3 The Value-Added Tax — A Review of the Issues

11 Demand for Currency: Is the Underground Economy Undermining Monetary Policy?
The Review is published monthly by the Research Department of the Federal Reserve Bank of St. Louis. Single-copy subscriptions are available to the public free of charge. Mail requests for subscriptions, back issues, or address changes to: Research Department, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, Missouri 63166.

Articles herein may be reprinted provided the source is credited. Please provide the Bank's Research Department with a copy of reprinted material.
The Value-Added Tax—A Review of the Issues

R. W. HAFER and MICHAEL E. TREBING

RECENTLY many groups have called for major tax reform. These groups maintain that the existing U.S. tax system has altered economic incentives and negatively influenced the performance of the economy. One possible solution to this problem is for the United States to adopt a value-added tax (VAT) similar to that used by European governments.1 This article provides a general economic framework for examining the current debate over the introduction of a VAT into the U.S. tax system.

BACKGROUND

Value added is the difference between a firm's receipts from the sale of a product and the payment for the resources or raw materials used in producing it. In other words, value added is equivalent to a firm's payment to the factors of production — land, labor, and capital — in the form of rent, wages, interest, and profits.2 These payments represent the "base" to which a VAT would be applied.

A simplified example of how value-added can be determined is shown in table 1, which illustrates a hypothetical four-step process of converting wheat into a loaf of bread. The miller purchases wheat from the farmer at a price of 10 cents and sells the resulting flour to the baker for 20 cents. Since one-half of the price to the baker represents the farmer's receipts, 10 cents of the miller's receipts represent the value added to the raw wheat during this stage of the production process. After converting the flour into bread, the baker sells the loaf of bread to the retailer for


<table>
<thead>
<tr>
<th>Stage</th>
<th>Receipts</th>
<th>Value-Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Farmer</td>
<td>.10</td>
<td>.10</td>
</tr>
<tr>
<td>2 Miller</td>
<td>.10 + .10</td>
<td>.20 + .10 (=.20 -.10)</td>
</tr>
<tr>
<td>3 Baker</td>
<td>.10 + .10 + .10</td>
<td>.30 + .10 (=.30 -.20)</td>
</tr>
<tr>
<td>4 Retailer</td>
<td>.10 + .10 + .10 + .10 = .40</td>
<td>.10 (=.40 -.30)</td>
</tr>
<tr>
<td>Final sales price to consumer = $.40</td>
<td>Total value added = $.40</td>
<td></td>
</tr>
</tbody>
</table>
30 cents. Since the cost of the flour was 20 cents, the value added at this stage (the baker's) is again 10 cents. Finally, out of the 40-cent retail price for the loaf, the remaining 10 cents represents the retailer's value added.

The usefulness of this example is twofold. First, it illustrates the fact that the final retail price to the consumer is really nothing more than the sum of the value added at each stage of production. Second, it provides a format by which a sales tax can be distinguished from a VAT. A sales tax involves a single payment by the consumer at the retail level and is levied at a constant rate on the purchase price. The VAT, on the other hand, is assessed at each stage of production. If the sales tax rate and the VAT rate are equal, the total tax revenue in the above example is the same irrespective of the collection process.

There are two basic methods of calculating the VAT. In the additive method, the tax base for the VAT is the sum of the firm's payments to the factors used in producing the good. Under this scheme, the VAT rate, say 10 percent, is applied to the firm's costs in terms of wages, interest, rents, and profits. Table 2 illustrates a hypothetical example of how the additive procedure is used to calculate the VAT due.

Table 2 provides the information needed to calculate the VAT payable under the second method, the subtractive procedure. Essentially, this approach makes the base to which the VAT rate is applied equal to the difference between a firm's sales receipts and its cost of production (again, its value added). The actual tax liability using this procedure is identical to that derived using the additive approach.

### The Tax Base and Its Calculation

As noted before, the VAT base is usually calculated as the difference between a firm's sales receipts and its cost of purchases from other firms. A debate exists, however, over whether all of a firm's costs should be included in the calculation of the base or whether special exemptions should be made. Exemptions from the tax base are important because they determine many of the economic effects of a VAT. Of particular concern is the treatment of capital goods. There are three general approaches to the treatment of capital goods: the gross product, the income, and the consumption approaches.

Under the gross product approach, firms are not permitted to deduct the purchase price of capital goods when calculating their VAT base, nor are they allowed to deduct the depreciation of existing capital. The tax base at each stage of production, therefore, includes wages, interest, rent, profits, and depreciation. Since there is no deduction for either capital goods depreciation or their purchase, the VAT applied to all income earned in the production of the firm's output. Thus, in an economy-wide sense, the tax base is associated with the value of gross national product — the total current market value of all goods and services produced or the total income derived from their production.

The income approach, on the other hand, does allow firms to deduct capital depreciation from their tax base. Since firms use up a percentage of their existing capital stock each year, this approach taxes firms on their net instead of gross income. The tax base of the income-type VAT is analogous, therefore, to the net national product of the economy.

---

3The following discussion is found in a number of writings. See, for example, Advisory Commission on Intergovernmental Relations, The Value-Added Tax and Alternative Sources of Federal Revenue (Washington, D.C.: ACIR, 1973); John F. Due, "The Value-Added Tax," The Western Economic Journal (Spring 1965), pp. 165-71; and, Richard W. Lindholm, Value-Added Tax and Other Tax Reforms (Chicago: Nelson-Hall, 1976), pp. 30-35.

4An algebraic expression similar to the national income account identities taught in elementary economic textbooks can aid in distinguishing the income base of each type of VAT discussed. Ignoring government expenditures, taxes, and the foreign sector, the expression

\[ \text{GDP} = C + I \]

shows the equivalence of total output or income earned in the economy and the sum of consumption (C) and investment (I) expenditures. With no VAT exemptions, GDP is the base of the gross income VAT. We define

- \[ I - D = I_n \]
- \[ \text{NNP} = I_n + C \]
- \[ \text{NY} = \text{NNP}, \]

where \( I_n \) is net investment, \( D \) is depreciation, \( \text{NY} \) is national income.

---

Table 2
Calculating VAT Liability Via Additive Procedure

<table>
<thead>
<tr>
<th>Wages</th>
<th>$500,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>50,000</td>
</tr>
<tr>
<td>Rent</td>
<td>25,000</td>
</tr>
<tr>
<td>Profit</td>
<td>25,000</td>
</tr>
<tr>
<td>Total</td>
<td>$600,000</td>
</tr>
</tbody>
</table>

Value added = $600,000  
Tax payable = $600,000 X 10% (VAT rate) = $60,000

### CRITERIA FOR EVALUATING THE VAT

The effect of taxation on economic efficiency is linked to the concept of potential output. Potential output, as it relates to a market-type economy, may vary according to incentives available in the marketplace. The economist defines these incentives in terms of income and relative prices. To the extent that taxes can alter economic incentives or the return to productive activity, they induce changes in economic behavior.

Economists use several criteria to judge the desirability of a given tax or tax system. These criteria involve both efficiency and equity considerations.

### Neutrality

Price theory suggests that prices established in a competitive market are reflections of both consumer preference for goods and services and the least-cost combination of inputs used to produce those goods. A tax system is considered optimal if it does not interfere with the price allocation mechanism established in the market—that is, it is neutral toward economic behavior. Although all taxes distort or alter economic behavior to some degree, certain taxes can be judged "superior" on efficiency grounds. VAT proponents claim that the value-added tax would achieve the objectives of neutrality toward both economic behavior and allocative efficiency.

A tax on each firm’s value added can be thought of as a proportional tax on the firm’s use of the factors of production. If each firm combines land, labor, and capital in the most efficient (least-cost) manner, resources are bid into their most productive use. Payments to these factors in a competitive market are approximately equal to each factor’s contribution to the market value of each firm’s output. Thus, a uniform VAT on all firms imposes the same proportional tax cost—with respect to each firm’s payments to the sum of factor payments (value added)—and is, therefore, neutral toward the choice of production methods or the use of productive resources.

---

**Table 3**

Calculating VAT Liability Via Subtractive Procedure

| Total sales of goods and services | $1,000,000 |
| Total interest, dividends, and rents received | 200,000 |
| **Total (A)** | **$1,200,000** |
| Taxes | $200,000 |
| Purchase of materials, services, power, and capital on which VAT has been paid | 300,000 |
| Interest, dividends, and rents paid to other firms | 100,000 |
| **Total (B)** | **$600,000** |
| Value added = A - B | $600,000 |
| Tax payable = $600,000 × 10% (VAT rate) | $60,000 |

**SOURCE:** Lindholm, *Value-Added Tax and Other Tax Reforms*, pp. 30-31.

5Note that the dollar amounts of the income and consumption bases are equivalent over time. “The main difference between the two types is that, under the consumption concept, the tax base is smaller the first year and slightly larger in the following years. The total base added over the full depreciation period is identical under both methods.” Tax Foundation, “A Value-Added Tax for The United States?” (June 1979), p. 1.

6Norman Ture argues that the “consumption-type” label has been misnamed. The name promotes the view that it is a tax exclusively on consumption. Savers, who directly or indirectly are owners of capital, do not escape the tax because value added by capital is subject to the tax. For a given interest rate, the tax on the goods produced with capital raises the amount of net income firms must derive from a capital asset in order to justify its purchase. In other words, the tax reduces the demand for funds and, other things remaining the same, the interest earned by savers. See Charles E. McLure, Jr. and Norman B. Ture, *Value-Added Tax: Two Views* (Washington, D.C.: American Enterprise Institute for Public Policy Research, November 1972), pp. 88-92.

VAT: The European Experience

The VAT is by no means a new idea. Following World War I, the possibility of implementing a VAT-type system was discussed in certain European countries and in the United States. Although the debate continued over the next several decades, it was not until the 1950s that a VAT system was introduced in France and in the state of Michigan.

Table A1

<table>
<thead>
<tr>
<th>Country</th>
<th>Effective Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>1971</td>
</tr>
<tr>
<td>Denmark</td>
<td>1967</td>
</tr>
<tr>
<td>France</td>
<td>1954</td>
</tr>
<tr>
<td>Ireland</td>
<td>1972</td>
</tr>
<tr>
<td>Italy</td>
<td>1973</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>1970</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1969</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1973</td>
</tr>
<tr>
<td>West Germany</td>
<td>1968</td>
</tr>
</tbody>
</table>

In an attempt to foster trade between its members, the Council of the European Economic Community (EEC) issued a directive on April 1, 1967, which required all member nations to establish a VAT system by January 1, 1970. Although the directive concerned only the current members—Belgium, France, Germany, Italy, Luxembourg, and the Netherlands—it was extended to the United Kingdom, Denmark, and Ireland upon their joining the EEC in 1973 (table A1). Since most popular discussions of the VAT refer to the “European experience,” it may be useful to briefly examine the basis of its adoption and its record to date.

1The idea of the VAT originated with a German industrialist by the name of F. Von Siemens. In 1918 he advocated the substitution of the VAT for the newly implemented turnover tax. T. S. Adams suggested using such a tax in the United States in 1921. In the same year, the tax was included in an amendment to the Revenue Act of 1921 proposed by Sen. Reed Smoot. For a general discussion of the VAT refer to the “Value-Added Tax,” p. 165.

2France adopted the VAT as a replacement for existing turnover taxes. Michigan used a modified form of the VAT from 1953 to 1967 when it formed the Single Business Tax. This tax was introduced primarily to replace existing business income taxes.

3This directive is reprinted in G.S.A. Wheatcroft, Value-Added Tax in the Enlarged Common Market (New York: John Wiley & Sons, 1973), Appendices I and II.

Table A2

<table>
<thead>
<tr>
<th>Country</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>16.0%</td>
</tr>
<tr>
<td>Denmark</td>
<td>20.2</td>
</tr>
<tr>
<td>France</td>
<td>17.6</td>
</tr>
<tr>
<td>Ireland</td>
<td>20.0</td>
</tr>
<tr>
<td>Italy</td>
<td>14.0</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>10.0</td>
</tr>
<tr>
<td>Netherlands</td>
<td>18.0</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>8.0</td>
</tr>
<tr>
<td>Germany</td>
<td>12.0</td>
</tr>
</tbody>
</table>


The VAT was introduced in various European countries as an alternative to turnover taxes. These turnover taxes, sometimes referred to as cascading turnover taxes, are characterized by the payment of taxes at each level of production without regard to relief for the tax paid at a prior stage. Because of the way in which these taxes were implemented, there was considerable incentive for a firm to control all the stages of production (vertical integration). Also, there was a variety of rules and regulations about the taxation of certain exports and imports, which reduced the foreign trade potential between the countries. The VAT provided for the general tax exemption of export goods and services. Thus, the goal of the EEC in introducing VAT was twofold: to “harmonize” the members’ tax systems and to encourage intra-EEC trade.

While introduction of the VAT system has increased foreign trade among the EEC nations, harmonization of the various tax systems is still far from complete. As table A2 indicates, the “standard” VAT rate varies considerably among the various EEC countries. Moreover, each country may choose to apply different rates to various groups of goods and services. For example, Belgium has two basic rates: a standard rate of 16.0 percent on necessities, and a 25.0 percent rate on luxury goods. Likewise, Germany and the United Kingdom each have different rates for necessity and luxury items.

In addition, each country exempts certain items from taxation. Although a complete listing would point out the wide diversity of possible exemptions, there are several areas commonly exempted from VAT. As mentioned earlier, exports are fully exempt from VAT. Also, sales of securities, stocks and bonds, doctors’ services, financial services such as insurance and banking, and postal services are generally exempt.
A question remains regarding the neutrality of the consumption or the income approaches. Economic theory suggests that the consumption approach is less distorting since it does not alter the individual's choice between consumption and saving (i.e., current vs. future consumption). The income approach, however, discriminates against saving since capital expenditures are not deductible from the tax base.

**Capital Formation and Growth**

A common complaint about our tax system is that it discourages economic growth. VAT proponents claim that a consumption-type VAT would not. Since they believe that a reliance upon income taxation has altered individual consumption-savings choices through the income tax's impact on relative prices, proponents argue that the VAT would support both a higher level of savings and investment. Behind this argument lies the "double-tax issue," one of the oldest controversies in economics. In other words, "Should taxes be based on consumption or on income?" Consumption-base tax advocates claim that a tax levied on income is also a tax on savings and is therefore inefficient because it raises the "price" of saving (future consumption) relative to current consumption. Since a saver eventually is taxed on the interest earned from present savings, some analysts argue that the income is taxed twice. Hence, the double-tax issue.

The exemption of capital expenditures with a VAT would offset the non-neutral aspects of the current tax system, which relies heavily on income taxation, and may result in a higher level of overall savings and investment.

In addition, VAT would directly affect labor supply. Some analysts contend that existing income and

---

8Despite the positive economic attributes of the VAT in its more favorable consumption-type form, the tax is likely to end up being less neutral than its theoretical construction. Critics claim that political necessity requires that the tax be applied with numerous exemptions. It has often been asserted that exemption of certain categories of food, clothing, housing, medicine, and a long list of other special categories would be a political necessity. In addition to narrowing the overall tax base of the VAT, numerous exemptions would erode the neutrality of the tax, distorting the allocation of resources.

9See, for example, Musgrave and Musgrave, Public Finance, pp. 468-69.

10Of particular concern is the responsiveness of savings to the after-tax rate of return. Preliminary evidence indicates that this response may be greater than was once thought. See Michael J. Boskin, "Taxation, Saving, and the Rate of Interest," Journal of Political Economy (April 1978) Part 2, pp. S3-S27. For an additional discussion of this issue, see E. Phillip Howrey and Saul H. Hymans, "The Measurement and Determination of Loanable Funds Saving," Brookings Papers on Economic Activity (3: 1978), pp. 655-705.


social security taxes have altered the relative price of work and leisure. These taxes, they claim, have induced individuals to reduce their supply of labor at prevailing market wages below which might otherwise have been supplied without these taxes. Whether VAT would correct these distortions depends upon the relative responsiveness of labor to the removal or reduction of the existing taxes and the changes in relative commodity prices that may result from the VAT.

**Distribution of the Tax Burden**

Another question of interest is "Who bears the burden of taxation?" An economic examination of this issue can be made through what economists call "incidence analysis." Incidence is the change in an individual's real income (nominal income adjusted for changes in the price level) that results from the imposition of a tax.

The key to understanding incidence lies in distinguishing between statutory incidence (the legal liability of the tax) and economic incidence (the final burden of the tax). The sales tax provides a clear example of this distinction. Although its legal liability is imposed upon the retailer, the sales tax often is assumed to be fully passed on to the individual who purchases goods from the retailer. However, since the prices of the retailer's goods are now higher, the quantity demanded will be reduced, and the owner's income therefore will be changed. In this case, the economic incidence of the tax differs from the statutory liability since the former takes into account both the final resting point of the tax as well as the economic ramifications on others (the retailer, for example).

Tracing the final burden of a VAT is difficult because it ultimately depends upon the number of exemptions. Assuming that the tax is applied equally to all goods and services and is an "additional" tax, its burden would be fully shifted to consumers. If
exemptions are allowed or if other taxes are reduced, however, there will be changes in relative prices and in the production mix of the economy.

**Equity Criteria**

Equity criteria are often used to evaluate different taxes (and thus incidence patterns). These criteria attempt to answer the questions, "Who should pay taxes?" and "How should the burden of taxation be distributed among different individuals?" Although objective economic analysis provides few answers to these questions, it can provide a perspective on relative costs and benefits and on the trade-offs between economic efficiency and equity considerations.

A principal equity concern is over the relative tax burden imposed on individuals of different income levels. The concept behind our income tax system, for example, is that those individuals who earn a higher income should pay a higher percentage of their income in taxes. This is a "progressive" tax system. A "regressive" tax, in contrast, takes a lesser proportion of income from individuals who earn a larger income.

Critics of the VAT claim that it would be a regressive tax since it taxes consumption and since lower income individuals spend a larger proportion of their incomes on consumption goods relative to higher income individuals. Although this argument is straightforward, its validity depends on several unsettled issues.

First, one might question the preceding analysis of regressive or progressivity on the grounds that it ignores the benefit side of fiscal policy. In other words, the net benefits derived from government spending or transfer payments is not taken into account. Thus, a more meaningful evaluation would include the benefits received from government-related programs as well as the costs.14

On the other hand, the regressiveness of the tax could be alleviated either by adjustments to the income tax rate and/or by special tax credits. For example, the exemption of necessity items such as food, shelter, clothing, and medical care has been proposed. These exemptions, however, would reduce the tax base and thus necessitate higher tax rates to insure the same yield as under a comprehensive base approach. In general, therefore, attempts to alleviate the regressiveness of the tax are likely to complicate the administrative problems and interfere with the neutrality criteria discussed earlier.

**Administrative Costs**

The cost of administration is another criterion for evaluating the merits of a tax system. The initial administrative costs of introducing a VAT (or any other tax for that matter) would be relatively high.

Perhaps the most important factor affecting the cost of using a VAT is the degree of complexity of the tax. The use of multiple rates and numerous exemptions, in contrast to a single uniform rate, would raise the administrative costs. Indeed, it has been estimated that if the VAT system used involves more than a single rate, administrative costs may rise by 50 to 80 percent.15 Most of this increase would be due to increased personnel costs caused by a rise in the amount of paperwork required of both business and government.16

Another consideration is whether the tax is used to replace part of an existing tax or as a supplementary source of government revenue. If the VAT replaces only part of an existing tax or is merely added to the present system, administrative cost savings may be negligible. In fact, adding a VAT may increase the current cost of the government's tax collecting apparatus. This can be seen by considering the influence on costs from a reduction in existing tax rates. Since the previous tax system still exists and since there is little or no change in collecting or reporting procedures, the time and manpower involved in collecting a 5 percent tax is essentially the same as that for a 10 percent tax.

Furthermore, adding a VAT to the existing array of taxes would have differential cost effects on business. For instance, soon after a VAT was introduced in Germany in 1968, small noncomputerized firms estimated that tax-related administrative costs increased up to 20 percent. Relatively large businesses, 15National Economic Development Office, *Value-Added Tax* (London: Her Majesty's Stationery Office, 1971) p. 41, cited in Dan Throop Smith, et. al., *What You Should Know About the Value-Added Tax* (Homewood: Dow Jones-Irwin, Inc., 1973), p. 53.

however, reported negligible administrative cost increases.17

The frequency of collection is also a significant cost factor. Most of the European countries currently using a VAT require monthly payments.18 In the United Kingdom, however, the collection period is quarterly. Although a monthly payment schedule may create cash-flow problems for some businesses, the European experience suggests that this frequency is feasible since most firms already record the data needed to calculate their VAT liability on a monthly basis.

ECONOMIC EFFECTS OF VAT

Price Effects

A major concern about the implementation of a VAT is the resulting price effects. The introduction of the VAT, it has been widely asserted, will lead to a one-time increase in the general level of prices. In fact, it can be easily demonstrated that, although a uniform VAT rate (an identical rate for all goods and services) applied without exemption will not alter relative prices within the economy, it will result in a one-time increase in the overall level of prices.19 It should be noted, however, that this conclusion is based on several assumptions.

Perhaps the most important assumption is that all sellers raise their prices by the exact amount of the VAT, thus passing the additional cost on to the consumer. There are several reasons why this actually may not occur.

A VAT is usually applied with a varying rate structure—for instance, a 5 percent rate on food and a 10 percent rate on nonfood items—and/or with some items exempt from taxation.20 If sellers face different cost increases due to the VAT, then it is uncertain that all prices will be raised by an equal amount. Since none of these conditions will lead to identical increases in all prices, the consequent relative price changes would induce a shift-in the pre-VAT pattern of demand for goods and services. In this way, the production (and employment) decisions of producers are affected by the introduction of a VAT.

Introducing a VAT may cause a one-time change in the level of prices if the tax is merely a supplement to existing taxes. It is generally assumed that firms face higher costs because of the VAT and, therefore, attempt to pass this on to consumers. However, if the tax burden is reduced in some other area—for instance, the reduction in firms’ contributions to social security—then this assumption is unwarranted. To determine the final price effect of a VAT in this case requires knowledge of the tradeoff between reductions in existing taxes (if any) and the cost to the firm of administering the VAT.

Finally, it is unlikely that a rise in the general level of prices can be maintained without an increase in the money stock held by the public or an increase in the velocity of the existing money stock (decline in real money balances). If neither of these situations occurs, then changes in production and employment will occur since consumers would not be able to maintain previous consumption levels at the higher level of prices. Thus, the reaction of the government to observed changes brought on by the initial price effects of a VAT may further complicate the foregoing analysis.

Interjurisdictional Intrusion

How VAT might interfere with existing state and local taxes is also an important issue. State and local governments likely will be apprehensive about a VAT, given that the tax appears to be an intrusion into an area upon which these governments traditionally have relied—namely, the sales tax. Continued use of the retail sales tax should not conflict with federal use of the VAT from an economic efficiency view, however. If the VAT is applied at a uniform rate, relative prices will not be affected.

Some have argued, however, that a VAT may change taxpayers’ perceptions of the cost of government expenditures and may alter the level of public services demanded. If the tax is “invisible” or hidden in the price of goods, taxpayers may demand a higher level of public expenditures since the costs are not fully perceived. On the other hand, if the tax is fully perceived at the retail level (for example, if the VAT portion is stated separately at purchase), consumers may object to additional retail sales taxes, and state and local governments will find it more difficult to tax by this method.


18For a discussion of the various collection periods, see the articles contained in G.S.A. Wheatcroft, Value-Added Tax in the Enlarged Common Market (New York: John Wiley & Sons, 1973). This monthly scheme also appears to be favored by some U.S. officials. A comparison of views may be found in Daily Report for Executives (October 5, 1979), pp. K7-K8.


Stabilization

An important feature of a tax system is its overall response to cyclical fluctuations in the economy. For a given set of tax rates, the growth of the tax base normally will change with the level of aggregate economic activity. Provided that the change in tax revenues does not affect the level of government spending, these changes will cause variations in the overall budget surplus or deficit. "Built-in flexibility" is said to exist if tax liabilities rise (and fall) at a faster rate than income. Our tax system, based primarily on a progressive income tax, posits that the overall federal budget has a stabilizing effect on the economy since the tax acts to dampen fluctuations in economic activity. A substitution from income and social security taxes to a VAT likely would reduce such built-in flexibility.\(^{21}\)

Summary

The purpose of this article has been to provide a general background for understanding the value-added tax. Recent examinations have narrowly addressed the tax's price effects and its potential to reduce the existing tax burden imposed by federal income and social security taxes. While these are indeed important issues, a complete analysis of VAT must go deeper, examining the tax base, its neutrality, regressiveness, incidence, and effects on capital formation. In addition, the administrative costs of the tax itself must be considered as well.

\(^{21}\)Stated alternatively, a tax whose income elasticity of yield is greater than one is considered an automatic stabilizer. Elasticity is defined as the percentage change in tax yield divided by the percentage change in income (consumption) base.
The Demand for Currency: Is the Underground Economy Undermining Monetary Policy?

NORMAN N. BOWSHER

Currency holdings of the public have risen at a brisk pace for many years, reflecting a progressive expansion in the demand for currency.

Since 1960 the currency component of the money stock has risen at an average 7 percent annual rate. It is reasonable that the public, faced with higher prices, would seek to hold more currency. However, currency holdings adjusted for changes in consumer prices have risen at a 2 percent rate since 1960 while real demand deposits have been unchanged on balance. As a result, currency rose from 25 percent of demand deposits in 1960 to 29 percent in 1970, to 38 percent in 1979 (table 1). In late 1979 the average holdings of currency for a family of five amounted to an astonishing $2,500.

This greater demand has been readily accommodated since the supply of currency, as provided for in the Federal Reserve Act of 1913, is elastic. In the U.S. monetary system, any depositor can obtain more currency by withdrawing it from deposit balances at commercial banks. Commercial banks that are members of the Federal Reserve System, in turn, obtain currency to restock inventories from their Federal Reserve District Banks by having their reserve accounts (deposits) charged. Other banks obtain currency in exchange for deposits at their member correspondent banks.

The persistent increase of currency into circulation has been surprising in view of several developments which discourage its use. Credit cards, for example, have been substituted for cash in an expanding number of transactions, and the number of checking accounts and traveler’s checks issued have continued to grow. Moreover, interest rates have moved to higher levels, raising the cost of holding wealth in the form of currency and coin, which does not earn...
interest. In addition, the possession of currency in large amounts is risky; it can be lost, destroyed, or stolen. Holding liquid funds in the form of deposits, on the other hand, has been virtually free of such risks since the advent of deposit insurance.

Because of the sustained growth of currency outstanding relative to demand deposits (chart 1), questions have arisen about what purposes the currency serves. More specifically, does this currency support a huge and growing amount of unreported economic activity? Moreover, since currency is one major use for the monetary base, are there other monetary policy implications of the alleged change in demand for currency? This article explores the answers to these questions.

The Use of Currency in Unreported Activities

Suspicion has arisen that the pronounced growth in currency relative to demand deposits reflects an expansion of illegal transactions or other activities that the participants wish to hide.¹ According to Gutmann, currency lubricates a vast amount of unreported income and employment, an entire subterranean economy whose total product, estimated to be almost $200 billion in 1976, exceeded the nation's entire gross national product in the middle of World War II.²

A number of recent developments certainly have provided increased incentives to conduct unreported activities. Inflation has increased markedly the effective income and estate tax rates for most individuals. For example, a family of four who earned $12,000 in 1977 and who received an 8 percent increase in income in 1978 owed 21.8 percent more in federal income taxes (using standard deductions).³ As inflation pushes individuals into higher effective tax brackets, it provides an incentive to receive income in forms that cannot be traced readily. Currency payments, of course, are less traceable than deposit transfers. Increasing regulations, prohibitions, and reporting burdens provide other incentives for unreported transactions. Similarly, certain illegal activities such as drug dealing, which are almost always transacted in currency, apparently have expanded.

Nearly everyone receives a minor amount of unreported income in currency from "friendly" bets, personal services, garage sales, etc. Any firm or person who receives currency in a business transaction has the option of joining, at least in part, the subterranean economy of unreported income. Such opportunities exist for many engaged in the professions and those performing repair work and personal services, where it is traditional to compensate individuals with cash payments. Wages for part-time work, especially for child care and domestic workers, also are frequently paid in cash.

Since reliable information on currency usage is unavailable, indirect evidence has been relied on to support conclusions about the dollar volume of unreported transactions. Gutmann's estimates of "underground" activity financed by currency were based on several simple relationships and assumptions. He observed that the ratio of currency to demand deposits had risen markedly. In the five years before World War II, which he selected as a base for comparison, the ratio of currency to demand deposits averaged


21.7 percent. By 1976 this ratio had risen to 34.4 percent. If the ratio of currency to demand deposits required to support legal activities had not changed, the 12.7 percentage-point increase in currency (or $28.7 billion) over and above the rise in demand deposits represents funds devoted to increased illegal activity. Furthermore, if each dollar of currency used for illegal transactions supports the same amount of activity as each dollar of money held for legal purposes ($275.3 billion of demand deposits plus legal currency produced $1,693 billion of reported GNP in 1976), then the $28.7 billion held for illegal activities produced an estimated $176 billion of unreported income.

This estimate is conservative according to Gutmann. It makes no allowance for unreported income in the pre-World War II base period, it takes for granted that the earlier downward trend in the currency to demand deposit ratio would not have continued into the postwar period, it assumes all demand deposits are held for transactions and none for compensating banks for services rendered, and it ignores the growth in credit card usage and other currency saving devices.

Several more recent studies provide support for the broad Gutmann conclusion that the volume of unrecorded transactions is indeed substantial. The Internal Revenue Service, from comprehensive audits of a sample of tax returns, estimated that unreported income on 1976 individual income tax returns was as much as $100 billion from legal sources. It also revealed no taxes were paid on an additional $25 billion to $35 billion of individual income from three types of criminal activity — narcotics, illegal gambling, and prostitution.4 One result of this study is that the Treasury is seeking to obtain fuller disclosure of sizable currency transactions.5

In another study, Feige estimated that unreported transactions in 1976 totaled $369 billion (and a fantastic $704 billion in 1978).6 He calculated the total value of all transactions in the economy by using outstanding stocks of demand deposits and currency and estimating turnover rates for each. The derived total transactions series has been expanding faster than reported income (GNP). Feige used this excess as a measure of the irregular economy of unreported activity. His research into other possible explanations of the faster growth in transactions than reported income — changes in relative prices, relative volume of financial transactions, and the degree of vertical economic integration — concluded that, if anything, his estimates were low.

According to Gutmann, the underground economy for 1976 was about 10 percent of reported gross national product; the Feige estimate was nearly 22 percent. Using Feige’s calculations, unreported transactions had jumped to roughly 33 percent of GNP by 1978. From just before World War II until 1976, reported GNP increased at an average 7.8 percent annual rate; were the subterranean economy included, the trend growth rate would have been 8.1 percent according to Gutmann and 8.3 percent according to Feige. Feige’s research indicates that from 1976 to 1978 total nominal output, reported and unreported, expanded at a 16.9 percent annual rate, compared with the 11.8 percent rate of reported GNP.

On the other hand, questions have been raised by other analysts concerning the logic of the assumptions used in arriving at such large estimates of the subterranean economy. For example, Gutmann’s estimate ultimately depends on the use of demand deposits as the yardstick for judging currency growth. Were the comparison made with total deposits, the results would have been considerably different.7 Indeed, currency has changed little relative to total bank deposits since 1939, which suggests that the growth of demand deposits has slowed rather than that the growth of currency has accelerated (table 2). Hence, the subterraneaus economy is indeed of staggering proportions and growing rapidly.8

---


6Edgar L. Feige, “How Big is the Irregular Economy?” Challenge (November/December 1979), pp. 5-13. In the report Feige observed, “I began this investigation suspecting that the irregular economy was smaller than previous estimates had suggested. I am now convinced that the irregular economy is indeed of staggering proportions and growing rapidly.”

The Greek economy may not have grown relative to the rest of the economy.

Although time deposits are close substitutes for demand deposits and currency for many purposes, the value of such criticism is uncertain. It seems reasonable to relate currency to demand deposits for transactions analysis since these two are the only monetary assets that can be directly exchanged for goods and services; other deposits must first be converted into currency or demand deposits before they are spent.

In another criticism Laurent suggested that increases in total transfers of demand deposits and currency are more indicative of the trend of economic activity than increases in the outstanding stocks. Certain available data indicate a much larger increase in the turnover of demand deposits than of currency. Debits to demand deposit accounts at reporting banks increased by more than 30 times over the period 1939 to 1976. In marked contrast, according to Laurent, currency turnover slowed over the same period. Although there is no firm information on currency transactions, a rough measure of velocity was inferred by observing currency redeemed and destroyed. Currency is redeemed and destroyed when the paper notes show signs of wear, presumably a function of its use in transactions. From 1937 to 1941, the average life of the currency destroyed was 3.1 years; in 1976 it was 5.3 years. To some extent, this lengthening was due to the Federal Reserve's policy of extending currency life by screening it less rigorously. But an examination of data for the early 1970s, before such policy changes occurred, also shows an increase in currency life. The combination of the increased deposit turnover and a slowing of currency turnover markedly changes the significance of the relative growths of the currency and demand deposit stocks. The growth in the currency stock, viewed in this light, appears too small to accommodate both the growth in reported transactions and large expansion in unreported activities.

The Feige study, however, concluded that currency turnover had accelerated since the forties and early fifties, not decelerated. He noted that the quality of paper used in currency was improved dramatically by a melamine-formaldehyde resin additive in 1957. Tests demonstrated that, after this development, each bill could be used in nearly twice as many transactions before wearing out.

In addition, Laurent suggested that the rapid growth in currency outstanding is due, in part, to an increased use of currency as a store of value. The increase in large denomination bills suggests hoarding. In fact, there is now more money outstanding in $100 bills than in any other currency denomination (larger denominations are no longer issued).

On the other hand, some argue that the abundance of large denomination notes may just as well indicate an increased use of currency for illegal activity as for a store of value. Furthermore, the growth of illegal activity itself increases the stock of currency hoarded. Reportedly, the large tax evader or drug dealer frequently accumulates sizable amounts of currency in the larger denominations. These cash balances are maintained, despite high interest rates, to avoid arousing suspicion in converting this currency into other assets. When it becomes profitable to do so, the holder may purchase a legitimate business to "laundry" the illegal gains.

From a casual comparison of the growth of currency and the expansion of spending, there does not appear to be a great unexplained rise in currency. In fact, currency outstanding has expanded at a pace slightly less than the increase in total recorded spending on goods and services (see table 3). The marked rise in the ratio of currency to demand deposits, giving rise to suspicions of a rapidly expanding underground economy, has reflected sluggish growth in demand deposits. The lack of demand deposit growth can be explained by a proliferation of alternatives for such deposits, including repurchase agreements, NOW accounts, telephone and automatic transfers between savings and demand deposits, and money market mutual funds.

<table>
<thead>
<tr>
<th>Year</th>
<th>Currency</th>
<th>Demand deposits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1939</td>
<td>.07</td>
<td>.33</td>
</tr>
<tr>
<td>1959</td>
<td>.06</td>
<td>.24</td>
</tr>
<tr>
<td>1969</td>
<td>.05</td>
<td>.17</td>
</tr>
<tr>
<td>1976</td>
<td>.05</td>
<td>.13</td>
</tr>
<tr>
<td>1979</td>
<td>.04</td>
<td>.11</td>
</tr>
</tbody>
</table>

1All data are annual averages of daily figures except 1939 which are end-of-year data.

9Laurent, "Currency and the Subterranean Economy."

10Ibid.
In another examination of currency developments, Garcia argues that the rapid increase in currency outstanding should have come as no surprise to forecasters of currency demand. She notes that, since the 1950s, the amount of currency outstanding has approximated the predictions based on economic models. Using the econometric equations developed by Goldfeld, Garcia found that currency as a function of reported income, interest rates, and lagged currency holdings behaved as anticipated, while demand deposits rose at a surprisingly slower pace. Hence, she concludes that the subterranean economy operating on currency has not grown spectacularly since the early fifties; more likely it is roughly the same portion of GNP.

Gutmann responded to the Garcia study by pointing out that an equation may do a satisfactory job of forecasting without affording much insight into underlying causes. He faulted the currency equations for their omission of both a variable proxying the subterranean economy and the implicit yield on demand deposits. In the latter case, he argued that including the yield on the closest currency substitute — demand deposits — would, ceteris paribus, have led to a forecast of smaller currency growth over the period. This is because the implicit yield on demand deposits has risen, or, alternatively, the price of the substitute has fallen. He claimed that the success of the Goldfeld formulation in predicting currency demand, despite the omitted variables, is explained by the fact that (1) both currency and two of the model's explanatory variables (GNP and interest rates) rose rapidly over the period examined and that (2) the previous period's currency was included as an explanatory variable.

In brief, it is well known that much income goes unreported — this has been true for decades. Without more accurate information on currency exchanges, however, reliable estimates of unreported activities will be elusive. The great stock of currency outstanding and its persistent expansion in the face of higher costs of holding currency and the widespread use of currency substitutes (e.g., credit cards) certainly support suspicions of increasing irregular transactions. Moreover, heavier tax burdens and government regulation seem to have provided powerful incentives for attempting to hide more activities.

On the other hand, currency has not risen as rapidly as total spending. In addition, the velocity of demand deposits has increased markedly in the past several decades, while evidence on currency turnover indicates only a moderate acceleration (or even a deceleration). It is probable, therefore, that unreported transactions associated with currency have not risen as fast as the relative importance of currency outstanding in the money stock might indicate. However, the persistent expansion in currency outstanding at a time of strong incentives to avoid reporting could indicate that unreported activities have been expanding faster than economic activity generally.

**Monetary Control**

The rapid increase of currency outstanding in recent years also has implications for economic stability in addition to the unrecorded direct effects on spending, production, and employment. It is possible that flows of currency out of commercial banks could adversely affect the ability of the Federal Reserve System to control the total money stock. A useful framework for analyzing the impact of currency movements on monetary control is provided by the sources and uses of the monetary base or, more specifically, the net monetary liabilities of the Federal Reserve System and the U.S. Treasury.

Sources of the base include the monetary gold stock, Treasury currency outstanding, member bank borrowings from Reserve Banks, and Federal Reserve float (credit extended on checks received but not yet collected). The largest and dominant source of the base, however, is Federal Reserve holdings of government securities. Fluctuations in the base caused by irregular movements in the other sources can be read-

---

12Other studies have also concluded that the Gutmann and Feige estimates of the subterranean economy are excessive. See Richard D. Porter and Stephan S. Thurman, "The Currency Ratio and The Subterranean Economy: Additional Comments" (Board of Governors of the Federal Reserve System, January 26, 1979, processed); and Richard D. Porter, "Some Notes on Estimating The Underground Economy," (Board of Governors of the Federal Reserve System, August 10, 1979, processed).

13Other more minor factors that must be considered in analyzing the net sources of the base include Treasury cash holdings (having a negative impact), foreign deposits at Reserve Banks (having a negative impact), other Federal Reserve assets, other Federal Reserve liabilities, and capital (having a net negative impact).
ily offset as soon as detected (usually within one business day) by System purchases or sales of government securities. Several factors, such as data revisions, changes in member bank borrowing (in response to credit demands and System purchases and sales of securities), and the two-week lagged imposition of reserve requirements (forcing member banks to obtain reserves by borrowing from Reserve Banks if not otherwise made available) can prevent the System from hitting precise base targets within a month or two. Over a period of several months, however, the growth rate of the base can be determined within a narrow range by System purchases and sales.\textsuperscript{17} As studies generally have shown, fluctuations in monetary aggregates over periods of less than three months have seldom had any measurable effect on prices, employment, or incomes. Thus, the Federal Reserve has the ability to adequately control the magnitude and trend of the monetary base by buying or selling securities in the open market.\textsuperscript{18}

Analysis of the uses of the base is more complicated than an examination of its sources. Although there are just two uses of the base, currency in circulation and reserves to support member bank deposits, the analysis of money determination must go behind these aggregates. The U.S. monetary system is complicated by the existence of both member and non-member banks, different classes of member banks, different reserve requirements on different types of deposits, and graduated reserve requirements for deposits of different amounts. It is thus necessary to allocate reserves among numerous types of deposits for both those included in the money stock and those not included (such as government and time deposits), as well as to estimate the amount of excess reserves that the banking system will find expedient to hold.

In addition, a chief factor influencing the amount of money the banking system can create, given an increase in the monetary base, is the proportion of currency to demand deposits that the public desires to hold. If the public held a fixed amount of currency, all increases in the supply of monetary base by the Federal Reserve would remain in the banking system as reserves and would be reflected entirely in a multiple change in deposits, the amount depending on reserve requirement ratios for different types of deposits.

When currency holdings of the public rise rapidly as they have in recent years, however, the total monetary expansion from a given monetary base injection is greatly reduced. This results because an outflow of currency from banks absorbs an equal amount of reserves. With fewer reserves, the deposit-creating potential of the banking system is reduced by some multiple.

Clearly this increased currency demand must be considered in determining how much monetary base must be supplied to achieve a desired increase in the money stock. Fortunately, the System receives prompt data on currency outstanding through frequent reports, and trend movements of currency have been projected fairly accurately by forecasting models. Hence, prompt action to offset any undesired effects of currency movements has been possible.

During the seventies, currency holdings of the public rose rapidly. Unless offset, this acceleration in the public's demand for currency would have caused the growth rate of the money stock to slow or even to decline at times. The trend growth rate of the money stock, however, was more rapid in the seventies than in any other decade since World War II. During the 1970s there were periods in which the growth rate of the money supply slowed long enough to depress the growth of spending. There is no evidence, however, that the slower money growth reflected an increased demand for currency that could not have been offset by net System purchases of securities promptly enough to avoid any material contractive effects on spending.

**Conclusions**

The almost steady rise in currency in the past decade has fostered a controversy over whether it reflected a tremendous growth in illegal or otherwise unrecorded activities. Although reliable evidence on the extent of the subterranean economy is lacking, it is no secret that numerous unrecorded transactions take place daily. It also is probable that the volume of such transactions has been expanding faster than the volume of recorded transactions. Certainly, greater tax burdens and additional restrictions on legal activities provide incentives for an increase in subrosa pursuits. A number of studies, however, indicate no pronounced acceleration in unreported activities.

Lack of firm knowledge of the volume of unreported transactions does not reduce its economic importance. The implications of a large, growing, but unrecorded sector in our society are numerous and


affect measurements of economic activity and the economic policies based on these measurements. For example, because much income goes undetected, tax revenues are smaller than they would be otherwise.\textsuperscript{19} Official statistics, which are based on recorded transactions, understate both the true magnitude and growth rates of sales, income, employment, and production. As a result, economic policies are more apt to exacerbate inflation than if all activities were reported.

Unusually rapid growth in currency outstanding, given our monetary system, can potentially contract the total money stock (demand deposits plus currency). Nevertheless, since information on currency movements into and out of banks is current, and since offsetting transactions can be made by the Federal Reserve within a brief period, potentially depressing effects on the economy from this source have been avoided.

REPRINT SERIES

OVER THE YEARS certain articles appearing in *Review* have proven helpful to banks, educational institutions, business organizations, and others. To satisfy the demand for these articles, our reprint series has been made available on request. The following articles have been added to the series in the past few years. Please indicate the title and number of the articles you need and mail your request to: Research Department, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO. 63166.

<table>
<thead>
<tr>
<th>Number</th>
<th>Title of Article</th>
<th>Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>86</td>
<td>A Primer on the Consumer Price Index</td>
<td>July 1974</td>
</tr>
<tr>
<td>87</td>
<td>Channels of Monetary Influence: A Survey</td>
<td>November 1974</td>
</tr>
<tr>
<td>89</td>
<td>The FOMC in 1974: Monetary Policy During Economic Uncertainty</td>
<td>April 1975</td>
</tr>
<tr>
<td>90</td>
<td>A Monetary View of the Balance of Payments</td>
<td>April 1975</td>
</tr>
<tr>
<td>91</td>
<td>A Monetary Model of Nominal Income Determination</td>
<td>June 1975</td>
</tr>
<tr>
<td>92</td>
<td>Observed Income Velocity of Money: A Misunderstood Issue in Monetary Policy</td>
<td>August 1975</td>
</tr>
<tr>
<td>93</td>
<td>Crowding Out and Its Critics</td>
<td>December 1975</td>
</tr>
<tr>
<td>94</td>
<td>The FOMC in 1975: Announcing Monetary Targets</td>
<td>March 1976</td>
</tr>
<tr>
<td>95</td>
<td>The Link Between Money and Prices — 1971-76</td>
<td>June 1976</td>
</tr>
<tr>
<td>96</td>
<td>The Unemployment Rate as an Economic Indicator</td>
<td>September 1976</td>
</tr>
<tr>
<td>97</td>
<td>Development of Electronic Funds Transfer Systems</td>
<td>September 1976</td>
</tr>
<tr>
<td>98</td>
<td>Derivation of the Monetary Base</td>
<td>November 1976</td>
</tr>
<tr>
<td>99</td>
<td>Are You Protected From Inflation?</td>
<td>January 1977</td>
</tr>
<tr>
<td>100</td>
<td>The FOMC in 1976: Progress Against Inflation</td>
<td>March 1977</td>
</tr>
<tr>
<td>101</td>
<td>The Effects of Changes in Inflationary Expectations</td>
<td>April 1977</td>
</tr>
<tr>
<td>102</td>
<td>The Effects of the New Energy Regime on Economic Capacity, Production, and Prices</td>
<td>May 1977</td>
</tr>
<tr>
<td>103</td>
<td>Energy Resources and Potential GNP</td>
<td>June 1977</td>
</tr>
<tr>
<td>104</td>
<td>Revision of the Monetary Base</td>
<td>July 1977</td>
</tr>
</tbody>
</table>