variants of this discrepancy. The first, [EXP. - INC.] shows the discrepancy between the GDP expenditure side estimates and income side estimates of the household sector. The second related discrepancy [INC. - OUT.] is based on the Income and Outlay household accounts discrepancy, and the third [NA - TA], represents the difference between entrepreneurial income and other income based on the NIPA accounts and the corresponding figures obtained directly from tax assessment statistics. The reported discrepancies range between approximately 3 - 6.5% of GDP. Hansson (1982) reports adjustments and recalculations of the expenditure discrepancy measure suggesting a decade average of 6.4% of GDP. None of the discrepancy measures reveal a pronounced trend.

The intuitive rationale underlying the use of the expenditure - income discrepancy as a measure of tax evasion is that the income side of the NIPA rely heavily on tax source data whereas the expenditure side is independently constructed from household budget studies. These discrepancies will reflect unreported

---


3 The first and second discrepancy estimates have been provided by A. Tengblad [10/17/1980]. With the exception of the years 1977-78, the two discrepancies have been reconciled to one another. The [NA - TA] discrepancy is based on data provided by Per Ericson in private correspondence (8/9/83).
income if and only if:

1) Income side estimates are downward biased to the full extent of unreported income, and

2) Expenditure side estimates fully reflect all income that should be reported to the tax authority.

In fact, neither of these necessary conditions are fulfilled. Although income side estimates are based in part on tax source data, they require considerable adjustments before they can be directly compared with expenditure side estimates. The income side estimates that enter the discrepancy measure are contaminated by errors in these adjustment items. The second condition is even less likely to be fulfilled. Tax evaders would be reluctant to be respondents in budget survey studies. If they do respond, they are unlikely to fully

---

4 The problem arises because NIPA income definitions do not exactly correspond to the assessment base for taxation statistics. Particular adjustments include such items as capital consumption and imputations for owner-occupied housing. [Tengblad; 1981]
report their true expenditures out of fear that honest reporting of expenditures will reveal income sources that have not been disclosed to the tax authority. These behaviors have two implications. First, the discrepancy measures will underestimate the size of tax evasion, and second, an increase in actual evasion will not be reflected as an upward trend in the discrepancy measures since expenditure side responses are likely to be downward biased in proportion to actual evasion. Therefore, the foregoing discrepancy measures simply reflect the final excess of NIPA expenditure estimates over NIPA income estimates of the net operating surplus of the household sector. Since both sides of the NIPA accounts are likely to systematically exclude unreported income, the discrepancy can not be taken as a valid measure of either the size or the growth of tax evasion.

c.) Currency Data

FIGURE 2

PER CAPITA CURRENCY HOLDINGS
NOMINAL AND REAL
Another source of casual evidence that has reinforced the suspicion of growing tax evasion is the surprisingly large amount of currency in circulation in Sweden and the changes in the denomination structure of its currency. According to this view, currency is a superior medium of exchange for the conduct of transactions that firms and individuals wish to hide from the scrutiny of tax or regulatory authorities. The superiority of currency over checks or giro transfers derives from the fact that currency transactions do not leave an audit or paper trail. Figure 2 displays the temporal trend in real and nominal per capita currency holdings in Sweden from 1950 - 1982. By 1982, nominal per capita currency holdings outside of the banking system approached 4250 Kr. and the average Swedish household held approximately 9000 Kr. in currency. These currency holdings are higher than those of any western nation other than Switzerland, whose currency is often used as an international medium of exchange to avoid detection by other government's fiscal authorities.

Sweden’s large and growing nominal currency holdings run counter to economic predictions of currency trends. Economists have anticipated the onset of the "cashless society", induced by technological innovations in bank and post giro institutions. The growth of services offered to customers using check and giro facilities were believed to induce shifts away from currency toward these media of exchange in response to the decline in their relative cost. Moreover, Sweden’s inflation experience imposed an implicit tax on currency holdings that individuals could have avoided by holding liquid assets in interest bearing form.

Equally surprising, is the changing composition of the Swedish currency supply. As displayed in Figure 3, the share of 1000 Kr. notes rose from 13% in 1950 to 43% by 1982. Higher prices will tend to shift the entire denomination structure toward higher denomination notes, however, this explanation can not account for the dramatic increase in the share of the 1000 Kr. denominations. A more plausible hypothesis is that these notes are being used for cash transactions which individuals wish to hide from government scrutiny.

II. A GENERAL CURRENCY RATIO MODEL

If currency is regarded as a superior medium of exchange for transactions that are not reported to tax authorities, and may therefore also avoid inclusion in NIPA income estimates, then it is possible to obtain aggregate time series estimates of unreported and unrecorded income from observations on the temporal
growth in currency holdings. The link between currency and underground income sources however requires
the specification of a model of economic activity that relates currency holdings directly to income.

a.) Model Specification

The specific assumptions required to implement the estimation of unrecorded income by means of a
currency ratio method are clarified by reference to a particular model. The currency ratio model presented
below is sufficiently general to encompass all previous currency ratio methods [Cagan, 1958; Gutmann,
1977; Tanzi; 1982; 1983; Klovland, 1980; 1984] as special cases. Let,

\[ C = \text{Actual currency stock} \]
\[ D = \text{Actual stock of checkable and giro deposits} \]
\[ Y_o = \text{Recorded or official income.} \]
\[ u = \text{subscript to denote the unrecorded sector.} \]
o = subscript to denote the "official" or recorded sector.

\( k_o \) = the ratio of currency to checkable deposits in the recorded sector

\( k_u \) = the ratio of currency to checkable deposits in the unrecorded sector

\( v_o \) = unrecorded sector income velocity

\( v_u \) = recorded sector income velocity

The general currency ratio (GCR) model contains the following specifications:

1. \( C = C_u + C_o \)
2. \( D = D_u + D_o \)
3. \( k_o = \frac{C_o}{D_o} \)
4. \( k_u = \frac{C_u}{D_u} \)
5. \( v_o = \frac{Y_o}{C_o + D_o} \)
6. \( v_u = \frac{Y_u}{C_u + D_u} \)
7. \( \beta = \frac{v_o}{v_u} \)

Equations (1) and (2) decompose the actual stocks of currency and checkable deposits into their

---

5 The appropriate definition of the assets that constitute the money supply must be limited to those assets that function as a final medium of exchange. In Swedish practice, currency, checkable deposits and giro demand deposits in banks and the post giro comprise the final media of exchange. Travelers checks and credit cards are sometimes considered to be media of exchange, however, since the purchase of travelers checks or the settlement of credit card accounts requires the use of currency, giro or checkable deposits, to
unrecorded and recorded components. Equations (3) and (4) are definitions of the terms \( k_o \) and \( k_u \) which can be specified as constants or stable functions of other variables. Similarly, equations (5) and (6) define income velocity in the two sectors. To solve the model for \( Y_u \), we evaluate (6) in terms of the model's observable variables, namely \( C \), \( D \) and \( Y_o \). Repeated substitution and rearrangement of terms yields the general solution for \( Y_u \) as:

\[
Y_u = \frac{1}{\beta} \cdot \frac{(k_u + 1) \cdot (C - k_o D)}{(k_o + 1) \cdot (k_u D - C)}
\]

which expresses unrecorded income as a function of the observable variables \( Y_o \), \( C \) and \( D \) and three parameters or functions: \( \beta \), \( k_u \) and \( k_o \).

The simplest variants of the (GCR) model employ the following restrictive assumptions:

1. That currency is the exclusive medium of exchange for unrecorded transactions \([D_u = 0; k_u \rightarrow \infty]\).
2. That the amount of unrecorded income produced by a dollar of currency transacted in the unrecorded sector is the same as the amount of reported income produced by a dollar of currency transacted in the reported sector. \([\beta = 1]\).

include these assets would amount to double counting. The same is true of time deposits which must first be converted into currency, demand or giro deposits before being used as a medium of exchange. Klovland's (1980) otherwise careful study makes the error of employing a broad definition of the money supply for the denominator of the currency ratio.

If the object of the analysis is to obtain estimates of unreported income, the appropriate income variable is the amount of income actually reported to the tax authority, which in the case of Sweden would be tax assessment income. Since the conceptual reconciliation between reported income for tax purposes and recorded income in NIPA in Sweden is poorly documented, the analysis proceeds on the basis of attempting to estimate unrecorded income. To the extent that unreported income moves pari passu with unrecorded income, the analysis can also be interpreted as a rough estimate of unreported income.

These are the assumptions employed by Cagan (1958) and Gutmann (1977).

Assumptions (1) and (2) are implicit in Klovland's (1980; 1984) approach. Klovland was forced to abandon the currency ratio version of the model because of the difficulty of constructing consistent time series data for Sweden's demand deposit series. The inconsistency problem which arises as a result of the reclassification of Swedish deposit accounts in various periods is addressed in Feige (1985; Appendix B) and the calculations that follow are based on a newly derived consistent series of deposits.
3. That the ratio of currency to checkable deposits remains constant except for changes induced by the growth of unrecorded income. \([k_{ot} = k_o\) for all \(t)\]. Imposing these restrictions on the GCR model, yields a simpler form of equation (8), namely,

\[ Y_u = Y_o \cdot \left\{ \frac{C - k_o D}{(k_o + 1) D} \right\} \]

c.) Econometric Specifications of \(k_o\).

The simplest versions of the (GCR) model assumed that the benchmark estimate of \(k_o\) was a constant, rather than a function of other economic variables. In the spirit of Cagan's (1958) original investigation of the currency ratio, \(k_o\) is specified as a function to be estimated by econometric means. The advantage of this approach is that it takes explicit account of those economic factors that are believed to have effected observed variations in the currency ratio over time.

Figure 4 displays the ratio of currency to bank and post giro demand deposits \([C/B+P]\) as well as the ratio of currency to each of the separate deposit components \([C/B\) and \(C/P]\). The component series, although corrected for classification changes still exhibit some important discontinuities, particularly in 1959-1960 reflecting the introduction of bank giro sight deposit accounts. The overall currency ratio exhibits a gradual secular decline reflecting such financial innovations as the introduction of wage and salary accounts in Sweden. These innovations as well as other economic factors have brought about a decline in the use of currency relative to deposit and giro accounts in a manner consistent with the anticipations of a "cashless society". On the other hand, it is possible that this anticipated relative decline in the use of currency would have proceeded at a much more rapid pace in the absence of incentives to continue to use currency for unrecorded transactions. In short, what is required is the development of the counter factual case which asks, what would have been the pace of currency decline in the absence of unrecorded transactions? The answer to this question can be obtained by decomposing the observed currency ratio into two components. The first component \([k_o]\) reflects the use of currency in the recorded sector, the second, \([C_u/D]\) reflects the use of currency in the unrecorded sector.
Since $k_o$ is itself unobserved, it is necessary to derive an expression for $k_o$ in terms of observed variables. Maintaining the simplifying assumption that demand deposits are never used for the payment of unrecorded incomes, $(D_u \to 0)$, it follows that the observed $(C/D)$ ratio is defined as:

\[
\frac{C}{D} = \frac{C_o}{D_o} + \frac{C_u}{D_o} = k_o + \frac{C_u}{D}
\]

The ratio $k_o$ can be approximated by a function such as:

\[
k_o = f_1(y,r,p,g,\gamma)
\]
12

and;

\[ \frac{C_u}{D} = f_2(r) \]

where, \((y)\) is real per capita recorded income; \((r)\) is the opportunity cost of holding currency relative to an interest bearing deposit, \((p)\) is the rate of inflation reflecting the implicit tax on currency balances, \((g)\) is the government expenditure share of GDP reflecting the preponderant use of demand deposits in government transactions and \((\gamma)\) is a dummy variable reflecting the innovation of bank giro deposits. Finally \([f_2(r)]\) reflects the tax rate \([\tau]\) incentive to use currency as a means of avoiding detection of unrecorded transactions.

Since the left hand side ratios in equations (10) and (11) are unobservable, an estimate of \(k_0\) must be obtained from a regression of the observed C/D ratio as shown in equation (9). Thus:

\[ \frac{C}{D} = f_1(y,r,p,g,\gamma) + f_2(r) \]

Equation (12) reveals that the observed currency ratio is the sum of two functions. Kloveland (1980), following Tanzi (1982) chooses a multiplicative (log linear) functional form to estimate the currency ratio therefore violating the additive specification implied by equation (12). To derive the predicted counterfactual time path of \(k_0\), it is necessary to estimate Equation (12) and then to obtain the dynamic forecast of \(k_0\) after setting \(f_2(r) = \tau_b\) were \(\tau_b\) is the tax rate obtaining in some base period. For the calculations that follow, the base period was chosen to be 1950. The dynamic simulation of \(k_0\) therefore represents the predicted time path of the currency ratio that would have taken place if tax incentives had remained at their 1950 level.\(^9\) The forecast values of \(k_0\) are then substituted into Equation (8a) in order to obtain a

\(^9\) The dynamic simulations are based on an estimating equation employing a lagged endogenous variable; a first order autoregressive disturbance and a single lag for \(\tau\).
Figure (5) presents the estimates of $\frac{Y_u}{Y_o}$ for the period 1955 - 1982 that result from a dynamic simulation of the (GCR) model allowing for a variable $k_o$ and assuming that $\beta = 1$. The results suggest that unrecorded income averaged about 3.5% during the late 1950's, rising to the vicinity of 6% during the 1960's and reaching a peak of 21% by the end of the 1970's. The estimates suggest that unrecorded income actually declined during the early 1980's. Several of the erratic movements in the estimated series may be the result of asset reclassifications affecting the original data series that could not entirely be removed in the attempt to construct consistent time series estimates.

It is also possible to relax the restriction that $\beta = 1$. To the extent that unrecorded income is largely derived from the service sector, it is likely that $\beta < 1$, since the ratio of intermediate to final transactions in the service sector is likely to be lower than in the non service sector. On the other hand, a lower propensity...
to consume unrecorded income would imply that $\beta > 1$. The former assumption produces higher estimates of unrecorded income, whereas the latter assumption reduces estimated unrecorded income.

Figure (5) includes estimates of $Y_u / Y_0$ for alternative values of $\beta$. While the value of $\beta$ affects the estimated size of the unrecorded sector, it leaves invariant the temporal path of $Y_u / Y_0$.

Figure (6) displays the estimates of total GDP obtained by the GCR model with the variable $k_o$ specification [$\beta = 1$]. The estimates imply that per capita unrecorded income amounted to approximately 9000 Kr. in the early 1980's.

III. TRANSACTIONS ESTIMATES OF UNREPORTED INCOME

The foregoing estimates of unreported income are based on a conventional monetary model in which economic activity is specified as depending upon the stock of monetary assets rather than on the monetary transaction flows which actually underlie all economic activity. In a separate monograph [Feige, 1985] the
author has attempted to specify a general macroeconomic framework based on Fisher's (1911) equation of exchange, which permits empirical estimation of both the payment side of Fisher's identity \([MV]\) and the transactions side of the identity \([PT]\). The measurement of total payments \([MV]\) requires independent estimates of the payment velocity of each medium of exchange. The transactions side of the equation of exchange involves the aggregation of all current national accounting frameworks, including NIPA; Input-Output Accounts; Income and Outlay Accounts; Balance of Payment Accounts and Flow of Funds Accounts. The conceptual framework uses Swedish data in a pilot study to demonstrate the feasibility of replacing Keynes's \([Y = C + I + G]\) identity with Fisher's \([MV = PT]\) identity as the basis for macroeconomic accounting and analysis.

FIGURE 7

PERCENT OF TOTAL INCOME UNRECORDED

Independent estimates of aggregate payments in Sweden \([MV]\) are shown to exceed estimates of total
recorded transactions [PT]. The discrepancy between estimated total payments and estimated recorded transactions has grown dramatically over time and this discrepancy is taken as an indicator of unrecorded transactions. Figure (7) displays the percent of total income that is unrecorded as estimated by the transactions method. For comparative purposes, Figure (7) also displays the comparable estimates derived from the currency ratio model. Both estimates suggest a secular increase in unrecorded income with the transactions method producing the higher set of estimates. This result is not surprising since the transactions approach is not subject to two of the restrictive assumptions implicit in the currency ratio approach. The currency ratio estimates limit all transactions in unrecorded income to currency whereas the transactions approach permits both checks and giro deposits to be used for some unrecorded transactions. The second difference in the approaches is that the transactions estimates take account of the different velocities of circulation of each medium of exchange.

The foregoing estimates of unrecorded income suggest that unrecorded income in Sweden grew substantially during the past two decades with its growth subsiding during the early 1980's. On the basis of the available evidence it appears that unrecorded income now amounts to between 12 - 25% of total income in Sweden.

IV. IMPLICATIONS

As part of the research project on Sweden's "underground economy" Feige and McGee (1983) developed a model of Sweden's Laffer curve which could be empirically implemented given knowledge of supply elasticities; the progressivity of the tax system and independent estimates of the unrecorded economy. The model was simulated for 1979 under various assumptions, and produced estimates of the revenue maximizing average tax rate for Sweden. The results of the simulations indicated that under the most plausible assumptions, an unrecorded sector between 10% - 20% of total output would imply a revenue maximizing tax rate for Sweden between .58 and .56. The average effective tax rate for Sweden in 1979 was calculated to be .62, suggesting that Sweden was the only country\(^\text{10}\) which appeared to be on the downward

---

\(^{10}\) The Feige-McGee model was also simulated for the U.K; the Netherlands and the U.S. In each case,
sloping portion of the Laffer Curve. The empirical evidence presented in this report supports the conjecture that Sweden can afford to cut its rate of taxation without fearing a consequent loss of revenues. Indeed, on the basis of the simulation experiments combined with the empirical findings on the size of the unrecorded economy, it appears that a cut in Sweden's tax rates can bring about the salutary consequences of increased output, declining tax evasion and increased revenues for public expenditures.

these countries were found to have effective tax rates well below the revenue maximizing tax rate.
REFERENCES


Klovland, Jan (1980) "In Search of the Hidden Economy: Tax Evasion and the Demand for Currency in Norway and Sweden" Norwegian School of Economics and Business Administration, 18/80, 1098.


