DALLASFED

# Economic Review

Federal Reserve Bank of Dallas
January 1985

# What Is Money Today?

Dale K. Osborne

*Professor of Finance*
The University of Texas at Dallas
*Consultant*
Federal Reserve Bank of Dallas

> *Much ingenuity has been spent upon attempts to define the term* money. . . .*All such attempts at definition seem to me to involve the logical blunder of supposing that we may, by settling the meaning of a single word, avoid all the complex differences and various conditions of many things, each requiring its own definition.*
>
> —William Stanley Jevons, *Money and the Mechanism of Exchange*

> *It is in treating the simple elements that we require the most care and precision, since the least error of conception must vitiate all our deductions.*
>
> —Jevons, *The Theory of Political Economy*

In a previous article[1] I examined ten approaches to the definition of money and found three that promise rewards to further study. These approaches emphasize routine circulation (money is that which circulates routinely as a medium of exchange), simultaneity of payments (money is that, the entire stock of which can be spent at once), and finality of payment (money is that which pays for goods or services without creating a further debt for the payer).[2] These approaches provide three criteria for the identification of money, three tests to be passed by anything that deserves the name. Something that passes one or two of the tests might nevertheless fail another. Travelers' checks, for instance, pass the simultaneity and finality tests when spent by someone other than the issuer, but they fail the routine-circulation test because they do not circulate but are returned to the issuer by the first recipient. Insofar as we insist on all three tests, then, travelers' checks are not money.

The finality test turns out to be redundant in the

*The views expressed are those of the author and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.*

1. "Ten Approaches to the Definition of Money, *Economic Review*, Federal Reserve Bank of Dallas, March 1984.

2. See Leland B. Yeager, "Essential Properties of the Medium of Exchange," *Kyklos* 21, no. 1 (1968): 45-69, and "What Are Banks?" *Atlantic Economic Journal* 6 (December 1978): 1-14; G. L. S. Shackle, discussion of "Theoretical Foundations of Monetary Policy," by Robert W. Clower, in *Monetary Theory and Monetary Policy in the 1970s*, ed. G. Clayton, J. C. Gilbert, and R. Sedgwick (London: Oxford University Press, 1971), 32-34; and Harry G. Johnson, discussion of "Theoretical Foundations of Monetary Policy," by Robert W. Clower, in the same book, 29-31.

United States today (and in the United Kingdom too), for it is passed by everything that now passes the simultaneity and routine-circulation tests. Whether the finality test is redundant in other monetary systems is a question for further research. But both the simultaneity and routine-circulation tests are essential in all systems; the former to identify a *stock* instead of velocity or a mixture of stock and velocity, and the latter to ensure that the stock is *money*. Regarding the latter test, Leland Yeager (1978, p. 9) has explained very clearly that the routine circulation of money is what allows individual holdings to be adjusted so readily even in cases of aggregate excess demand, meaning that a positive or negative aggregate excess demand for money necessarily causes a corresponding positive or negative excess supply of things other than money. The property of routine circulation thus not only inheres in the concept of money but also explains the grave consequences of monetary disequilibrium. These consequences have been the central theme of monetary thought for two centuries, and it was a brilliant insight that traced them to such a simple and obvious property of exchange media.

Still, the capacity for routine circulation is as much a property of money on the move as of money at rest, and this is why we need the simultaneity criterion. G. L. S. Shackle introduced this criterion in the following passage:

> If we are allowed to take the meaning of *payment* as known we can...define money as the means of payment. Then the quantity of money existing at some moment, the stock of money, can be defined as the means of *strictly simultaneous* payment. That is to say, it is equal to the total of all those payments which could be made without the payers receiving or counting on the payments to be made by others. Simultaneity must be insisted on here, lest we mix up quantity and velocity. A single coin circulating fast enough can carry a payments flow of unlimited size. In thus defining the size of the stock of money, we must require all payments to wait for the gun and each of them to be represented, when the gun goes off, by its value in coins or something equally unconditional. The distinction between *size of stock* and *frequency of turnover* is doubtless best preserved until we see whether we need it. (1971, p. 32)

In this passage Shackle provides a simple and infallible means of distinguishing between the quantity of money and its velocity. The brilliance of his insight is the equal of Yeager's. It is the fate of such insights to seem obvious after the fact.[3]

The main purpose of the present article is to explain that one, and only one, identification of money satisfies both the simultaneity and routine-circulation criteria in the United States today, namely the monetary base (also called "high-powered money" and defined as the sum of currency outstanding and depository institutions' reserve balances at the central bank). Though the article aims at explanation not exhortation,[4] it attempts to deal with anticipated objections to the simultaneity criterion and to its combination with the routine-circulation criterion into a definition of money. Some of these objections spring from points of view that were shown in the first article to be problematical; thus they lead naturally to a review of the seven unsatisfactory approaches in the light of the proposed identification. Although those approaches do not yield a satisfactory definition of money, they might throw additional light on it.

As pointed out in the first article, three species of exchange media pass Yeager's routine-circulation test: currency, reserve balances with the central bank, and public transaction balances with depository institutions.[5] Our first task is to see

---

3. These insights can be found in the earlier literature but nowhere so clearly expressed. See, for example, Knut Wicksell, *Lectures on Political Economy*, translated from the 1906 Swedish edition (London: Routledge and Kegan Paul, 1935), vol. 2, *Money*, 19: "Money, on the other hand, always remains in the market, though in different hands." Does this point to Yeager's later emphasis on routine circulation? Again (p. 23): "and since the money in his possession cannot, at the same time, serve as a medium of payment or exchange for somebody else...." Isn't this an indirect expression of Shackle's criterion?

4. "Philosophical argument, trying to get someone to believe something whether he wants to believe it or not, is not, I have held, a nice way to behave toward someone; also, it does not fit the original motivation for studying or entering philosophy. That motivation is puzzlement, curiosity, a desire to understand, not a desire to produce uniformity of belief. Most people do not want to become thought-police." Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981), 13.

5. Transaction balances are balances in accounts payable on demand (to the owner or a third party). Most transaction balances in the United States are in demand accounts at cc mercial banks, NOW accounts at banks and thrift institutio

which of them pass Shackle's simultaneity test.

**Means of simultaneous payment**

It is clear that currency and the reserve balances of depository institutions at the central bank ("bank reserve balances" hereafter) can all be spent simultaneously. Payments in these media can be made "without the payers receiving or counting on the payments to be made by others." It is sometimes implicitly assumed that the public's transaction balances at banks could also be spent simultaneously. This would clearly be so in a 100-percent-reserve banking system and might be so in a monopoly banking system; but is it so in the fractional-reserve, multiple-bank system of the United States or the United Kingdom? The answer depends on how rigorously we administer Shackle's test.

Consider a society consisting of three persons, Ann, Beth, and Carla, and their respective banks, A, B, and C. Each person has a transaction balance of $100 at her bank. Each bank has reserves of $10, other assets of $90, and deposit liabilities of $100. Clearly, the $300 of transaction balances can be spent simultaneously if they are spent in the right way. For example, let Ann write a $100 check to Beth, Beth write a $100 check to Carla, and Carla write a $100 check to Ann; then all three checks can clear simultaneously and all three banks will maintain their positions unchanged. But suppose Beth and Carla write $100 checks to Ann, who simultaneously writes a $100 check to Beth; then Bank A's reserves and deposit liabilities rise by $100, Bank B's are unchanged, and Bank C's fall by $100. But as Bank C only had $10 of reserves to begin with, it cannot honor Carla's check with its present funds. It has to borrow funds or sell its other assets, and its payment of Carla's check is "counting on" payments received from some other source.

This second pattern of payments is admittedly unusual. Most of the time, deposit flows into and out of a bank are approximately equal, and any net outflow is within the bank's reserve. Such was the case for all three banks in the first payment pattern considered above and for Banks A and B in the second pattern. But this just means that the inflows "counted on" by a bank usually arrive in time to support the outflows. Therefore, I believe that the spirit of Shackle's test implies a rigorous enforcement: simultaneity must be insisted on no matter what the pattern of payment.

If this conclusion is reasonable, it means that transaction balances fail Shackle's test. Any outflow of funds greater than a bank's reserve is possible only upon inflows from some source: deposits of currency, deposits of checks written on other banks or the central bank, or sales of assets. Only the outflows that don't depend on such support are simultaneously possible in all banks, and their maximum equals total bank reserves.

This conclusion is consistent with financial practice. Banks do not grant their depositors access to funds represented by checks written on other banks (called transit checks) until they collect from those banks.[6] They don't even credit a loan customer's compensating balance with deposited transit checks prior to collection (the compensating balance is usually measured as the "average collected balance"). And securities dealers do not accept ordinary checks but require payment in Immediately Available Funds.[7] Indeed, the U.S. Treasury and Federal Reserve require Immediately Available Funds even with noncompetitive tenders for Treasury issues. Transaction balances are not means of payment in the money or capital markets.

Transaction balances probably ought not to be regarded as means of payment or media of exchange in any market. Strictly speaking, checks written on transaction accounts are orders to bankers to pay reserves. It is the reserves that really serve as means of payment. But to insist on such a

and share draft accounts at credit unions. See the box for definitions of the financial aggregates most commonly discussed in connection with money.

6. The exceptions are cashiers' and certified checks, wire transfers, and other orders to pay "immediately available funds." In addition, banks often cash transit checks for a fee or as a favor to valued customers. For a brief report of bank practices regarding transit checks, see Mary Williams, "Banks' Slowness in Clearing Checks Prompts State and Federal Actions," *Wall Street Journal*, 6 March 1984, Southwest edition, sec. 2.

7. The customary delays between agreements to trade and settlements of trades are two business days in the money markets and five business days in the capital markets. They give the buyer time to obtain Immediately Available Funds and are reflected in price quotations.

## Definitions of Some Financial Aggregates

| Item | M1 | M2 | M3 | L |
|---|---|---|---|---|
| Currency in circulation | X | X | X | X |
| Travelers' checks of nonbank issuers | X | X | X | X |
| **At commercial banks** | | | | |
| Demand deposits (except those due to domestic banks, U.S. Government, foreign banks, and official institutions) | X | X | X | X |
| NOW (negotiable order of withdrawal) accounts | X | X | X | X |
| ATS (automatic transfer service) accounts | X | X | X | X |
| **At thrift institutions** | | | | |
| Demand deposits (at mutual savings banks) | X | X | X | X |
| NOW accounts | X | X | X | X |
| ATS accounts | X | X | X | X |
| Credit union share draft balances | X | X | X | X |
| **At commercial banks** | | | | |
| Overnight RPs (repurchase agreements) | | X | X | X |
| Small time deposits (less than $100,000) | | X | X | X |
| Savings deposits | | X | X | X |
| **At thrift institutions** | | | | |
| Savings deposits (at mutual savings banks and savings and loan associations) | | X | X | X |
| Small time deposits (less than $100,000) | | X | X | X |
| **Other** | | | | |
| Overnight Eurodollar deposits of nonbank U.S. residents at Caribbean branches of member banks | | X | X | X |
| Money market mutual fund shares | | X | X | X |
| **At commercial banks** | | | | |
| Large time deposits ($100,000 or more), including large negotiable certificates of deposit | | | X | X |
| Term RPs | | | X | X |
| **At thrift institutions** | | | | |
| Large time deposits ($100,000 or more) | | | X | X |
| Term RPs at savings and loan associations | | | X | X |
| Term Eurodollars held by nonbank U.S. residents at Caribbean branches of member banks | | | | X |
| Bankers' acceptances | | | | X |
| Commercial paper | | | | X |
| U.S. Treasury bills and other liquid Treasury securities | | | | X |
| U.S. savings bonds | | | | X |
| Consolidation component[1] | | X | X | X |

Total credit = total credit-market debt owed by domestic nonfinancial sectors.

Monetary base = currency held outside the Treasury, Federal Reserve banks, and the vaults of depository institutions
+ reserve balances at Federal Reserve banks (current)
+ vault cash used to satisfy reserve requirements at all depository institutions (held two weeks earlier)
+ surplus vault cash at depository institutions.

1. Consolidation component: less cash items in the process of collection, interbank deposits, Federal Reserve float, and estimated proportion of demand deposits used by thrift institutions to service their transaction accounts.

distinction is to risk a charge of hairsplitting, and I shall honor established usage by treating such balances as consistent with Yeager's test. For our purpose the point is moot, as the balances clearly fail Shackle's test.

Above, I mentioned in passing that transaction balances would pass Shackle's test in a 100-percent-reserve banking system and might do so in a monopoly banking system. Let us now consider these systems. In a 100-percent-reserve system, total bank reserves equal total transaction balances, but the sum of reserves and transaction balances is not, of course, added to public holdings of currency when counting the money stock. The writing of a check sets in motion two transfers that are part of a single transaction in the goods or securities markets—a transfer of transaction balances from buyer to seller and a transfer of reserves from the buyer's bank to the seller's. Both transfers are part of the same payment, the amount of which is the cost of the things bought. Only this amount is spent. It may be entered either in the transaction balances column or the bank reserves column. If the first, money is identified as (essentially) M1; if the second, it is identified as the monetary base. Either way, the total money stock is the same. Therefore, even in a 100-percent-reserve system, the Yeager-Shackle money stock equals the monetary base.

In a monopoly banking system, the monetary status of transaction balances depends on two factors. The first factor concerns the existence of currency or other exchange media into which the balances could be converted on demand by the depositor. If no such alternative media exist, transfers of balances cannot cause the bank to lose reserves—indeed, the bank need not have any reserves to lose—and it can honor all checks written by its depositors. In this case all transaction balances can be spent simultaneously, and they constitute the money stock.[8] But if alternative media exist, into which the balances are convertible on demand, the bank must keep its reserve in them and can honor conversions only up to its reserve. Now the second factor comes into play: ought we to regard a conversion as a kind of payment, which is therefore subject to Shackle's test, or as something distinct that is not covered by the test? If the former, transaction balances are not money; if the latter, they are. A decision either way strikes me as arbitrary. Nothing in Shackle's approach obviously forces the issue one way or the other. We have to leave the matter open.

The Yeager and Shackle tests do not decisively identify the money stock in case the circulating media consist of transaction balances at a monopoly bank and currency or other media that also serve as the bank's reserve. This is the case of Russia and the other socialist states that own their banking systems. Its resolution evidently requires the formal recognition of some additional criterion for the identification of money not reached by Yeager's and Shackle's criteria. But Yeager's and Shackle's criteria suffice for the United States and United Kingdom today, where they decisively identify money as the monetary base.[9]

**Objections**

The identification of money today as the monetary base is sure to encounter objections. Objections could spring from any of the seven unsatisfactory approaches to a definition reviewed in the earlier article. For example, someone might object that the monetary base has not been shown to have the stablest demand function or the closest relation to GNP. Although such objections do not really call for a reply, as they do not appeal to defining characteristics of money, they might contain the germs of theorems. Although money cannot be *defined* as that financial aggregate which has the stablest demand function, empirically it might turn out to have that property. I shall review the seven approaches from this point of view in the next section. In the present section I'll attempt to answer objections that might be provoked by the novelty of the identification.[10] Most of us are so used to think-

8. This is the case of the example (given in section 2 of the March 1984 article) of balances in the Storekeeper's ledger. The Storekeeper functions as a monopoly bank in the example.

9. This statement is not strictly correct. Yeager's and Shackle's criteria collectively imply that *if* there is money in the United States today, it is the monetary base. They do not rule out the possibility that additional pertinent criteria might be inconsistent with them. My analysis and conclusions do not disprove the assertions of those who maintain that we haven't had any real money in this country since we went off gold.

10. The identification is novel only in terms of contemporary usage, which did not become dominant until well into the

ing of transaction balances as money that we naturally object when told to change.

1. *Why refuse to call transaction balances money? After all, they circulate routinely as media of exchange.* This objection overlooks the fact that routine circulation is only a necessary and not a sufficient identifying property of money. Not only transaction balances but also currency and bank reserves routinely circulate as exchange media. In order to decide which combination of these media ought to be called money (including the combination of all three), we have to consider additional properties of money. The additional property here considered is simultaneity, and it rules out transaction balances for the reason explained in the preceding section.

2. *Why not just identify money as the largest stock of routinely circulating exchange media (essentially M1), as is customary?* The largest stock is not M1 but the sum of M1 and reserves. No one calls this sum money, and the reason is instructive. Under the customary identification of money as M1, reserves are the *assets* of issuers of most of the money, and by convention only the monetary *liabilities* of issuers are counted as money. Reserves are, of course, liabilities of the central bank, but conventional monetary analysis consolidates the central bank's balance sheet with that of the banking system to obtain the aggregate "money" stock, and this consolidation nets out reserves. This consolidation would not make sense if the decision had not already been made to treat transaction balances as money. Thus it is the prior decision to call transaction balances money that explains why reserves are not also called money. This prior decision is the real objection; like the first objection, it is answered in terms of simultaneity.

3. *Why insist on simultaneity? Why can't we call transaction balances money even though they can't all be spent simultaneously? After all, the question of spending them simultaneously does not arise in practice.* This is true (it is also true of the monetary base, for holders of the base never attempt to spend all their holdings at the same time), but it is beside the point. The simultaneity criterion does not appeal to any supposed practical constraint on total spending but to an elementary characteristic of the money stock. This characteristic (that it can all be spent at the same time, without any payment "counting on" any other payment) ought to be an intuitively obvious characteristic of the total stock of money. But in case the characteristic is not obvious, consider the following argument based on two assumptions.

> Assumption A. *There is some period of time during which each dollar of the entire money stock can be spent.*
>
> Assumption B. *If each dollar can be spent during a period of length $L$, there is some number $k$, $0 < k < 1$, such that each dollar can be spent during a period of length $kL$.*

Assumption B asserts, in effect, that those money holders who spend near the end of the period (during the interval $L - kL$) could, if they wish, spend their money sometime closer to the beginning of the period. This assumption is clearly permissible. Since Assumption A does not require any dollar to be spent more than once during period $L$, the payments made near the end of the period need not "count on" payments made earlier, that is, no dollar need be respent during the period.

Assumptions A and B immediately imply the simultaneity characteristic. Let $L$ be the period mentioned in Assumption A. Then Assumption B implies that $kL$ is also such a period for some $0 < k < 1$. But now $kL$ is a period mentioned in Assumption A, and by Assumption B, $k(kL) = k^2L$ is also such a period and so are $k^3L$, $k^4L$, . . . . Each period in the sequence $L$, $kL$, $k^2L$, $k^3L$, . . . is shorter than its predecessor. If extended far enough, the sequence will approach arbitrarily close to an instant of time, that is, to simultaneous payability of the entire money stock.

In terms of this argument, there are only two ways to deny the simultaneity characteristic. One way is to count some things as money when we con-

---

20th century. For examples of the older usage, which denied monetary status to transaction balances, see Irving Fisher, *The Purchasing Power of Money* (New York: Macmillan Company, 1911), 11; F. W. Taussig, *Principles of Economics*, 3d ed., 2 vols. (New York: Macmillan Company, 1921), 1:432; and H. Parker Willis and George W. Edwards, *Banking and Business* (New York: Harper & Brothers, 1922), 11. For a survey of common usage up to about World War II, see Charles Rist, *History of Monetary and Credit Theory*, translated by Jane Degras from the 1938 French edition (New York: Macmillan Company, 1940).

sider the whole period $L$ that we do not count as money when considering some part ($k^rL$) of the period ($r = 1, 2, \ldots$). In other words, we can alter the concept of money somewhere in the sequence. This choice would make our measure of the money stock intrinsically time-dependent. We could not say that the money stock is $M$ dollars during period $L$ but must say that it is $M(k,r)$ dollars, depending on the fraction ($k^r$) of the period that we considered. This would mean that the quantity of money could not be regarded as fixed throughout any period $L$, no matter how short.

The other way to deny the simultaneity characteristic is to count some money more than once during the whole period $L$ but not during some fraction of the period (that is, to recount some money near the beginning of the sequence but not near its end). To do this would be compounding the stock concept of money with its flow concept: it would be "mixing up quantity and velocity."

4. *Why is it necessary to distinguish between quantity and velocity?* For many practical purposes it is not necessary. The product of quantity ($M$) and velocity ($V$) equals total spending ($PQ$[11]), and predictions of $PQ$ require predictions of $MV$ but not of $M$ or $V$ separately. Still, if we want to understand the cause of some particular change in $PQ$, we'd better be able to distinguish changes in $M$ from changes in $V$.

Economists have fallen into the habit of calling any ratio, $PQ/M_a$, the "velocity of $M_a$" even when $M_a$ is a broad financial aggregate consisting largely of assets that don't circulate in payment. Expressions such as the "velocity of $M_2$," the "velocity of $L$," or even the "velocity of total debt" are harmless enough if we always remember that they refer to the ratios $PQ/M_2$, $PQ/L$, or $PQ/$*(total debt)* and not to a rate of turnover. Only the circulating media have velocities, for only they are spent.[12] The broader aggregates are often used in something that looks like an equation of exchange; for example, with $V_2$ defined as $PQ/M_2$, we have $M_2V_2 = PQ$ identically. But this is an identity only because of the residual definition of $V_2$. Anyone who is content to define velocity in this manner can obtain equations of exchange in anything, even chairs: define the "velocity of chairs" as the ratio of $PQ$ to the stock of chairs, multiply this ratio by the stock of chairs, and—presto!—an identity that looks like an equation of exchange. Such an identity has only a superficial resemblance to equations of exchange in the monetary base or in M1.

A conceptually correct equation of exchange contains two kinds of velocity: the actual velocity, defined as the average number of times a dollar is spent, and the "virtual" velocity, defined as the ratio, to the stock of money, of total purchases of goods and services financed by the creation of new debt or the surrender of someone else's existing debt. Actual velocity refers to the transfer of money, virtual velocity to the transfer of claims to money.[13] Where money is identified as M1, the equation of exchange puts transfers of transaction balances under the head of actual velocity, for such balances are regarded not as claims to money but as money itself. When money is identified as the monetary base, the equation puts such transfers under this head only so far as they induce a transfer of reserves between banks, all the rest going under the head of virtual velocity. The latter approach is logically and economically sounder, for the economic effects of purchases financed by the surrender of existing debt are the same whether the debt is that of a bank or some other party.

11. Total spending, $PQ$, is strictly speaking the inner product of the vectors $p = (p_1, p_2, \ldots, p_n)$ and $q = (q_1, q_2, \ldots, q_n)$ of prices and traded quantities, respectively, of the $n$ goods. Most economists assume that these vectors can be represented by index numbers $P$ of prices and $Q$ of quantities. In the interest of brevity I shall do the same, for nothing in my argument depends on it.

12. "Velocity" is, of course, a misnomer for *turnover*, or the average number of times a dollar of currency, reserves, or transaction balances changes hands. (If exchanges for goods only are counted, the result is income velocity; if exchanges for securities are also counted, the result is transaction velocity.) Wicksell's discussion of velocity in chapter 3 of the *Lectures* (vol. 2) is still very useful.

13. The term "virtual velocity" was introduced (I think) by Wicksell (*Lectures* 2:67ff.) in recognition of the role of circulating debt (or equivalently, circulating credit) in economic activity and the economic distinction between this role and that of money. The concept long predated the term. When Wicksell wrote (the *Lectures* were published in 1906), most circulating debts were bills of exchange; today they are transaction balances. I will attempt to derive a conceptually correct and intuitively satisfying equation of exchange in a future article.

5. *An excess demand for M1 is a classical monetary disequilibrium. Why not therefore identify M1 as money?* This is potentially a powerful objection. Money is important in macroeconomic theory precisely because its excess demand is thought always to be transmitted inversely to the entire economic system. If money is in positive excess demand, other things (financial assets or goods or both) are in positive excess supply; the effect will be a contraction in prices or real activity and perhaps in both. If money is in negative excess demand (positive excess supply), other things are in positive excess demand; the effect is an expansion in prices or real activity. This is the chief lesson of two centuries of monetary thought. If an excess demand for M1 produces such effects (as we have all been taught and as is assumed in the objection), my refusal to call it money might well be a case of irrelevant hairsplitting.

Before dealing with the objection, it might be well to explain that excess demand for the monetary base has the macroeconomic effects noticed above. There are four main cases to consider, for the excess demand may be positive or negative and it may originate on the demand side or the supply side. For convenience we assume that each case represents a disturbance to an initial equilibrium. This assumption is not essential to the argument, and in practice an excess demand for money might well arise endogenously from preceding or concurrent developments in other markets.

*Case A1: Positive excess demand owing to an increase in demand.* This case springs from an exogenous increase in the banks' desired reserve ratios or in the public's desired currency holdings. If the former, banks will reduce their lending or raise the rate they pay for deposits and other sources of funds; if the latter, banks will lose reserves to the public and will have to reduce their lending. On either supposition, the supply of loanable funds will decrease. And insofar as the public is trying to build up its currency holdings, it will be reducing its holdings of other assets or its purchases of goods. All these actions by the public and the banks tend to depress prices or production or both. The effect is clearly contractive.

*Case A2: Positive excess demand owing to a decrease in supply.* Here there is an exogenous withdrawal of currency or reserves by the central bank. Banks will try to restore their reserve ratios by cutting back their loans or increasing their prices for funds. The supply of loanable funds again falls. Insofar as the public tries to restore its currency holdings, it will cut back on holdings of other assets or purchases of goods. The effect is qualitatively the same as in Case A1.

*Case B1: Negative excess demand owing to an increase in supply.* Here there is an exogenous increase in currency or reserves by the central bank. Since by assumption the demand has not changed, banks find themselves with excessively high reserve ratios and will try to reduce them by lending more. The supply of loanable funds increases. If the public has more currency than it desires, it will try to get rid of the excess by purchasing other financial assets (possibly including bank deposits) and goods. The actions of the public and the banks clearly produce an expansive effect.

*Case B2: Negative excess demand owing to a decrease in demand.* This case springs from an exogenous decrease in the banks' desired reserve ratios or in the public's desired currency holdings. It is the mirror image of Case A1 and is described in precisely the same terms but with a change of sign (change "increase" to "decrease" throughout, and vice versa).

Thus whatever the source of disturbance, the effects of a positive excess demand for the monetary base are unambiguously contractive while the effects of a negative excess demand are unambiguously expansive.[14] The objection assumes that the same effects follow from excess demand for transaction balances. But do they?

Consider a positive excess demand. If this arises from a decrease in supply, it is because bank reserves fall or the banks' desired reserve ratios rise. The effects are unambiguously contractive, as in Case A2, but since they spring from a disturbance in the reserves market, they contribute nothing new to the analysis of that case. Therefore suppose an increase in the public's demand for transaction

14. For a fuller discussion, see Don Patinkin, *Money, Interest, and Prices*, 2d ed. (New York: Harper & Row, 1965), which, after all, is mainly concerned with "outside" money (in our terms, the monetary base). Regarding the issues at hand, Patinkin's treatment agrees in all essentials with (say) Knut Wicksell, *Interest and Prices*, translated by R. F. Kahn from the 1898 German edition (1936; reprint, New York: Augustus M. Kelly, 1965), especially chap. 5.

balances. We saw in Case A1 that an increase in demand for the *monetary base* has unambiguous contractive effects. But the effects of an increased demand for *transaction balances* depend on what the public is turning away from, that is, on which of its demands are decreasing. There are three possibilities: (1) currency, (2) other financial assets, and (3) goods.

1. If the public is trying to build up its transaction balances at the expense of currency, it will send currency to the banks. The banks then find their reserves increasing and their costs of attracting deposits decreasing; they will increase their lending and their deposits. The effects are *expansive.*

2. If the public is trying to build up its transaction balances at the expense of other financial assets, the analysis is more complicated. On the one hand, the public's diminished demand for other financial assets tends to depress the prices of those assets and, by raising the cost of capital to business firms, to depress real investment. On the other hand, the public's increased desire for transaction balances tends to reduce the banks' cost of funds; this, together with the depressed prices of securities, raises the profits of intermediation and stimulates the banks to try to increase their activities.

If the banks are not bound by a regulatory required-reserve constraint, they will increase their holdings of earning assets. The actions of banks and the public have opposing effects on economic activity. The net effect depends on the elasticities of several demand and supply functions and may be expansive or contractive. But if the banks *are* bound by a required-reserve constraint, they will not be able to increase their aggregate security holdings and, therefore, will not exert an opposing force to the public's actions. If the reserve constraint is absolute, the story ends here with a contraction in economic activity. The banks have only succeeded in bidding up the interest rate on interbank funds (federal funds in the United States) until they are content with the same level of earning assets as before.

But in practice the reserve constraint is not absolute for three reasons, two of which involve the central bank. First, the central bank might be trying to keep the federal funds rate constant, in which case it will create additional reserves for the banking system. (This was the case in the United States during the 1970s.) Second, the central bank's lending rate (the discount rate in this country) might be below the federal funds rate (as it almost always is), leading the banks to exploit the growing differential between the two rates by borrowing more from the central bank. If the central bank is operating with an unborrowed-reserves target, it will accommodate the banks.[15] Third, the banks can turn to a number of devices, such as overnight repurchase agreements or Eurodollar borrowings, that reduce their reserve requirements against a given level of deposit liabilities. For all these reasons, indirect finance through the banks will be increasing as direct finance by the public is decreasing. The expansive tendencies of indirect finance might well dominate the contractive tendencies of the diminished direct finance, but strictly speaking the outcome is uncertain.

3. If the public is attempting to build up its transaction balances at the expense of goods, it exerts a positive effect on the supply of loanable funds (for banks' deposit costs fall) and a negative effect on the demand for goods. The analysis is similar to that of paragraph 1 above but a bit more complicated because it must deal with a probable shift in output from consumption goods to capital goods. The point is that the net effect is indeterminate.

Thus a positive excess demand for transaction balances has contractive, expansive, or indeterminate effects, depending on its source. A similar analysis of excess supply (which is a straightforward modification of the above) leads to a similar conclusion. By contrast, excess demand for the monetary base has unambiguous effects regardless of its source. We conclude that it is excess demand for the monetary base, not for M1 or transaction deposits, that invariably creates the troubles so well described in all leading monetary theories. The money of monetary theory is the monetary base.

### Searching for theorems

The approaches to a definition examined in the

15. That borrowings from the central bank vary directly with the differential between discount and federal funds rates has been shown by Stephen M. Goldfeld and Edward J. Kane, "The Determinants of Member-Bank Borrowing: An Econometric Study," *Journal of Finance* 21 (September 1966): 499–514; and Donald Dutkowsky, "The Demand for Borrowed Reserves: A Switching Regression Model," *Journal of Finance* 39 (June 1984): 407–24.

earlier article emphasized a number of properties or asserted properties of money, such as its tangibility, its relation to GNP, and the stability of its demand function. Now that we have identified our current money with the aid of Shackle's and Yeager's approaches, let us see whether the other approaches throw additional light on it.

1. *Tangibility*. Tangible currency is the greater part of the monetary base. Bank reserve balances, the remaining part, though not tangible, can be converted entirely into currency on demand. Therefore, for all practical purposes, the money stock is tangible today. Eugene Fama's emphasis on tangibility thus leads approximately to a correct identification of today's money.[16] But it is not an infallible approach because tangibility is not intrinsic to money. The earlier article presented an example of a model monetary economy in which money consisted entirely of the intangible balances in the Storekeeper's ledger.[17] And if we should ever become a "cashless" society, paying for all purchases by check or electronic transfers[18] while still blessed with the existence of a central bank, our money stock would consist of the central bank's intangible liabilities to commercial banks. Money would then be defined, as it is now, in terms of routine circulation and simultaneous payability, but it would have a different identity.

2. *Liquidity*. Money is always more liquid than any other asset, for the measure of liquidity in a monetary economy is the ease with which an asset can be sold for money. Liquidity is therefore intrinsic to money. But as we saw in the earlier article, to call an asset money because it is very liquid, or to deny the economic significance of money because other assets perform some of its functions, is to commit a fallacy of composition or a fallacy of illicit conversion.

3. *Means of payment*. Both components of the monetary base are important means of payment. Bank reserve balances are especially important because they are the means of payment in the transactions between banks that accompany check transactions between customers of different banks. But the base is only a subset of the total means of payment. Travelers' checks and transaction balances are also means of payment, but the former fail Yeager's test and the latter fail Shackle's.

Though the means-of-payment property is intrinsic to money, it logically derives from the fundamental medium-of-exchange property. These properties are not equivalent. Every circulating medium necessarily serves as a means of payment. The converse is false, as is shown by travelers' checks. Even means of payment that are also circulating media need not be money, for, like transaction balances, they might fail Shackle's test.

4. *Means of payment creating no further obligations*. That something is money only if its use as such creates no further obligations is Johnson's criterion. The scope of the criterion is restricted to money used by someone other than the issuer if the issuer is obliged to repay or redeem it (as in the case of private issuers or governments that issue paper money convertible into a commodity serving as the monetary standard).[19] Although it is free of obvious logical defects, Johnson's criterion does not seem as basic as Shackle's or Yeager's. Perhaps the property it describes ought to be regarded as derivative rather than fundamental. It is certainly derivative in the United States today, for everything that now passes Shackle's or Yeager's test also passes Johnson's. The real force of Johnson's criterion seems to bear more heavily on the concept of payment than on the concept of money. This is appropriate in view of the context in which Johnson wrote.

16. Eugene F. Fama, "Banking in the Theory of Finance," *Journal of Monetary Economics* 6 (January 1980): 39–57.

17. See pages 5 and 6 of the 1984 article. Transaction balances pass the routine-circulation and simultaneity tests in that model economy. They are economically equivalent to our banks' reserve balances at the central bank, not to our public balances at commercial banks.

18. This is extremely unlikely, as the public will always want to hold some of its wealth in a form that can be spent without the assistance of banks or other institutions. The importance of money to personal freedom has been emphasized by Georg Simmel, *The Philosophy of Money*, translated by Tom Bottomore and David Frisby from the second German edition (London: Routledge and Kegan Paul, 1978); Rist, *History of Monetary and Credit Theory*; and S. Herbert Frankel, *Money and Liberty* (Washington, D.C.: American Enterprise Institute for Public Policy Research, 1980).

19. The restriction has no force in the United States today, where only the government issues money and is not obligated to redeem it.

Table 1
**CORRELATION BETWEEN MONEY OR VARIOUS FINANCIAL AGGREGATES AND GNP**
(Quarterly data, 1960–83)

| Aggregate | Coefficient of correlation with GNP |
|---|---|
| Monetary base | .99505 |
| M1 | .99636 |
| M2 | .99654 |
| M3 | .99699 |
| L | .99670 |
| Total credit | .99772 |

5. *Correlation with GNP.* The facts about contemporaneous simple correlations from 1960 through 1983 are shown in Table 1. The aggregates are listed in order of increasing size, and the table shows the tendency of the larger credit aggregates to be somewhat more highly correlated with GNP. This is easy to understand. The larger aggregates are reasonable proxies for total financing in the economy, which varies with the economic activity measured by GNP. This high correlation sheds no light on money, for it would just as surely exist in a barter economy.[20]

Contemporaneous simple correlation is, of course, only the simplest method of investigating the relations between GNP and the financial aggregates. But more complicated methods, whether employing reduced forms of theoretical models or time series analysis free of theory, do not paint a clearer picture. The reduced forms (primarily of IS-LM models) achieve better fits with large credit aggregates than with monetary aggregates, although the results are sensitive to specification and vary somewhat across countries.[21] Time series analysis, especially of vector-autoregressive or moving-average processes, is extremely sensitive to specification.[22] Since this kind of analysis does not derive from theory, it confers plenty of opportunity to torture the data. As Ronald Coase has said, the data will confess if tortured long enough.

Thus neither reason nor evidence suggests that GNP is more closely related to money than to other assets. "Having the closest relation to GNP" does not appear to be an intrinsic property of money.

This conclusion does not imply that a broad credit aggregate should be adopted as a target of monetary policy. The observed relations between GNP and the various aggregates depend, in part, on the monetary policy in effect during the observation period. There is no reason to believe that similar relations would be observed under a different policy. The perceived policy of the central bank affects the public's expectations, which in turn affect the public's behavior. This behavior generates the data and is unlikely to be invariant to different policy regimes.[23]

The velocity of the monetary base, $V_B$, is related to the velocity of M1, $V_1$, by $V_B = V_1/m$, where $m$ is the "money multiplier" M1/MB. Although $m$ is not constant, its variation can affect GNP. Therefore, in principle, $V_B$ could be as stable as $V_1$. Who knows? Perhaps a policy of stable growth in the monetary base would so sufficiently stabilize (or avoid destabilizing) GNP that the base would have the closest relation to GNP. Then the proposition

20. I am indebted to John Wood for this observation.

21. Geoffrey E. J. Dennis, *Monetary Aggregates and Economic Activity: Evidence from Five Industrial Countries*, BIS Economic Papers, no. 7 (Basle: Bank for International Settlements, Monetary and Economic Department, 1983).

22. For a demonstration of this point see Richard D. Porter and Edward K. Offenbacher, "Empirical Comparisons of Credit and Monetary Aggregates Using Vector Autoregressive Methods," *Economic Review*, Federal Reserve Bank of Richmond, November/December 1983, 16–29.

23. "The error lies in using empirical data accumulated in a history when there existed no policy rule as evidence for or against the efficiency of such a rule, had such a rule been in existence. Do we really want to assume that individual behavior in monetary matters would remain invariant as between two quite distinct monetary regimes?" James M. Buchanan, "Monetary Research, Monetary Rules, and Monetary Regimes," *Cato Journal* 3 (Spring 1983): 143. In speaking of a "policy rule," Buchanan is referring to the type of monetary policy that keeps money growing at a constant rate. His remarks about behavior under different monetary regimes are a clear application of a more general proposition enunciated by Robert E. Lucas, Jr., "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and Labor Markets*, ed. Karl Brunner and Allan H. Meltzer, Carnegie-Rochester Conference Series on Public Policy, vol. 1 (Amsterdam: North-Holland Publishing Company, 1976), 19–46.

"Money has the closest relation to GNP of all the financial aggregates" would be a *theorem*. It does not appear to be a theorem in our present economy.

6. *Stability of the demand function*. Although stability of the demand function cannot be imposed as a defining criterion of money, it is an interesting subject of research and a very important one for those who support some kind of utilitarian monetary policy, whether Keynesian or monetarist.[24] Owing, however, to the dominant practice in empirical research of identifying money as M1 or M2, the demand for the monetary base MB has received insufficient study to permit a confident judgment of its stability. The results published so far seem favorable. James Lothian reported a stabler demand for MB than for M1 or either of M1's components in international cross sections for the period 1952–66. Bruce Brittain found both the velocity and holdings of MB to be better explained than those of M1 by a few variables, including foreign interest rates, over the period 1960–79. Gillian Garcia and Simon Pak estimated the demand for currency over the period 1952–67 and found it able to forecast currency holdings during 1974–76 with substantially smaller errors than those that plagued forecasts of deposits or M1 holdings.[25] But the jury is still out because these studies might not be free of the methodological problems of money demand so clearly explained by Thomas Cooley and Stephen LeRoy.[26]

It is likely that empirical study of the demand for MB requires better theory than is now popular. The public's demand for currency almost certainly responds in a different manner to its determinants—and certainly to different determinants—than does the banking system's demand for reserves. Empirical estimates of the aggregate demand for MB ought not to merge these different demand functions into a single estimating equation as if everything were linear.

Nor should demand for MB be estimated by adding together separate estimates of the public's demand for currency and the banks' demand for reserves as if these demands were independent. The two demands are part of a simultaneous system that includes the supply of and demand for deposits (and the central bank's supply of reserves and currency). Deposit pricing affects the public's desired currency–deposit ratio; changes in this ratio, especially sharp changes, interact with the banks' desired reserves. If banks are bound by a required-reserve constraint, their demand for reserves is derived simply from the public's demand for deposits. If banks are not bound by this constraint, their demand for reserves is no longer simply derived from the public's demand for deposits but depends also on interest rates and the probability distribution of net currency withdrawals: if the public is trying to increase its ratio of currency to deposits, then the banks will try to increase their reserve ratio.[27] In such cases the banks' demand for their part of the base depends on the public's demand for its part. The two demands interact.

Most of the time during recent decades banks have held negligible excess reserves. Their demand for reserves has been a very stable function of its

24. Classical liberals would prefer to abolish central banks and take their chances with private arrangements, even if the additional freedom meant less prosperity or stability—although it is possible that both prosperity and stability would increase along with freedom. See F. A. Hayek, *Denationalisation of Money*, Hobart Paper Special 70, 2d ed. (London: Institute of Economic Affairs, 1978); and Lawrence H. White, *Free Banking in Britain* (Cambridge: Cambridge University Press, 1984). Milton Friedman, too, would dismantle the central bank but would not go so far as to take the issuance of money away from government; see *A Program for Monetary Stability* (New York: Fordham University Press, 1960).

25. James R. Lothian, "The Demand for High-Powered Money," *American Economic Review* 66 (March 1976): 56–68; Bruce Brittain, "International Currency Substitution and the Apparent Instability of Velocity in Some Western European Economies and in the United States," *Journal of Money Credit, and Banking* 13 (May 1981): 135–55; and Gillian Garcia and Simon Pak, "The Ratio of Currency to Demand Deposits in the United States," *Journal of Finance* 34 (June 1979): 703–15.

26. Thomas F. Cooley and Stephen F. LeRoy, "Identification and Estimation of Money Demand," *American Economic Review* 71 (December 1981): 825–44. The argument of Cooley and LeRoy is summarized in section 6 of the 1984 article.

27. This is what happened in the early 1930s, owing to bank failures (see Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867–1960* [Princeton: Princeton University Press for National Bureau of Economic Research, 1963], chaps. 7–9), although a different interpretation has been advanced by Peter Frost and others. See Peter A. Frost, "Banks' Demand for Excess Reserves," *Journal of Political Economy* 79 (July/August 1971): 805–25, and the references there cited.

arguments (though the arguments themselves have been unstable). This fact, together with the Garcia-Pak finding of a stable currency demand, suggests that the demand for MB might be the stablest of all the demands for monetary aggregates whenever the public does not fear losses from bank failures. It remains to be seen whether a theoretically sound demand for MB was stable during the depression of the 1930s.

7. *Temporary abode of purchasing power.* Money necessarily has the ability to serve as a temporary store of value during the interval between its receipt and its expenditure. It is not the main store of value even for short periods. Transaction balances, time deposits, and money market assets such as Treasury bills store far more value. Yet none of these assets possesses the absolute liquidity of money, and because nothing confers freedom of mobility like absolute liquidity, the majority of currency appears to be held for this freedom. The public holds about $150 billion in currency, more than $600 for each man, woman, and child in the United States (though foreigners hold some of it). Less than a third of this staggering sum appears to be in use as exchange media.[28] Some part of the remainder has undoubtedly been used to light cigars, but most of it is apparently hoarded against civil disturbances, natural disasters, and nasty divorces. More than 90 percent of $100 notes and 80 percent of $50 notes are estimated to lie in hoards (Anderson 1977; Whitehead 1982). Since 1980, estimated hoards of all denominations have remained an almost constant 3 percent of yearly GNP (while the ratio of currency in use to GNP has decreased by a third).

These estimates—and even the raw figures for per capita currency outstanding—never cease to amaze economists, who immediately start calculating the implied opportunity cost. This cost is indeed impressive (hoardings of $100 billion cost $10 billion *a year* when interest rates are 10 percent), but it just means that freedom of mobility has value.

As currency is now about 80 percent of the monetary base, and some two-thirds of currency is hoarded, more than half of money is hoarded.[29] Hoards are not "temporary" stores of value. Only the part of the base that circulates actually serves as a temporary abode of purchasing power, and it is less than half. Emphasis on temporary abodes throws no light on money.

8. *Exogeneity.* The monetary base is exogenous in the narrow sense of being beyond the control of private parties. It is formally under the control of the central bank, which is usually regarded as an exogenous factor with respect to the economic system. In this narrow sense, emphasis on exogeneity leads to a correct identification of today's money.[30] In a broader sense, the central bank is endogenous to the politico-economic system, and its behavior can be modeled with the same tools that serve for strictly private parties.[31] Whether the central bank—and therefore the monetary base—is best regarded as endogenous or exogenous need not be decided here. Even if we agreed to regard it as exogenous, we could not take exogeneity as a defining characteristic of money. That our money is exogenous (in the narrow sense) would be a theorem, true today but not true at all times. Money was not exogenous when it was gold. Individual countries gained or lost money when their prices fell or rose relative to those of other countries, and the whole world gained money when the production of gold grew profitable. Besides, the acceptance of exogeneity as either a defining characteristic or a theorem true of all future moneys would mean that a system of privately produced money is a contradiction in terms. However hard the essentials of

28. See Paul S. Anderson, "Currency in Use and in Hoards," *New England Economic Review*, Federal Reserve Bank of Boston, March/April 1977, 21-30; and David D. Whitehead, "Explaining the Cash Explosion," *Economic Review*, Federal Reserve Bank of Atlanta, March 1982, 14-18. Anderson argues that very little of the extraordinary per capital total of outstanding currency can reasonably by attributed to illegal activities. On this question see also Robert D. Laurent, "Currency and the Subterranean Economy," *Economic Perspectives*, Federal Reserve Bank of Chicago, March/April 1979, 3-6.

29. Hoarded money does not circulate, but it does not thereby cease to be money. It is hoarded *because* it is money.

30. James Tobin, "Redefining the Aggregates: Comments on the Exercise," in *Measuring the Money Aggregates*, Compendium of Views Prepared by the Subcommittee on Domestic Monetary Policy of the House Committee on Banking, Finance and Urban Affairs, 96th Cong., 2d sess., 1980, Committee Print 96-10, 317-26.

31. See John H. Wood, "A Model of Federal Reserve Behavior," in *Monetary Process and Policy: A Symposium*, ed. George Horwich (Homewood, Ill.: Richard D. Irwin, 1967), 135-66.

such a system may be to grasp, it does not appear to be self-contradictory.[32]

### Unanswered questions

We have answered the question, What is money today? by converting Shackle's emphasis on simultaneous payability and Yeager's emphasis on routine circulation into axioms that collectively define money far enough to identify it as the monetary base. Although the answer is sufficient to the question, both it and the analysis that yielded it are too specific to be a satisfactory account of money *per se*. As pointed out earlier, the analysis is indecisive in case all public transaction balances are issued by a private monopoly bank that holds reserves in currency or other media issued by government. This is not the only indication of a need for additional axioms. The need is especially strong when we try to understand the monetary systems of the United States and the United Kingdom as they were when privately issued bank notes circulated and as they might be if their official monetary authorities were to disappear.

Private banks had the freedom to issue notes until well into the 19th century,[33] and one of the controversial questions in the monetary thought of the time was whether the notes were money. The Currency School, composed chiefly of followers of David Ricardo, regarded the notes as money; the Banking School, led by Thomas Tooke, regarded them not as money but as circulating credit.[34] (Both schools, indeed nearly all economists then writing, agreed that transaction balances were circulating credit and not money.) The Currency School made little attempt to show any economically relevant distinction between notes and balances, as if it were too obvious to require explanation. The Banking School, however, especially its leader, took great pains to trace out the effects of private issues of notes and deposits and to show that they were economically identical. On this point, at least, the Banking School was surely the more persuasive.[35] Yet our present axioms force us to disagree, for all notes can change hands simultaneously. It is true that attempts to convert notes issued by one bank into notes issued by another bank would induce the same sort of reserve transfers that follow check payments involving two banks, thus creating the same lack of simultaneous payability that inheres in transaction balances, but conversions lie beyond our analysis because neither axiom addresses them. Our present analysis forces us to identify private bank notes as money even though they are economically indistinguishable from intangible balances. This unacceptable result proves the incompleteness of our definition of money.

The need to complete our definition is especially strong when we try to understand what a monetary system would be like that lacked a central monetary authority (whether king, treasury office, or central bank). Suppose everyone had the freedom to issue media of exchange or means of payment but no one had legal authority to force them upon sellers of goods or securities. Would such free competition produce anything that ought to be called money? If so, what axioms, besides Shackle's and Yeager's, would describe it? Most important, how would such an economic system behave? Would it be subject to the same macroeconomic disturbances that plague

32. See Hayek, *Denationalisation of Money*, for example.

33. White, *Free Banking in Britain*, gives a beautiful account of the British experience. Arthur J. Rolnick and Warren E. Weber, "Free Banking, Wildcat Banking, and Shinplasters," *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall 1982, 10–19, offer an interesting reinterpretation of American experience.

34. See Rist, *History of Monetary and Credit Theory*, chaps. 4 and 5, for a most interesting account of the war between these schools.

35. See Rist, *History of Monetary and Credit Theory*, 203–9, for a summary. The definition of money was the least important of the controversies between the Banking and Currency Schools. Far more important was the Currency School's belief in the sufficiency, for purposes of stabilization policy, of control of the quantity of money.

36. Wicksell, *Interest and Prices*; Hayek, *Denationalisation of Money*; Benjamin Klein, "The Competitive Supply of Money," *Journal of Money, Credit, and Banking* 6 (November 1974): 423–53; Fischer Black, "Banking and Interest Rates in a World Without Money," *Journal of Bank Research* 1, no. 3 (Autumn 1970): 8–20, and "Bank Funds Management in an Efficient Market," *Journal of Financial Economics* 2 (December 1975): 323–39; Fama, "Banking in the Theory of Finance"; White, *Free Banking in Britain*; Leland B. Yeager, "Stable Money and Free-Market Currencies," *Cato Journal* 3 (Spring 1983): 305–26; and Robert L. Greenfield and Leland B. Yeager, "A Laissez-Faire Approach to Monetary Stability," *Journal of Money, Credit, and Banking* 15 (August 1983): 302–15.

our present system? These questions, as difficult as any in economics, have been wrestled with by Knut Wicksell, Friedrich Hayek, Benjamin Klein, Fischer Black, Eugene Fama, Lawrence White, Leland Yeager, and Robert Greenfield.[36] Though they are not to be answered by a definition of money, a definition is almost certainly a psychological prerequisite. How else can we explain the effort devoted to its search?

---

# Analyzing Deficit Finance in a Regime of Unbacked Government Paper

John Bryant

*Fox Professor of Economics*
Rice University
*Consultant*
Federal Reserve Bank of Dallas

The U.S. federal budget has been in deficit almost continuously since the 1950s. Moreover, as has been widely reported, this deficit has recently increased substantially both in dollar terms and as a percentage of gross national product.[1] Usually it is assumed that when an economic agent borrows, it will repay in the future. If this economic agent is a government, it will run a surplus in the future, either from future tax collections or from the return on its investments, to pay back principal and interest.

However, a case can now be made that the Federal Government does not intend to stop borrowing, much less repay. It is not engaging in extensive and highly profitable investment, and massive future tax increases do not seem likely. This does not imply that the Federal Government is unreliable or unstable. It also does not imply that the government is a dishonest borrower, with the intent of repudiating its debt. Instead, it simply suggests that the government is following a policy of permanent deficit finance.

The Federal Government is able to sustain this policy because the public accepts its unbacked paper. Paper issued by most economic agents is "backed" by real assets, by an explicit commitment to maintain convertibility (as with currency under a commodity standard), or by an implicit commitment to a standard of fiscal discipline. However, the Federal Government can issue paper that is backed neither explicitly by real assets nor implicitly through future taxes or returns on investments.

This ability gives the Federal Government a unique role in the economy. All levels of government are special economic agents. They each can raise revenue through taxes, whereas businesses must ultimately rely on sales. However, only the Federal Government can issue unbacked paper. Indeed, from a macroeconomic perspective, this is the important characteristic that distinguishes the Federal Government from state and local governments. The Federal Government's freedom from the discipline of having to balance its budget over the

*The views expressed are those of the author and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.*

1. See Leroy O. Laney, "The Strong Dollar, The Current Account, and Federal Deficits: Cause and Effect," *Economic Review*, Federal Reserve Bank of Dallas, January 1984, 1-14.

long haul explains, in part, why more attention is focused on the fiscal and monetary policy of the Federal Government than on the activities of state and local governments.

Ever since the United States went off the gold standard for currency, the Federal Government has had the explicit authority to issue unbacked paper in the form of Federal Reserve notes. For several decades, however, this authority has not been fully exercised. Currency has not been backed by any commodity, but bonds have been partially backed by future tax revenues. What seems to have emerged recently is an increased willingness to issue bonds that it seemingly intends to roll over forever.[2] That is, the government will redeem old bonds by issuing new ones and, in turn, will redeem those by issuing yet more bonds, and so on forever.

Insofar as the government intends to roll the bonds over forever, they are backed neither by future taxing authority nor by investments. The bonds are promises to pay currency, and the currency to pay off the bonds will be obtained through future bond sales or, possibly, issuance of additional currency, rather than through taxes or sales of goods. The bonds are indeed unbacked paper, just like the explicitly unbacked Federal Reserve notes.

This puts the economy in a new situation; we are in a new economic regime that differs from the old in an important respect. The economy includes an economic agent (a large one), the Federal Government, that borrows without promise of future surplus. How does this change in regime matter from a macroeconomic perspective? In other words, how will our economy's aggregate allocation of goods and services, and inflation, be affected?

This article demonstrates analysis of fiscal and monetary policy under such a regime with a model that explicitly incorporates the Federal Government's unique position as the sole issuer of unbacked paper. An important element of the model is the Federal Government's monopoly on the issuance of unbacked paper. Many of the model's implications arise because this new regime of unbacked paper is amenable to standard modes of analysis of monopoly behavior.

To pursue this analysis, it is useful to consider separately the consequences for fiscal and monetary policy of the new economic regime of unbacked government paper. Fiscal policy—government spending and taxing policies—determines the volume of unbacked paper. A deficit increases the volume of unbacked paper, and a surplus reduces it. The composition of unbacked paper between currency and government bonds is determined by monetary policy. Actually, fiscal and monetary policy are not completely separate. The interest bill of government bonds is a substantial portion of expenditures, and the composition of unbacked paper between currency and bonds affects the interest bill. Nonetheless, it is useful to treat fiscal and monetary policy separately.

The basic result of the inquiry is that fiscal deficits cause inflation but appropriate monetary policy can mitigate the distortions caused by inflation. This result is obtained by first examining the effects of deficits financed entirely through currency expansion, which produces distortions analogous to the inefficiencies observed when a monopolist maximizes profits while charging the same price for each unit of output. Next, the effects of financing deficits with a mixture of currency and bonds are analyzed. By paying interest on bonds but not on currency, the government collects less "seigniorage" from each dollar's worth of bonds than it collects on each unit of currency. This enables the government to behave as a price-discriminating monopolist and thereby reduce the inefficiencies of single-part pricing. The inquiry begins with an analysis of unbacked paper.

### A simple model

The first question that must be addressed is why the private sector accepts the government's unbacked paper. Why is the worth of unbacked currency or bonds more than the value of the paper on which they are printed? The private sector accepts the government's unbacked paper either because there is a useful role for the government's unbacked paper in the economy or because the government imposes enough regulations to force a demand for its unbacked paper.[3] For the purposes of this article, it suffices to assume that unbacked government

2. See W. Michael Cox, "What Is the Rule for Financing Public Debt?" *Economic Review*, Federal Reserve Bank of Dallas, September 1984, 25-31.

3. This legal restrictions view of unbacked paper appears in John Bryant, "A Clower Constraint Model of Unbacked Money," *Journal of Banking and Finance*, forthcoming.

paper performs a useful function not provided by any commodity supplied by the private sector.

Paul Samuelson's overlapping-generations model is useful to illustrate the special properties of unbacked government paper.[4] This model isolates the notion that unbacked currency is only accepted by someone because she believes that someone else will accept it later, an individual who, in turn, accepts it only because she believes someone else will accept it yet later, and so on. Moreover, agents have some confidence in how much they will receive in the future in exchange for their currency holdings. This self-fulfilling prophesy, which "drives" the overlapping-generations model, must be the crucial property of valued unbacked currency. The model also has the advantage of yielding implications about how the utility of individual citizens is affected by changes in government policy. Most traditional models predict effects on aggregate income and inflation, but the effects on the welfare of people are not explicitly incorporated.

**Assumptions.** The particular model used here is simplified in a number of respects. The economy consists of individuals who live for two periods. Except for date of birth, they are all identical. Therefore, conclusions concerning the behavior of one individual extend to the economy as a whole. There is no production, only trade. Because a major concern of the analysis of unbacked currency is the effects of inflation on trade, this is a useful simplification.

Each individual is endowed with $X$ units of a single consumption good when young and with nothing when old. The consumption good is tradable but not storable. At the beginning of each period, the old people die, the young people become old, and the same number of young people are born. Therefore, in any given period there are equal numbers of young and old people. The young trade some of their endowment to the old for currency. In the next period, when the young become old, they trade the currency for goods with the next generation of young people.

Each individual makes a complete trade just once in a lifetime. An individual's trading of goods for currency and then currency for goods amounts to a single trade of goods "today" for goods "tomorrow." This is, of course, exactly what accepting unbacked currency allows people to do: exchange goods now for goods later.

**The decision.** The individual must choose the amount of her endowment to trade for currency when she is young. When old—her action is obvious—she trades all her currency for goods. In choosing how much currency to hold, the individual is deciding on the allocation of consumption over her lifetime. Her decision depends on her tastes for consumption now versus consumption later and on the rate at which she can exchange goods today for goods tomorrow. This rate of exchange will be called the gross real rate of return, equal to 1 plus the net real rate of interest. With no inflation, the gross real rate of return on currency, $R_m$, is 1. It rises with deflation and falls with inflation.

The individual picks whatever amount of "real balances" (real value of currency) provides her preferred allocation of consumption over her lifetime. Let us call this amount of real balances $m^*$.

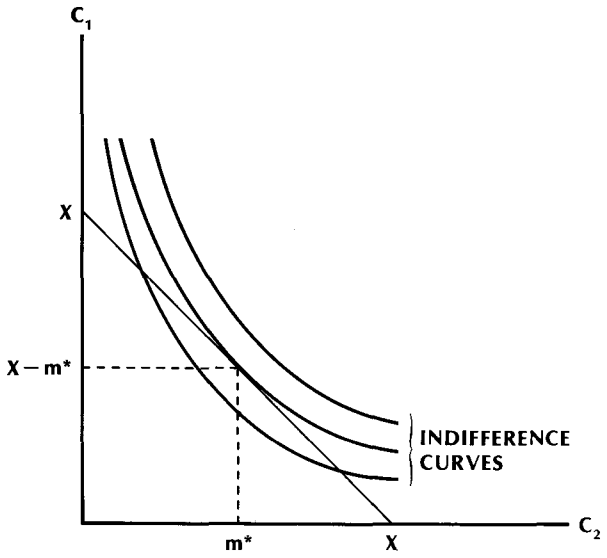
Notice that it is the demand for *real* balances—purchasing power—that is determined, not the demand for nominal, dollar, balances. Therefore, the level of the nominal stock of currency is irrelevant except for determining the price level. For example, if there were twice as much currency, the price level would be twice as high, thereby keeping real balances the same.

The young person's demand for real balances depends in part on her expectation of what the gross real rate of return on currency will be. This rate of return is determined by the inflation rate. The realized inflation rate will be influenced by the demand for real balances by the next generation of young people, demand that, in turn, depends in part on their expectation of the inflation rate they will encounter, and so on. For the market to be in equilibrium, expected inflation must equal realized inflation for each generation.

Suppose, however, that all generations expect no inflation. If all generations have the same tastes, endowments, and expectations, as assumed, each will demand the same real value of currency, $m^*$. With a continuously balanced government budget, there is a constant nominal supply of currency. In each period the price level is the nominal supply of money per young person divided by $m^*$, the demand

4. Paul A. Samuelson, "An Exact Consumption-Loan Model of Interest With or Without the Social Contrivance of Money," *Journal of Political Economy* 66 (December 1958): 467–82.

Figure 1
**Inactive Government**




$C_1$ is first-period consumption, and $C_2$ is second-period consumption. Stationary consumption bundles are represented by the straight line running from X on the $C_1$ axis to X on the $C_2$ axis. The individual chooses the feasible consumption bundle that is on the highest indifference curve, namely, $C_1 = X - m^*$ and $C_2 = m^*$; $m^*$ is the real value of currency per young person.

In general, with a real gross rate of return on money of $R_m$, the individual's perceived budget set is represented by a straight line running from X on the $C_1$ axis to $R_mX$ on the $C_2$ axis. Therefore, only $R_m = 1$, no inflation, is consistent with stationary consumption when the population is constant. If the population were growing at the rate g, the amount of currency demanded by the successive young generations would increase. The price level would decline at the rate of population growth, so $R_m = 1 + g$.

per young person for real balances. Therefore, with a constant nominal supply of currency and a constant real demand for currency, the price level is constant and there is no inflation. The expectation of no inflation is consistent with equilibrium. Determination of the young person's desired real balances under these conditions is shown in Figure 1.

### Fiscal policy

So far in the analysis the government has been completely inactive. To provide some perspective on fiscal policy in an economy with unbacked currency, first consider the simple case of a balanced budget and lump-sum taxes.[5] The results of this sort of fiscal policy can then be compared with those generated by other taxes and by deficit finance. All government spending is assumed to be transfer payments. This assumption is adopted because the composition of government expenditure is not a consideration here. Exactly analogous analysis applies to other forms of government expenditure.

A commonly used simplifying assumption in macroeconomics is that the effects of the redistributions caused by taxes and transfer payments cancel out in the aggregate. This occurs, for example, when the recipients of transfer payments increase savings by the amount that the taxed individuals reduce savings. A simple way to build this feature into a model is to assume taxes and transfer payments accrue to the same individual.

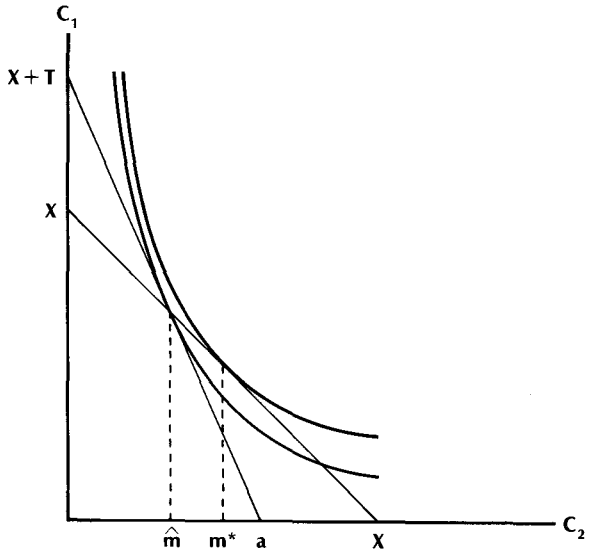
**Balanced budgets.** Suppose, for example, that the government in our model taxes each young individual $T$ units of real value of currency (or $PT$ "dollars," where $P$ is the price level) and makes a transfer payment to each young person of the same amount. This fiscal policy obviously has no effect, as the government is just taking with one hand and returning with the other. The tax and transfer cancel out. The first result, then, is that transfer payments financed by lump-sum taxes are inconsequential from a macroeconomic perspective.

Next, consider the more realistic case of transfer payments financed by taxes that do distort the choices faced by the individuals. Suppose the government charges a flat-rate tax on the holding of currency. As the amount of currency to hold is the individual's only decision, this is the only sort of distorting tax that can be used without complicating the analysis. As before, this tax on currency finances the transfer payment of $T$ units of real value of currency to each young person.

The government must choose the tax rate, $t$, so that tax revenues will equal expenditures. In terms of the before-tax real value of currency per young person in the current period, $m$, $t$ and $m$ must satisfy $tm = (1 - R_m)m = T$. Taxes must be based on correctly estimated currency holdings. That estimate must take into account the transfer payment and the effect of the tax on the young person's perception of her budget set. If the after-tax real value of

5. For an analysis of fiscal policy in an economy without unbacked government paper, see John Bryant, "How Fiscal Policy Matters," *Economic Review*, Federal Reserve Bank of Dallas, January 1984, 15–20, and "Government Irrelevance Results: A Simple Exposition," *American Economic Review* 73 (September 1983): 758–61.

Figure 2
**Taxes on Currency**




A tax on currency and a government transfer payment do not affect the set of feasible stationary consumption bundles, which is still represented by the straight line running from X on the $C_1$ axis to X on the $C_2$ axis. However, once the government sets the transfer payment (T) and the tax rate (t), the individual's perceived budget set will be represented by the straight line running from $(X + T)$ on the $C_1$ axis to $a = (1 - t)(X + T) = R_m(X + T)$ on the $C_2$ axis. The perceived budget set must induce the individual to choose a technologically feasible consumption bundle. Therefore, t will be chosen so that the consumption bundle in the budget set that is on the highest indifference curve is also feasible, as illustrated. This causes the individual to reduce second-period consumption from $m^*$ to $\hat{m}$.

currency per young person in the next period is $\hat{m}$, $\hat{m} = (1 - t)m$, then $t$ must be chosen so that the resulting $\hat{m}$ satisfies $[t/(1 - t)]\hat{m} = T$ (Figure 2).

Once again, no inflation is consistent with equilibrium. In this noninflationary equilibrium the real rate of interest on currency is not zero, but it is minus $t$. In other words, the gross real rate of return on currency, $R_m$, is $(1 - t)$. The tax reduces the incentive to hold currency, increases the price level, reduces second-period consumption, and makes the individual worse off. The tax on currency is, in effect, a tax on trade and therefore discourages trade.

**Currency-financed deficits.** Now suppose that instead of taxing, the government finances the transfer payment by continuously issuing more currency. To make transfer payments, the government simply prints the required currency and hands it to young individuals. The government continuously runs a current account deficit. The resulting continuous inflation and erosion of the value of currency act exactly as an explicit tax on currency.

The above analysis of an explicit flat-rate tax on currency carries over directly to this form of deficit finance. Instead of a constant price level, inflation at the rate $t/(1 - t)$ is a self-fulfilling prophesy, and instead of an explicit tax rate of $t$, there is an implicit inflation tax rate of $t$. That is, $[t/(1 - t)]\hat{m} = T$ or $tm = (1 - R_m)m = T$ still has to hold (Figure 2). The inflation reduces the incentive to hold currency, discourages trade, and makes the individual worse off.

Just as transfer payments financed by nondistorting taxes are equivalent to an inactive government (from a macroeconomic perspective), so are deficit-financed transfer payments equivalent to transfer payments financed by distorting taxes on currency. In using deficit finance, the government is inefficiently extracting rent on its monopoly of currency, just as if it charged directly for the use of its currency.

**Fiscal policy shifts.** In the model, changing from a budget balanced with an explicit flat-rate tax on currency to a policy of deficit financing has no real consequences, permanent or temporary. Not only are the stationary consumptions the same, but so are the real holdings of currency. However, this result does require that individuals perfectly foresee the implicit inflation tax generated by the currency issue.

Suppose, instead, that individuals at first ignore the inflationary consequences of deficit finance. In particular, suppose the government, which had been taxing currency explicitly, stops doing so and instead issues additional currency to make transfer payments.[6] Moreover, suppose that the first generation in this new "regime" sees the explicit tax disappear but ignores the implicit inflation tax. These individuals continue to assume no inflation. For simplicity of analysis, it is also assumed that the subsequent generations do correctly understand the inflationary consequences of deficit finance and that previous generations did not anticipate the policy change.

6. See Bryant, "How Fiscal Policy Matters," for an analysis of the effects of beginning to issue bonds to finance the transfer payments.

The first generation in this new regime, by failing to take future inflation into account, overestimates the rate of return on currency. Assuming current and future consumption are substitutes, the implication is that current consumption is reduced and demand for currency is increased. As a result, the price level drops temporarily. Suppose the first generation expects this lower price level to be sustained. When these individuals subsequently sell the unbacked paper to the next generation, a wiser one, they are unpleasantly surprised. The next generation, which has correctly foreseen the subsequent inflation, does not have a raised demand for currency. As a consequence, the price level jumps and then continues to rise thereafter at a (lower) constant rate (Figure 3).

**Applicability.** In reality, of course, direct taxes are not taxes on currency. What does carry over to the real world is the notion that in changing from tax financing to deficit financing, the government is trading explicit distorting taxes for an implicit distorting tax—inflation. Moreover, in changing to implicit inflation tax, the government may further distort economic decisions, and prices, by fooling people.

### Monetary policy

Financing deficits entirely by issuing currency has an unfortunate consequence: the resulting implicit tax on currency is distorting. The inflation causes people to consume more in the first period and less in the second period, and it reduces the demand for currency. The question that immediately arises is, Can the government reduce or eliminate the distortion of inflation by financing at least part of the deficit by issuing interest-bearing bonds? This possibility introduces a role for monetary policy. Under certain conditions, monetary policy will allow deficits to be financed with the equivalent of nondistorting, lump-sum taxes. Monetary policy amounts to two-part pricing by the government unbacked-paper monopoly.

**The special case of perfect price discrimination.** To illustrate this point, the model used so far will be expanded to allow for the coexistence of currency and interest-bearing government bonds. This approach requires further attention to the motive for using currency. With interest-bearing bonds available as an alternative store of value, currency must have some special property if there is to be any demand for it.

The special property of currency in this model is that it will be the only store of value available in small denominations. Bonds will only be issued in large denominations, and private substitutes for currency, such as bank deposits or money market mutual fund shares, will be outlawed.[7]

For simplicity, assume that bonds are issued in

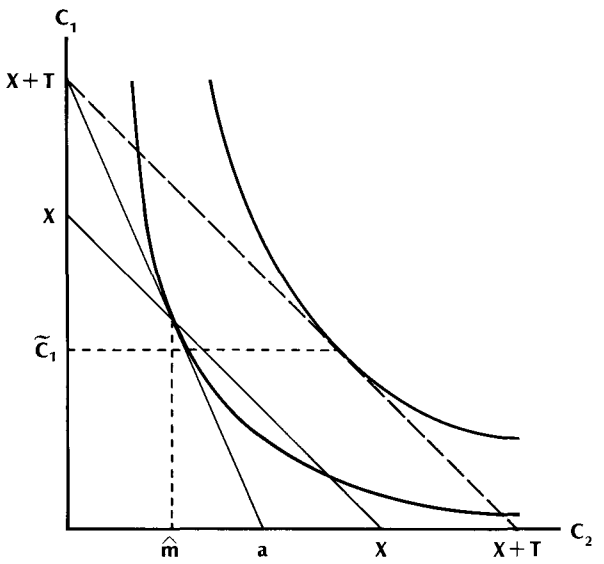
## International Finance

The distorting inflation tax view of deficit finance has important implications for international finance. Governments' fiscal policies can involve an inefficient beggar-my-neighbor "game."[1] Suppose, for example, that governments peg constant exchange rates; in effect, there is a single world currency. However, no one government is a monopolist in the provision of this currency. Rather, the currency is provided by the oligopoly comprising the various governments. As a consequence, the governments face the standard enforcement problem of cartels. Suppose the governments finance transfer payments by issuing currency. This action, as analyzed in the section on fiscal policy, amounts to imposing a distorting inflation tax on the world currency.

However, if one country uses deficit finance while the others balance their budgets, this deficit country is levying an inflation tax on the entire world currency. This amounts to an inefficient transfer of wealth from the fiscally conservative countries to the deficit country. The deficit country is handing increments of the world currency to its citizens, while the other countries are not doing so. Although the entire world might be better off with balanced budgets than with universal deficits and inflation, each country is best off being the only deficit country in an otherwise fiscally conservative world. In other words, a stable international monetary system requires coordinated fiscal policies.

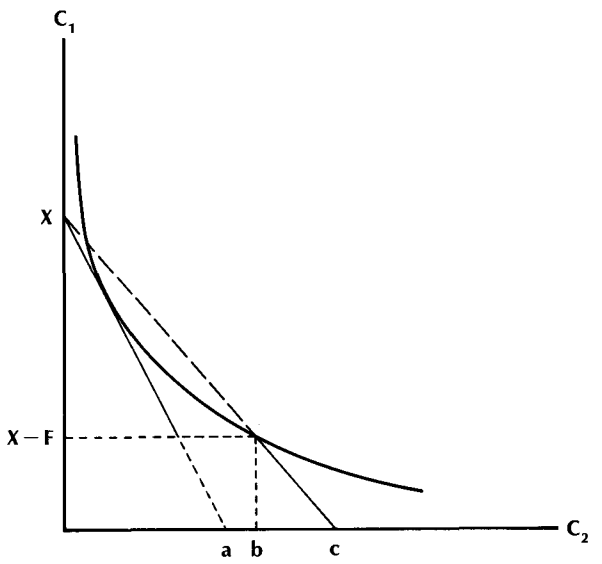
1. This discussion draws on John Kareken and Neil Wallace, "On the Indeterminacy of Equilibrium Exchange Rates," *Quarterly Journal of Economics* 96 (May 1981): 207-22, and "International Monetary Reform: The Feasible Alternatives," *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1978, 2-7.

Figure 3
**Unanticipated Inflation**




An inflation "tax" does not alter the set of feasible stationary consumption bundles, which is still represented by the straight line running from X on the $C_1$ axis to X on the $C_2$ axis, nor the perceived budget set of the individuals anticipating the implicit inflation tax, which is still represented by the straight line running from (X + T) on the $C_1$ axis to a = (1 − t)(X + T) = $R_m$(X + T) on the $C_2$ axis. However, if the first generation ignores inflation, its (mis)perceived budget set runs from (X + T) on the $C_1$ axis to (X + T) on the $C_2$ axis. These individuals overestimate the rate of return on currency; they assume a gross real rate of return on currency of 1. Therefore, they decrease their current consumption, if current and future consumption are substitutes, and increase their holdings of currency. However, the unanticipated inflation ultimately erodes the purchasing power of that paper to $\hat{m}$.

Figure 4
**Bonds**




With bonds available, the individual's perceived budget set is represented by a broken line running from X on the $C_1$ axis to c = $R_b$X on the $C_2$ axis, where $R_b$ is the gross real rate of return on bonds. The break occurs at $C_1$ = X − F, where F is the minimum real purchase price of bonds. If the bonds did not have a minimum real purchase price, the individual's perceived budget set would be represented by the straight line running from X on the $C_1$ axis to c = $R_b$X on the $C_2$ axis. If bonds were not available, the individual's perceived budget set would be represented by the straight line running from X on the $C_1$ axis to a = $R_m$X on the $C_2$ axis. The gap between the gross real rates of return on bonds and money is $R_b - R_m = i(R_m)$, where i is the net nominal rate of interest on bonds. As the bonds are sold at whatever discount the market will bear, this gap must make the individual indifferent between currency and bond holdings, as illustrated.

any denomination above some fixed minimum. These bonds are sure Treasury promises of currency in the next period. Let the gross real rate of return on bonds be $R_b$. Then $R_b$ is greater than $R_m$, the gross real rate of return on currency, because bonds bear nominal interest. Also, assume that all intermediation of bonds is illegal. (This assumption precludes private, interest-bearing substitutes for currency. In reality, restrictions such as prohibition of private note issue and imposition of reserve requirements on checkable deposits impede, but do not eliminate, intermediation.)

Under these assumptions, an individual must choose between holding currency or bonds between periods. The disadvantage of holding currency is that it pays no interest. The disadvantage of holding bonds is the minimum denomination. One may have to hold a larger value of bonds (and thereby defer more consumption to old age) than would normally be optimal when the prevailing rate of return is $R_b$. This approximates the situation in the United States, where Treasury bills bear interest but cannot be bought with a face value of less than $10,000.

7. This legal restrictions view is exposited in John Bryant and Neil Wallace, "A Price Discrimination Analysis of Monetary Policy," *Review of Economic Studies* 51 (April 1984): 279-88; Thomas J. Sargent and Neil Wallace, "The Real-Bills Doctrine Versus the Quantity Theory: A Reconsideration," *Journal of Political Economy* 90 (December 1982): 1212-36; and Neil Wallace, "A Legal Restrictions Theory of the Demand for 'Money' and the Role of Monetary Policy," *Federal Reserve Bank of Minneapolis Quarterly Review*, Winter 1983, 1-7.

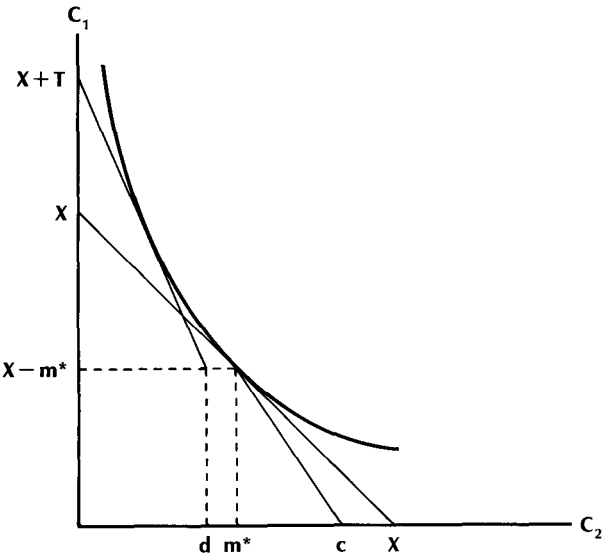
Determination of the young person's desired real bond holdings under these conditions is shown in Figure 4.

Suppose, once again, that the government wants to finance a transfer payment of $T$. However, instead of just printing currency or taxing, the government finances the transfer payment by selling bonds. As the model assumes identical individuals, also assume that the government sells one bond per young person. The sequence of transactions could occur as follows: (1) the young person trades goods for currency with an old person; (2) the government prints currency and hands it to the young person (the transfer payment); (3) the government sells a bond to the young person for currency; (4) the young person holds the bond between periods; (5) the young person becomes old, her bond matures, and the government prints currency to pay off on the bond. The sequence is repeated with the next generation. In this scenario, bonds completely "soak up" the currency. Currency is only used in trade, and bonds are the unique store of value. (This rigid sequence of transactions need not be followed exactly: the government could simultaneously sell bonds and use the proceeds for transfer payments and payoffs on bonds.)

Suppose the government sets the minimum real purchase price of the bonds (which will be denoted by $F$) equal to the transfer payment ($T$) plus the real value of unbacked government paper that would be held with an inactive government ($m^*$). Thus, $b$ is greater than or equal to ($T + m^* = F$), where $b$ is real bond holdings. For a market equilibrium, the interest rate and inflation must be such that individuals are just willing to hold the bonds.[8] Also, and similar to the case of pure currency financing, the revenue from the implicit tax on bonds must equal expenditures. Therefore, the government sets the gross rate of return on bonds ($R_b$) so that $(1 - R_b)b = (1 - R_b)F = T$.

The consequences are that consumptions are the same as with an inactive government, the holding of unbacked government paper is higher than with a flat-rate inflation tax, and the rate of inflation is lower than with a flat-rate inflation tax (see Figure 5). Bond-financing the deficit has the same real consequences, generates the same consumptions, as lump-sum tax financing. The individual does face inflation, but this inflation does not affect consumption or trade because the minimum purchase "price" of the bonds introduces a discontinuity in her perceived budget set.

Figure 5
**Bond-Financed Deficits**




A discriminatory inflation "tax" and a government transfer payment do not alter the set of feasible stationary consumption bundles, which is still represented by the straight line running from X on the $C_1$ axis to X on the $C_2$ axis. However, the individual's perceived budget set is now represented by a broken line running from $(X + T)$ on the $C_1$ axis to $c = [m^*/(T + m^*)](X + T) = R_b(X + T)$ on the $C_2$ axis. The break occurs at $C_1 = X - m^*$. The gross real rate of return on bonds, $R_b$, is $m^*/(T + m^*)$. The individual can choose to hold currency with gross real rate of return $R_m < R_b$ or, for holdings of $(T + m^*)$ or more, bonds. Notice that $d = R_m(T + m^*)$ on the $C_2$ axis is less than $m^* = R_b(T + m^*)$. The individual chooses to hold one bond at the minimum real purchase price of $(T + m^*)$ and consumes $C_1 = (X + T) - (T + m^*) = (X - m^*)$ and $C_2 = [m^*/(T + m^*)](T + m^*) = R_b(T + m^*) = m^*$.

The bonds amount to a volume discount on unbacked government paper. When the minimum denomination is chosen appropriately, this discount eliminates the distortion of individuals holding too little unbacked government paper. The avoidance of distortions under this policy is analogous to the "efficient" outcome achieved by a perfectly discriminating monopolist practicing multipart pricing.

8. Actually, because everyone is holding a bond and no currency between periods, the nominal interest rate and inflation are indeterminate but have unique minimum values. This is an uninteresting technical feature of such examples.

**Tightening policy.** This simple model of monetary policy suggests that a "tight" monetary policy can generate a one-time price-level drop at the cost of subsequent higher inflation. In this model a tighter monetary policy is one with a larger minimum real purchase price of bonds. In reality, the central bank tightens monetary policy by selling bonds for currency, thereby increasing the number of bonds outstanding. With identical individuals, each holding one bond, the amount of bonds is increased by increasing the size of bonds, not the number of bonds.

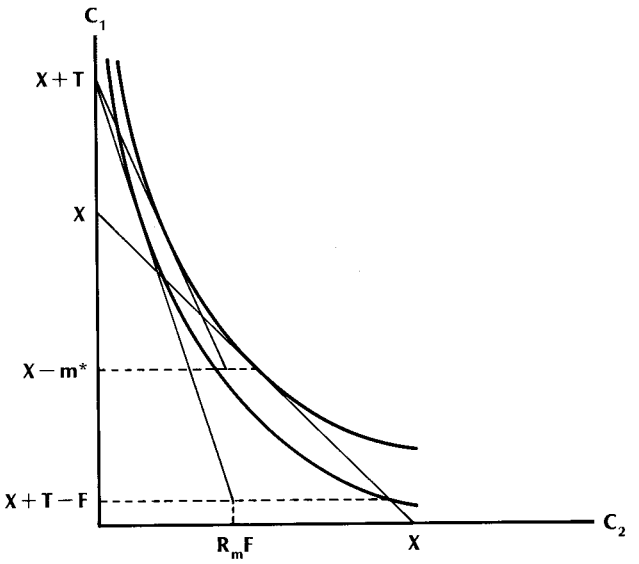
Increases in the minimum real purchase price of bonds result in larger real holdings of unbacked government paper (bonds) and, therefore, a lower initial price level. For minimum real purchase prices of bonds less than $(T + m^*)$, increases in the minimum real purchase price of bonds also slow the rate of inflation, as the distortion of too little second-period consumption is reduced. (Compare Figures 5 and 2.) However, increases in the minimum real purchase price of bonds beyond $(T + m^*)$ result in higher inflation, as too much second-period consumption is encouraged. (See Figure 6.)

While this model clearly distinguishes between price-level and rate-of-inflation effects, observed price data do not make this distinction. A too tight monetary policy would appear in the data as a temporary respite from inflation, followed by higher inflation. Tightening monetary policy always tends to reduce observed inflation at first. Whether it does so in the long run depends on just how tight that policy is.

This simple model of bond-financed deficits is one in which bonds are the only asset used as a store of value. Currency is used only in transactions, which consume no time. A direct real-world analogue would be a bond-backed mutual fund with a minimum deposit restriction and with transactions carried on by checks (which are currency for electronic instants). How closely this approximates the real world is, of course, problematic. Individuals are not identical, and one could not expect monetary policy to achieve perfect price discrimination. Neither could one expect that price discrimination is the only consequence of monetary policy. The simple model really just indicates the potential for multipart pricing of unbacked paper when a government engages in deficit finance.

**Imperfect price discrimination.** The analysis of imperfect price discrimination produces similar results.[9] While explicitly deriving these results is complicated, it is easy enough to motivate them. Let $D$ be the aggregate real deficit (net of interest payments), $M$ be the aggregate real currency holdings, and $B$ be the aggregate real bond holdings. $R_m$ is the gross real rate of return on currency, and $R_b$ is the gross real rate of return on bonds. The higher $R_m$ is, the lower the rate of inflation; the higher $R_b$ is, the higher the real rate of interest on bonds.

Figure 6
**Tight Monetary Policy**




The individual chooses to hold one bond at the minimum real purchase price $F > (T + m^*)$. The upper portion of the individual's perceived budget set is the line running from $(X + T)$ on the $C_1$ axis and terminating at $C_1 = X + T - F$. This line segment is steeper than the line segment running from $(X + T)$ on the $C_1$ axis and terminating at $C_1 = X - m^*$, which is the upper portion of the individual's perceived budget set when the minimum real purchase price of bonds is $(T + m^*)$ (see Figure 5). A steeper upper portion of the perceived budget set means that more $C_1$ must be sacrificed to increase $C_2$ by holding currency, which means a lower gross real rate of return on currency or a higher inflation rate. However, as F is greater than $(T + m^*)$, the real holding of government paper is higher with the minimum real purchase price of F, so the initial price level is lower.

9. Monetary policy as imperfect price discrimination is analyzed in an earlier version of Bryant and Wallace, "A Price Discrimination Analysis of Monetary Policy," which appeared as Federal Reserve Bank of Minneapolis Research Department Staff Report no. 51 (Minneapolis, 1979).

In the earlier cases, with currency financing of a transfer payment and an inflation rate of $t/(1 - t)$, the financing equation was $[t/(1 - t)]\hat{m} = T$; that is, the inflation tax per young person equals the transfer payment. In our new notation this equation is $(1 - R_m)M = nT = D$, where $n$ is the number of young people. With both currency and bonds outstanding, this equation becomes $(1 - R_m)M + (1 - R_b)B = D$. This financing equation means that the revenue from the implicit inflation tax on currency plus the revenue from the implicit interest tax on bonds equals the deficit.

Now, the assumptions that all individuals have the same preferences and endowments can be relaxed. The above financing equation holds generally in a "steady state," even with some individuals holding more than the minimum real purchase price of bonds. If individuals are holding more than the minimum real purchase price of bonds, the monetary policy is not perfectly discriminating with respect to these individuals. There also may be individuals who are holding the minimum real purchase price of bonds but for whom the minimum is too large or too small. None of this affects the validity of the financing equation.

The financing equation, $(1 - R_m)M + (1 - R_b)B = D$, is by itself very informative. A policy that increases the real rate of interest on bonds reduces the rate of implicit tax on bonds, $(1 - R_b)$. However, with the higher interest rate, that policy may also increase the demand for bonds and thereby increase the tax "base" $B$. The net effect on the implicit tax collection from bond holdings, $(1 - R_b)B$, is unclear.

If the elasticity of demand for bonds is high, increasing the real interest rate on bonds can actually raise the implicit tax revenue from bond holdings. (This is exactly analogous to the so-called Laffer curve of direct taxation.) In turn, an increase in the implicit tax collection from bond holdings reduces the collection necessary from currency holdings; that is, it leads to lower inflation. Thus, discriminating against money holders and in favor of bondholders (charging a higher implicit tax on currency than on bonds) can actually make money holders better off by reducing inflation. This point is important because minimum size restrictions on interest-bearing assets can appear to be discrimination against the poor and in favor of the rich. The result is that such price "discrimination" can actually benefit the "discriminated against" group, the poor.

The financing equation also makes clear that there is a point beyond which higher real interest rates cannot, in the long run, reduce inflation. If $R_b$ is greater than 1—that is, the real rate of interest on government bonds is positive—the tax take from bond holdings is negative because the implicit tax rate, $(1 - R_b)$, is negative.[10] The seigniorage the government collects by the bond issue is exceeded by the interest bill, which in effect increases the deficit. Implicit taxes on currency holdings are "subsidizing" both the government and the bondholders. Therefore, inflation rises. Our model assumes an economy of constant size. This result does have to be modified for a growing economy: a real interest rate on government bonds in excess of the rate of growth of the economy causes inflation to rise in the long run.

**Implications.** This price discrimination view of monetary policy has important implications for the deregulation of financial markets and for technological innovations in the financial industry. As with any form of price discrimination, monetary policy price discrimination depends on the ability of the discriminator to separate markets. In the case of monetary policy the discriminator is the central bank, and the markets are the markets for currency and bonds. Deregulation and technological innovations in the financial industry both tend to erode the barriers between the markets for currency and bonds. Both encourage the intermediation of bonds into better substitutes for currency. This intermediation reduces the ability of the central bank to use monetary policy to enhance the demand for government obligations and thereby improve the efficiency of deficit finance. Put simply, whatever their merits, deregulation and technological innovation of the financial industry may fan the fires of inflation.

## Summary

An economy with unbacked government paper is different in important respects from an economy with backed government paper. Unbacked government paper frees a government from balancing budgets, on average, but imposes a distorting inflation tax. However, the distortion of this inflation tax can be mitigated by appropriate monetary policy if

10. Therefore, $(1 - R_b)B$ typically starts falling for some $R_b < 1$.

legal restrictions are imposed to inhibit the intermediation of government bonds. Unfortunately, "tailoring debt issue to the needs of the public" requires a great deal of information. As with any form of price discrimination, demand functions must be known. This, in turn, requires knowledge of the effectiveness with which legal restrictions separate the markets for currency and bonds. Finally, changes in monetary policy regimes can involve misallocation insofar as individuals do not perfectly foresee the consequences of regime changes.

---

FEDERAL RESERVE BANK OF DALLAS
STATION K, DALLAS, TEXAS 75222

ADDRESS CORRECTION REQUESTED

The *Economic Review* is published by the Federal Reserve Bank of Dallas and will be issued six times in 1985 (January, March, May, July, September, and November). This publication is sent to the mailing list without charge, and additional copies of most issues are available from the Public Affairs Department, Federal Reserve Bank of Dallas, Station K, Dallas, Texas 75222. Articles may be reprinted on the condition that the source is credited.