Cowles Foundation Paper 288

Reprinted from THE JOURNAL OF POLITICAL ECONOMY
Vol 76 No 4 Part II July/August, 1968

Notes on Optimal Monetary Growth

James Tobin*

Yale University

The question of the optimal size and rate of growth of the money supply has at least as many meanings as there are definitions of "money." Three possible interpretations of the question are: (1) What are the optimal size and the optimal rate of growth of the central government's deadweight debt to its citizens? (2) What are the optimal size and the optimal rate of growth of the supplies of currency and other means of payment? (3) What is the optimal degree of financial intermediation in an economy, and what is its optimal rate of expansion?

In some models one or two of these interpretations vanish, or merge. An important example is the simplest monetary extension of the standard aggregative neoclassical growth model. In this extension money as government debt and money as means of payment are identical. It is assumed, in other words, that all government debt takes the form of means of payment —either directly as currency or indirectly as demand deposits backed 100 per cent by currency or other government obligations—and that all means of payment are directly or indirectly obligations of the central government. Under these restrictive assumptions, interpretations (1) and (2) merge. At the same time these models ignore private financial markets and intermediaries, so that question (3) does not arise.

I. "Money" as Government Debt

I shall begin with the first interpretation of the question. The crucial property of "money" in this role is being a store of value, an alternative to reproducible productive capital in satisfying the desires of the community to accumulate wealth. If the supply of government debt in real

* The research described in this paper was carried out under a grant from the National Science Foundation. I have benefited from discussions with William Brainard, Harry Johnson, Alvin Marty, and E. S. Phelps. Roger Grawe helped with the diagrams.

terms is increased, its acquisition may absorb private saving that would go into investment in productive capital. Thus the degree to which saving is absorbed in government debt helps to determine the equilibrium capital-labor ratio and the net marginal productivity of capital. The welfare question is whether this diversion of saving steers the economy toward or away from the optimal capital-labor ratio.

Properties of Neoclassical Growth Equilibrium

Let me review the well-known essential properties of a model capable of balanced growth paths of "moving stationary states" (for a good expository review, see Johnson, 1967, chap. iv). Output depends on two inputs, capital and labor, with constant returns to scale and diminishing positive marginal productivity of each factor; the production function remains constant over time except for technical progress that augments the effective labor input represented by a natural unit of labor; the "natural" rate of growth of the supply of effective labor is an exogenous constant, determined by population growth and technical progress. Under these conditions, there is a family of paths along which output, capital, and effective labor all grow steadily at the natural rate. Each path is characterized by its constant capital-output ratio and, related to that, its constant ratio of capital to effective labor. With a high capital-output path is associated a high efficiency wage and a low marginal productivity of capital. A path with low capital-labor and capital-output ratios will have a low wage and a high marginal productivity of capital.

Let g be the natural rate of growth, μ the capital-output ratio, and s the share of output that is invested in new capital. Along an equilibrium path the rate of growth of the capital stock, s/μ, must be equal to the natural rate, g. Therefore, a balanced growth path will be an equilibrium path if and only if it induces savers to hold capital in the technologically required ratio to output, μ, and accordingly to provide continuously the required addition to capital, namely a constant fraction of output equal to μg.

Figure 1 relates the capital-output ratio to the net marginal productivity of capital, R_K, for alternative balanced growth paths. Curve T represents the technological relationship between these variables implicit in the economy's production function. Paths with higher capital-output ratios will have lower marginal productivities of capital. A curve like S represents the amount of wealth savers desire, relative to national income, in a situation of balanced growth. As curve S illustrates, they may desire a higher wealth-income ratio along a path with a high real rate of interest than along one where the reward for saving is low. In an economy where the only store of value is capital, the wealth-income ratio must be the same as the capital-output ratio, and the rate of interest on saving is the net marginal productivity of capital. The equilibrium capital-output ratio is

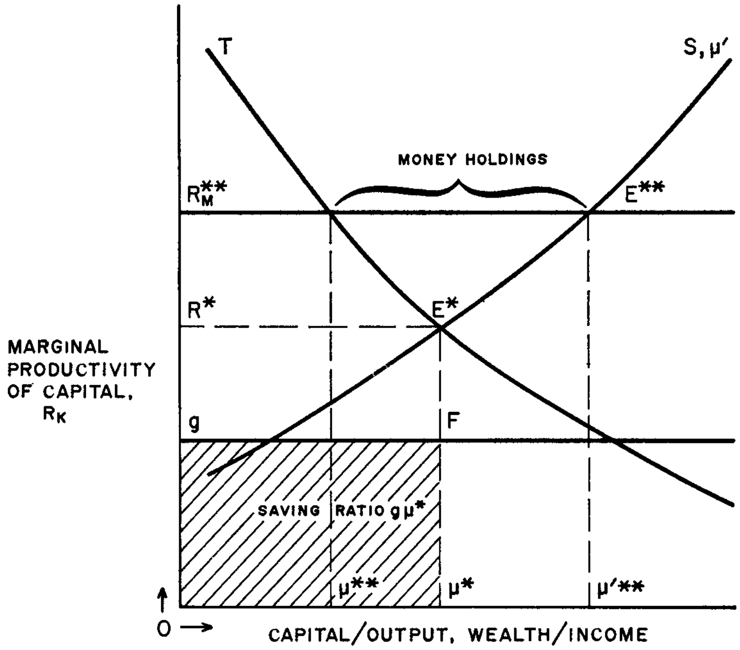

FIG. 1

therefore μ^*; the corresponding amount of saving or investment relative to income is $g\mu^*$, graphically illustrated by the area of the rectangle $O\mu^*Fg$. A curve like S, and therefore an equilibrium like E^*, will exist if the properties of a balanced growth path suffice to determine a constant desired wealth-income ratio. This means that desired wealth and saving, relative to income, must not depend on the absolute level of population or of per capita income.

A special assumption about saving behavior, which has received the most attention in the literature, is that the saving ratio is a constant $\bar{s}$. This means, of course, that the wealth-income ratio desired along a balanced growth path is the same for every path, namely $\bar{\mu} = \bar{s}/g$. In other words, the S curve is vertical. Of the technologically possible paths, the one that has a capital-output ratio $\bar{\mu}$ is the equilibrium. This is illustrated in Figure 2.

Another conceivable special case is that consumer savers have a fixed marginal rate of substitution of present for future consumption, will accumulate wealth indefinitely at a rate of interest greater than R^* or equal to that rate, and dissave indefinitely at any lower rate. If so, the S curve is horizontal, as in Figure 3. The equilibrium capital-output ratio is the one that provides a marginal productivity of capital equal to this rate.

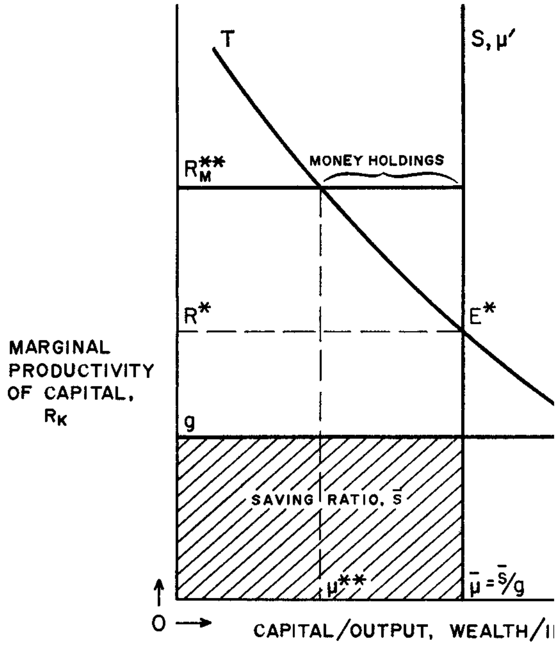
T
S, μ′
MONEY HOLDINGS
R_M^{**}
R*
E*
MARGINAL
PRODUCTIVITY
OF CAPITAL,
R_K
g
SAVING RATIO, $\bar{s}$
μ**
$\bar{\mu} = \bar{s}/g$
0
CAPITAL/OUTPUT, WEALTH/I

FIG. 2

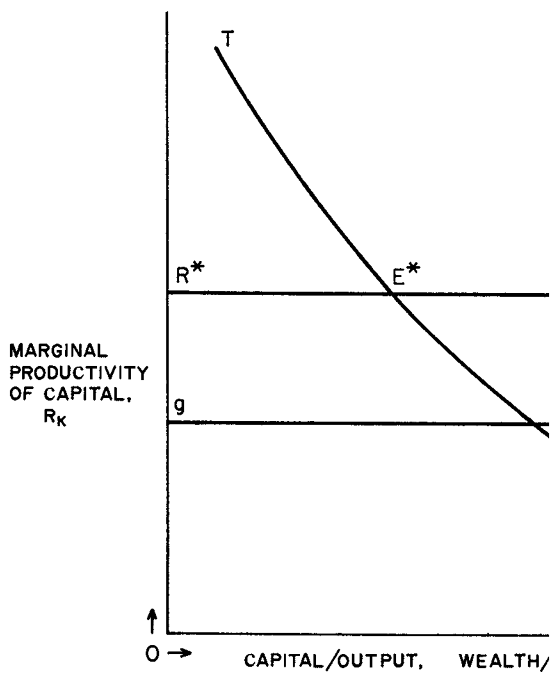
T
R*
E*
MARGINAL
PRODUCTIVITY
OF CAPITAL,
R_K
g
0
CAPITAL/OUTPUT, WEALTH/

FIG. 3

Growth with a Monetary Asset Perfectly Substitutable for Capital[1]

What happens when a second asset, a competing store of value, is introduced? Suppose that money—still in the sense of government debt—has a value $1/p$ in terms of goods; that the price level p is perfectly flexible; and that the government pays an own rate of interest, or nominal rate, of r on money.

One thing we know right away, of course, is that in a growth equilibrium the *real* quantity of government debt, like every other real magnitude, must be expanding at the natural rate of growth of the economy. That is, if M is the nominal quantity,

$$\frac{\dot{M}}{M} - \frac{\dot{p}}{p} = g. \tag{1}$$

We can also calculate the real rate of interest on money as

$$R_M = r - \frac{\dot{p}}{p} = r - \frac{\dot{M}}{M} + g. \tag{2}$$

Thus the question of the optimal value of $\dot{M}/M$ can be translated into the question of the optimal value of R_M, the real rate of return on money.

This question is easiest to answer in the most uninteresting case, that is, where it is assumed that money and capital are perfect substitutes in the portfolios of savers. This means that owners of wealth are indifferent about the proportions in which they hold the two assets so long as their real yields are equal (or differ by an exogenously determined constant) and otherwise will hold none of the lower-yielding asset. In this case, clearly, coexistence of the two assets requires:

$$R_K = R_M = r - \frac{\dot{M}}{M} + g. \tag{3}$$

Evidently the government can, by determining r and $\dot{M}/M$, determine R_K and, therefore, determine the equilibrium capital-labor ratio. In Figure 1, for example, by selecting R_M^{**}, the government steers the economy to a capital-output ratio μ^{**}. (An exception to this rule arises in the case exemplified by Figure 3. Here R_K is determined by the perfectly elastic supply of saving at the rate R^*. The government has no choice but to set R_M at the same level; otherwise there will be either no demand for capital or no demand for money.)

However, the capital-output ratio is no longer the same as the wealth-income ratio. Wealth is now $K + (M/p)$. And "disposable" income exceeds output by the growth of the real stock of money, $g(M/p)$. Let

[1] This and the next section draw on and elaborate my article, "Money and Economic Growth" (1965).

$m = M/pK$, the ratio of money to capital holdings; this will be constant in equilibrium. Let μ' be the ratio of wealth to disposable income. Then

$$\mu' = \frac{K(1 + m)}{Y + gmK} = \frac{\mu(1 + m)}{1 + gm\mu}, \tag{4}$$

$$\mu = \frac{\mu'}{1 + m(1 - g\mu')}. \tag{5}$$

Here $g\mu'$ will be recognized as the ratio of total saving, including accumulation of money in real terms as well as of capital, to disposable income. Since this is smaller than one, the capital-output ratio μ is smaller than, equal to, or larger than μ' according as m is positive, zero, or negative. Now to interpret Figure 1 remember that in the case under consideration money and capital are perfect substitutes. Therefore, the curve S can still represent the desired wealth-income ratio μ'. When, for example, the real rate of interest, whether R_K or R_M, is set at R_M^{**} in Figure 1, the desired wealth-income ratio is given by curve S as μ'^{**}. This exceeds the capital-output ratio μ^{**}, and the two must be reconciled by a positive value of m necessary to satisfy equation (5). Since any value of m is acceptable to wealth owners, this situation is an equlibrium.

In Figure 1, R^* is the lowest real rate of interest compatible with positive quantities of money. Should the authorities establish a real rate lower than R^*, μ' would be less than μ and m would have to be negative. That is, the government would have to be a net creditor of the private economy, its net credit position rising at the natural rate of growth.

In the special case of Figure 2, the saving ratio is fixed at $\bar{s}$. With the introduction of money, this ratio should now be applied to disposable income, so that the vertical S curve applies to $\mu' = \bar{s}/g$. From equation (5) we have:

$$\mu = (\bar{s}/g)[1/1 + m(1 - \bar{s})]. \tag{6}$$

It is still true that, as in the case of Figure 1, m can adjust to any horizontal difference between the curves T and S that exists at the established real interest rate.

If the authorities can in effect set any real rate of interest they want, which should they set? First of all, they should not set one below g. A balanced growth path with a marginal productivity of capital below the growth rate is inefficient: All generations could have higher consumption by saving less and having a lower capital-output ratio. Conceivably the economy could, in the absence of money, be stuck on an inefficient equilibrium path of this kind. If so, the government could improve matters by issuing money with a real rate of interest equal to g, absorbing some of the excessive thrift of the population in this paper form and raising the marginal productivity of capital to g.

Note that a rate of g means that $\dot{M}/M = r$ and that $\dot{p}/p = r - g$. The debt increases solely by the government's incurring new debt to pay the interest. The considerations so far developed give us no criterion for choosing further a common value of $\dot{M}/M$ and r, for example, choosing between (*a*) $\dot{M}/M = r = g$ and price stability, and (*b*) $\dot{M}/M = r = 0$ and price deflation at the natural rate g. In equilibrium the steady rate of price change is wholly anticipated, and one rate is as good as another.

If, in the absence of government debt, the equilibrium marginal productivity of capital would exceed g, it is not optimal to absorb any saving in government debt. In other words, if outside money is not competitive at a real rate g, the government should not try to make it competitive by offering a higher rate. The reasoning is as follows: a steadily growing economy with an indefinite life can always use claims against the government to trade present consumption for future consumption at the rate of interest g. In a moving stationary state, savers acquiring such claims can be assured of this return simply because the savers of the future will be more numerous and/or richer. By the time savers wish to cash in their paper assets, the market for them will have grown at the natural rate of increase g. But this is the maximum sustainable rate at which such trades can be made via government paper. If accumulation of physical capital offers a higher payoff in future consumption for current saving, then it dominates government paper as a vehicle for making such trades.

The question then arises whether the government should push the economy to the "golden rule" degree of capital intensity by augmenting the private saving of the economy with public saving from tax revenues. In equilibrium the government would lend at the interest rate g, which would also be the rate of increase of the private economy's debt to the government. There is a considerable literature discussing the sense, if any, in which the golden rule solution is optimal, and I will not review it here. The important conclusion of this discussion is that the situation is not symmetrical with the case in which government debt could rescue the economy from overcapitalization. If the society is originally endowed with a less-than-golden-rule capital stock, then the additional capital cannot be acquired without reducing the consumption and utility of consumers during the transition. These will be taxpayers who are, in effect, forced to save involuntarily in order to build up the government's stock of claims against the private economy.

Growth with a Monetary Asset Imperfectly Substitutable for Capital

The more interesting case is that in which capital and money are not perfect substitutes but can coexist in wealth-owners' portfolios even when they do not bear the same real rate of interest. The money-capital ratio is, by the same token, not a matter of indifference; it depends on the two rates R_K and R_M.

Let us provisionally suppose, although this assumption now has less logic, that the curves S still describe the relationship of the desired wealth-income ratio μ' to R_K and, what is more, do so independently of R_M. Given R_K and R_M, a portfolio-balancing value of m is determined. Equation (5) or (6) then shows how μ' and m together determine the value of μ, which satisfies both the saving and the asset preferences of the public. For positive m, μ will be smaller than μ'; for negative m, larger. Given R_M, increasing R_K will lower m and raise μ relative to μ'.

This is illustrated in Figures 4 and 5, where it is seen that the introduction of money, with a real rate that induces wealthowners to substitute it for capital, lowers the equilibrium capital intensity and raises the equilibrium marginal productivity of capital. The curve $\mu(R_M)$ will be shifted upward by a rise in R_M. Thus the diversion of saving will be greater with a high R_M than with a low R_M. Once again, any substitution of money for capital—any positive m—will move the capital-output ratio in the wrong direction unless the marginal productivity would otherwise be below g.

If there were danger of an inefficient, "over-capitalized" equilibrium, then there is an optimal combination of r and $\dot{M}/M$ that will raise the

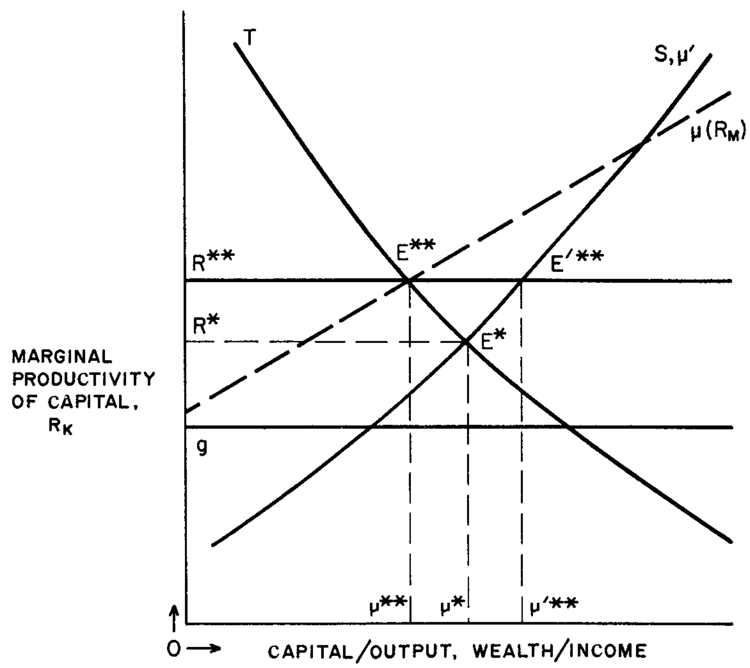

FIG. 4

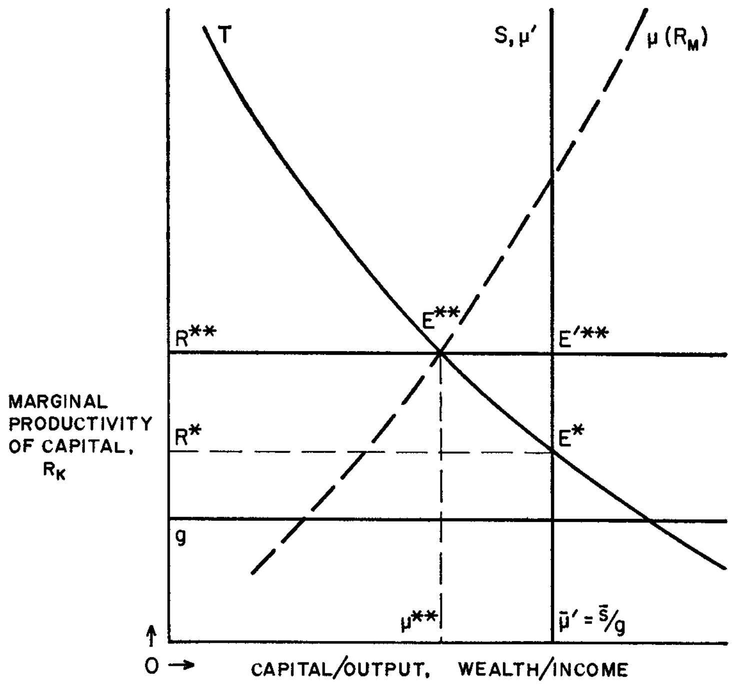

FIG. 5

marginal productivity of capital to the golden rule rate g. This is illustrated in Figure 6. In this equilibrium the real rate on money will not necessarily be equal to $R_K = g$, but probably lower by an amount that reflects those imputed advantages of money that lead people to hold it even when its explicit rate is not competitive with the return on capital. Therefore, the optimal $\dot{M}/M$ is larger than the nominal interest rate r.

The general conclusion is that there is an optimal rate of growth of the supply of outside money, equal to or smaller than the nominal interest rate r, only to the extent that diversion of saving into this vehicle is necessary to keep the marginal productivity of capital from falling below g. In the absence of such a tendency for oversaving, it is not optimal to absorb any saving in outside money or deadweight debt.

This does not mean that the government should not issue any liabilities for private savers to acquire. There may be good reason to do so, evidenced by the willingness of the public to hold such liabilities at a sacrifice of return. It does mean that the government should in turn invest the savings entrusted to it, either directly on its own account or indirectly by re-lending them to private investors, in capital bearing the prevailing real rate of return. The question which then arises, that is, how far such financial intermediation should be carried, will be discussed, though not answered, below.

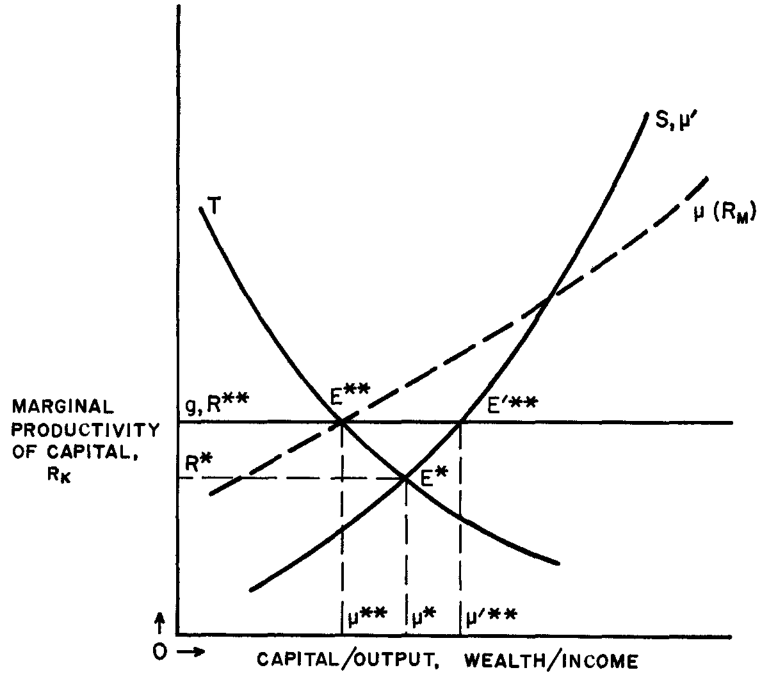

FIG. 6

Inflation and Unemployment

The preceding discussion has assumed that prices, and correspondingly money wages, are completely flexible, so that any rate of inflation or deflation can occur consistently with full employment of the exogenously given labor force. One implication is that the government cannot directly control the stock of outside money in real terms; whatever nominal stock is supplied will be instantaneously adjusted to the public's demand by movement of the price level. Thus the government controls the real supply, in the models so far discussed, only indirectly, by influencing the demand. As explained, the rate of change of the nominal supply helps to determine the size of the stock demanded.

A Phillips-curve model of wage and price determination has somewhat different implications. The rate of price inflation is an increasing function of the degree of utilization of the labor force. Steady growth at the natural rate can occur at any fixed degree of utilization; once this is chosen, a particular steady rate of inflation has also been chosen. Given the nominal rate r, the government's choice of $\dot{M}/M$ determines in equilibrium not only the rate of inflation $\dot{p}/p$ [equal to $(\dot{M}/M) - g$] but the degree of utilization associated with it. It also determines R_M [equal to $r - (\dot{p}/p)$] and, therefore, in the manner explained above, the equilibrium capital-output ratio.

Thus a deflationary policy—low $\dot{M}/M$—not only diminishes the amount of capital per employed worker but also diminishes the number of workers employed. Even the use of a relatively deflationary policy to discourage overinvestment in physical capital has an offsetting disadvantage in reducing utilization. But the possibility of manipulating r is another degree of freedom. An obvious way to have a rate of inflation compatible with high employment without having an excessively low real rate of return on money is to compensate for the inflation by a higher nominal interest rate.

The application of the Phillips-curve model to long-run growth equilibrium is questionable, but I will not enter that debate here. It is also questionable that the government's choice of a $\dot{M}/M$ policy can proceed as if all values of $\dot{p}/p$, negative and positive, were equally feasible.

II. "Money" as Means of Payment

The previous discussion has been concerned with the effects of the availability of monetary and financial assets on capital formation. In this context the important property of these assets is that they are stores of value with smaller yield uncertainties than physical capital. The optimal size and rate of growth of the supply of means of payment is quite another question. Means of payment can be supplied either as outside money or as inside money, without affecting in one way or another the optimal supply of saving for capital formation. What part of public or private debt should be monetized to provide a circulating medium is a narrower question.

Growth and the Demand for Means of Payment

In a balanced growth equilibrium or "golden age" of the type discussed above, our natural first presumption would be that the stock of means of payment, in real terms, must grow with every other real aggregate, at the pervasive growth rate of the economy. The nominal stock will then grow at this rate plus the rate of price inflation.

However, this initial presumption deserves closer examination. I shall consider it from the viewpoint of the inventory theory of the demand for cash, which focuses attention on the management of the temporary and fluctuating balances that people hold to bridge gaps between their receipts and their outlays (Tobin, 1956).

According to this theory, there are economies of scale in the management of these balances. These economies arise from the fact that at least a portion of the costs of transactions between cash and higher-yielding assets is independent of the size of the transactions and depends only on the number of transactions. Some individuals have temporary balances too small and short-lived to justify the costs of investing them; they simply hold the balances in cash. For others, the average amount of cash held depends inversely on the interest premium available and directly on the volume of receipts or outlays relative to the costs of making financial transactions. When the dollar volume of an individual's receipts and outlays increases, while his transaction costs remain unchanged, his average cash holdings rise less than in proportion.

It is an error of composition, however, to attribute to the theory the prediction that in a growing economy the demand for means of payment expands at a rate slower than total income and wealth. In the absence of technical progress, growth in neoclassical theory is simply an increase in the population of individuals, households, firms. But none of these units increases in average size—income, wealth, volume of receipts, and outlays. Nothing happens to transaction costs either; the opportunity costs of making transactions, measured either in human labor or in consumer goods, remain the same. Interest rates are constant. Consequently, the theoretical prediction is that the demand for means of payment increases like everything else at the natural rate of growth of the population and economy.

If there is labor-augmenting (Harrod-neutral) technical progress, the situation is more complicated. Now the scale of an economic unit—its income, wealth, and volume of transactions—increases at the rate of technical progress. What happens to transaction costs? If the making of transactions is purely a labor-using activity—either the labor of the transactor or of his agents—which does not benefit from labor-augmenting technical progress, then the costs are essentially wages. Wages rise at the rate of technical progress the same as the volume of transactions per economic unit. Once again, then, the conclusion is that the demand for means of payment rises at the natural growth rate—that is, the sum of the rates of population increase and technical progress. The non-linearity of the inventory-theory approach does not carry over to the economy as a whole.

However, it might be more natural and more realistic to assume that the making of transactions benefits from labor-augmenting technical progress at the same rate as productive activity in the economy. In this event the scale of a typical transactor increases at the rate of technical progress, while transactions costs do not rise; the average cash holding of a transactor rises less than in proportion to the increase in the volume of his transactions. Hence the aggregate demand for means of payment increases at a rate larger than n, the rate of population increase, but less than $n + \gamma$, the rate of growth of income and wealth—indeed, approximately at the rate $n + (\sqrt{1 + \gamma} - 1)$.

The conclusion that the velocity of means of payment should rise secularly in a growing economy would be even stronger if it were thought that technical progress in the making of transactions were faster than elsewhere in the economy. This is not farfetched. The innovations in business machines—calculating, automatic data processing, copying—are perhaps the most dramatic example of labor-augmenting technological progress. Among other things, these innovations have surely cut the costs of economizing cash balances. They have been accompanied by other cash-economizing innovations. Credit cards are a device by which small transactors are pooled into large units than can exploit the economies of scale in cash management.

It is possible that the growth of per capita income due to technical progress leads to a shift from work to leisure. If so, the supply of labor does not grow as fast as population. The natural rate of growth of the economy, both income and capital, is less than $n + \gamma$. Income and transactions volumes per unit grow less rapidly than the rate of technical progress, but the wage rate grows at the rate of technical progress. In this case it is possible that average cash holding requirements rise—provided little or no progress occurs in financial technology.[2]

On balance the inventory-theory model gives little support to the idea that the demand for means of payment should, at constant interest rates, rise secularly relative to national income. The model does not suggest that money behaves like a luxury durable consumer's good, generating services for which the income elasticity of demand exceeds one. This would work against the model if the empirical evidence advanced in support of the assertion that the long-run trend in velocity is downward were more convincing. But for the United States the "money" whose holdings have grown relative to income includes more than means of payment, specifically commercial bank time deposits. Before the first world war, these were the major monetary store of value available to savers. Since the second world war, the demand for time deposits has been increased by the dramatic rise in their yields relative to other vehicles of saving. There is no evidence that the services of means of payment per se are a luxury good.

The Optimal Supply of Means of Payment

The preceding discussion concerns the rate of growth in means of payment, in real terms, needed in a steadily growing economy. What is the optimal size of the growing stock of means of payment?

[2] If growth is not balanced but is accompanied by a general increase in capital per unit of output and per augmented unit of labor, this too leads to an increase both in income per capita and in the wage rate. But even so there is no need for an increase in cash holdings per capita unless the wage rate increases faster than per capita income. Whether it does or not depends on technology in ways that can be somewhat tediously described by the value of the elasticity of substitution. A higher wealth-income ratio may increase the demand for cash for transactions in capital account—a requirement somewhat similar to Keynes' finance motive. And the decline in interest

Scarcity of means of payment forces individuals, firms, and other economic units to economize their cash holdings. In order to gain the earnings possible from keeping their working balances heavily invested in assets, which are not means of payment but yield higher real returns, they must make frequent transactions in and out of cash. These transactions have real costs, for example, the labor of the transactors themselves or their agents. Diversion of productive resources into the handling of in-and-out transactions is socially wasteful, because there is no cost to society in creating means of payment.

The way to avoid this waste is to supply a large enough stock of means of payment to absorb all working balances. This requires that means of payment bear a high enough real rate of return to remove the incentive to economize them. No resources should be devoted to the making of transactions to invest working balances temporarily in higher-yielding assets, either physical capital or public and private debt instruments. Since optimality will generally require positive real rates of interest on these alternative assets, means of payment will also have to bear a positive real rate.

If the nominal rate on means of payment is stuck at zero by law or by convention, the incentive to economize means of payment cannot be removed without persistent and anticipated deflation. When there are rigidities of money wages and prices that prevent such deflation and convert deflationary impulses into unemployment of labor and other resources, the social waste of economizing means of payment becomes one of the costs of avoiding the larger social wastes of underutilization. Deflation, of course, will also contribute to the real returns on other assets denominated in the monetary unit of account. Their nominal yields will have to be close enough to zero to prevent them from, on the one hand, diverting saving from capital formation and, on the other hand, diverting working balances away from means of payment.

A better way would be to allow means of payment to bear a nominal interest rate. Or, to put the same thing another way, interest-bearing assets defined in the monetary unit of account could be allowed to serve as means of payment. There seems no reason, for example, why checking should not be permitted against savings accounts in commercial banks and thrift institutions, transforming these assets into interest-bearing means of payment. Freeing means of payment from the legal limitation of zero interest would make it theoretically possible to have an efficient growth equilibrium without deflation—efficient both in the sense that the real rate of interest is high enough to avoid overcapitalization and in the sense that real resources are not diverted into economizing means of payment.

---

rates due to higher capital intensity may also lower velocity. The price of transaction-making services will not change if labor productivity gains as much from higher capital intensity there as elsewhere. But it may rise if financial services are more labor intensive than productive activities.

III. Uncertainty, Saving, and Liquidity

I return now to the question dodged earlier by the assumption that the desired wealth-income ratio is independent of the menu of assets available to savers and the structure of their rates of return. The comparisons that this assumption tempts one to make between economies with and without outside money, illustrated in Figures 1 and 2, are likely to be misleading.[3] When money or other financial assets and real capital bear different real rates of return, clearly they are not perfect substitutes. If financial assets are not perfect substitutes for real capital in the portfolios of wealth owners, it is unlikely that they replace them dollar for dollar in saving decisions. The old Keynesian dichotomy—analyze separately decisions about total wealth and decisions about its composition—is useful for many purposes. I have generally found it convenient myself. But for our present problem it is not really appropriate, because the central question is precisely the bearing of alternative financial policies and institutional arrangements on the supply of saving available for capital formation.

Unfortunately, I have not found this to be a simple problem, and at the moment I cannot do more than indicate some directions for future analysis.

One of the principal reasons that savers hold financial assets with expected yields smaller than those on real capital is to diminish uncertainty about the amounts of their future consumption. We must begin, therefore, by considering the bearing of uncertainty on saving decisions—first without and then with financial assets available in addition to real assets.

General Remarks on Uncertainties and Saving Decisions

There are two ways in which uncertainty and risk aversion affect the quantity of saving for the future. One kind of uncertainty, *yield uncertainty*, relates to the return on saving, positive or negative—the individual does not know how much future consumption a dollar saved today will actually provide. The other kind of uncertainty, *need uncertainty*, relates to the size of the consumer's future needs and the degree to which resources other than saving will be available for meeting them—he does not know what his wage income will be or whether he will confront extraordinary consumption needs, for example, for medical reasons.

So far as yield uncertainty is concerned, aversion to risk may either deter saving or increase it. It deters it by diminishing the attractiveness of the reward for saving— a kind of substitution effect. Some risk-averse

[3] An economic historian would be puzzled by the implication of Part I that the development of monetary and financial institutions is in some sense bad for real investment. Without the safe assets made available by these institutions, how would the thrift of the cautious saver have been mobilized? The conflict is largely superficial. Financing of capital accumulation is the story of inside money, not of outside money.

consumers may save nothing for the future, even though they would save if the expected return on saving were sure. Uncertain of the return on saving, they prefer the certain utility of present consumption. On the other hand, yield uncertainty combined with risk aversion may increase saving via a calculation that, since the payoff may not be large, it is best to save enough to make sure of adequate future consumption, a kind of income effect.

Given this familiar ambiguity, it is not surprising that availability of less risky lower-yielding assets may work in either direction. Some savers, in particular those who previously saved nothing, will respond to the possibility of accumulating less risky portfolios by saving more. Conceivably they will even acquire more risky assets than when the opportunity to mix them with safer assets was not available. On the other hand, the possibility of saving in forms that give greater assurance of future return may diminish the need for saving felt by consumers dominated by the income effect.

Need uncertainty is the source of what might be termed precautionary saving—saving more than is actually required to meet future contingencies. Because of declining marginal utility of consumption, need uncertainty leads risk-averse savers to impute a higher value to a dollar provision for the future, relative to a dollar of current consumption, than they would if they anticipated with certainty the same average level of future consumption. Nevertheless, the uncertainties surrounding yield on risky assets may be so great that saving in this form will not improve the individual's position. The availability of safe assets will then increase precautionary saving. However, from an economy-wide standpoint, such saving may be excessive, providing ex post too generously for future consumption at the expense of current consumption. Insurance against unemployment, accident, illness, and longevity may be better social devices for accommodating individuals' reluctance to carry these risks.

Yield Uncertainty and Saving: A Two-Period Example

To reduce the problem to its most elementary terms, let us consider that old friend of the classroom and textbook, the consumer with a two-period life and horizon. Let his lifetime utility be the sum of two utilities, one depending on real consumption in period one, the other on real consumption in period two. That is, $U = u(c_1) + v(c_2)$. As is well known, this assumption of independence already gives us cardinal utility. When deciding how much to save in period one for use in period two, the individual knows for sure the consequences of his decision for c_1, but he does not know for sure what c_2 will result in. He is assumed to maximize expected utility: $E(U) = u(c_1) + E[v(c_2)]$. Further simplifying the problem, I assume that the expected value of second-period utility is a function of two parameters of the probability distribution of second-period consumption, the mean $\bar{c}_2$ and the standard deviation σ_{c_2}. Thus

$$E(U) = u(c_1) + \varphi(\bar{c}_2, \sigma_{c_2}).$$

A simple and pure case of yield uncertainty is constructed by endowing the consumer with known incomes y_1 and y_2 in the two periods, with the possibility of saving in a single asset with uncertain prospects. Let x be the value in consumption of period two of a unit saved from consumption of period one, $\bar{x}$ its expected value, σ_x its standard deviation. If $y_1 - c_1$ is saved, then $\bar{c}_2 = y_2 + (y_1 - c_1)\bar{x}$ and $\sigma_{c_2} = (y_1 - c_1)\sigma_x$. The expected value of period-two utility is then $\varphi(\bar{c}_2, \sigma_{c_2})$, and we define a certainty-equivalent level of second-period consumption as c_2^* such that $\varphi(c_2^*, 0)$—the same as $v(c_2^*)$—is equal to $\varphi(\bar{c}_2, \sigma_{c_2})$.

In Figure 7 the effect of yield uncertainty on the saving decision is illustrated. In the left panel, c_1 is measured right to left on the horizontal axis, and $\bar{c}_2$ and c_2^* are measured vertically. Indifference curves in the left quadrant, like U_A and U_B, represent constant levels of utility for consumption combinations actually achieved: $U = u(c_1) + v(c_2^*)$ For convenience y_2 is taken to be zero. The frontier of possible combinations of c_1 and $\bar{c}_2$ is the line from y_1 to $y_1\bar{x}$. If x were known in advance with certainty, or if the individual were risk neutral, he would equate $\bar{c}_2$ and c_2^* and divide his income y_1 between current consumption and saving in the manner indicated by point A on the frontier.

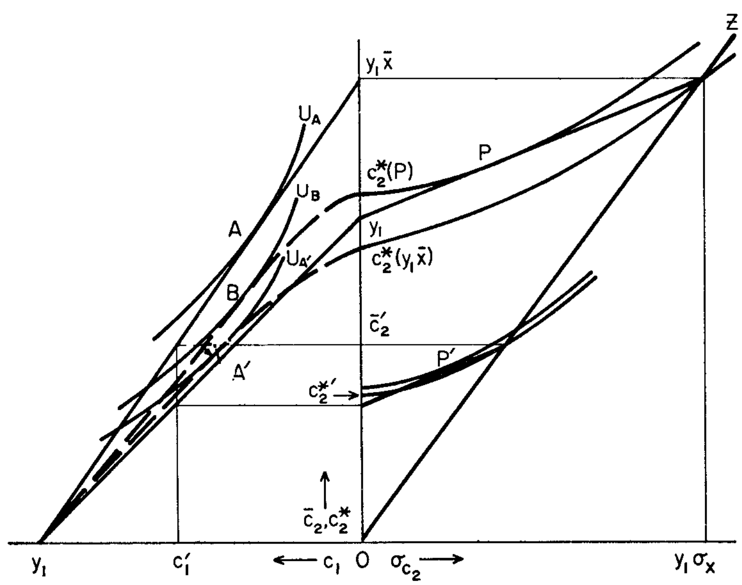

FIG. 7

Suppose, however, that x is not known with certainty and that our consumer is risk averse. On the right hand panel σ_{c_2} is measured horizontally. The line from the origin OZ shows for each value of expected consumption $\bar{c}_2$ on the vertical axis the corresponding dispersion σ_{c_2} of future consumption. The indifference curves in the right quadrant represent constant values of $\varphi(\bar{c}_2, \sigma_{c_2})$. Their upward slope indicates risk aversion. Their intercepts with the vertical axis, where $\sigma_{c_2} = 0$, are their certainty equivalents c_2^*. Thus it is possible to convert any expected consumption $\bar{c}_2$ into its certainty equivalent c_2^* by following back to the axis the indifference curve that cuts OZ at the level $\bar{c}_2$. Consider, for example, the future consumption made possible by saving y_1 in full. Expected future consumption is $y_1\bar{x}$; risk is $y_1\sigma_x$; the indifference curve through this point hits the axis at $c_2^*(y_1\bar{x})$. Similarly, an amount of current consumption c_1' leads to expected future consumption $\bar{c}_2'$ with a certainty equivalent $c_2^{*\prime}$. Joining all the points like $(c_1', c_2^{*\prime})$ results in a pseudo-opportunity locus. The individual determines his consumption and saving so as to arrive at A' on this locus rather than at A. In the diagram this is shown to involve less saving, although this result is not inevitable.[4]

It is conceivable that indifference curves in the right panel cut the line OZ from below, not from above. In this case, c_2^* diminishes as additional saving increases $\bar{c}_2$, and the effective pseudo-opportunity locus vanishes almost to the point y. This situation is illustrated in Figure 8. With such pronounced risk aversion, yield uncertainty is a great obstacle to saving and capital formation.

Now suppose that a riskless asset is provided by which a dollar of current consumption c_1 can be turned into a dollar of future consumption with perfect safety. This possibility is depicted in Figures 7 and 8 by the

[4] The optimal choice of c_1 is found by setting equal to zero the derivative of

$$\psi(c_1) = u(c_1) + \varphi[(y_1 - c_1)\bar{x}, (y_1 - c_1)\sigma_x] : \psi'(c_1^*) = u'(c_1^*) - \varphi_1\bar{x} - \varphi_2\sigma_x = 0,$$

provided this has a solution between 0 and y_1. The second derivative must be negative:

$$\psi''(c_1^*) = u''(c_1^*) + (\varphi_{11}\bar{x}^2 + 2\varphi_{12}\bar{x}\sigma_x + \varphi_{22}\sigma_x^2) < 0.$$

Consider the relationship of

$$c_1^* \text{ to } \sigma_x : (i)\ \frac{\partial c_1^*}{\partial \sigma_x} \cdot \psi''(c_1^*) = \varphi_{12}\bar{x}(y_1 - c_1^*) + \varphi_{22}\sigma_x(y_1 - c_1^*) + \varphi_2.$$

Given that ψ'', φ_{22}, and φ_2 are negative, $\partial c_1^*/\partial \sigma_x$ can be negative only if φ_{12} is positive. We know that if φ is the expected value of a quadratic utility function of c_2 then $\varphi = \bar{c}_2^2 - b(\bar{c}_2^2 + \sigma_{c_2}^2)$, (where $b > 0$) and $\varphi_{12} = 0$. Alternatively for a more general $v(c_2)$

$$\varphi = \int_{-\infty}^{\infty} v(\bar{c}_2 + \sigma_{c_2}z)N(z)dz,$$

where $N(z)$ is the normal probability density function with zero mean and unit standard deviation. In this case, $\varphi_{12} = \int Zv''(\bar{c}_2 + \sigma_{c_2}Z)N(z)dz$ and may have either sign.

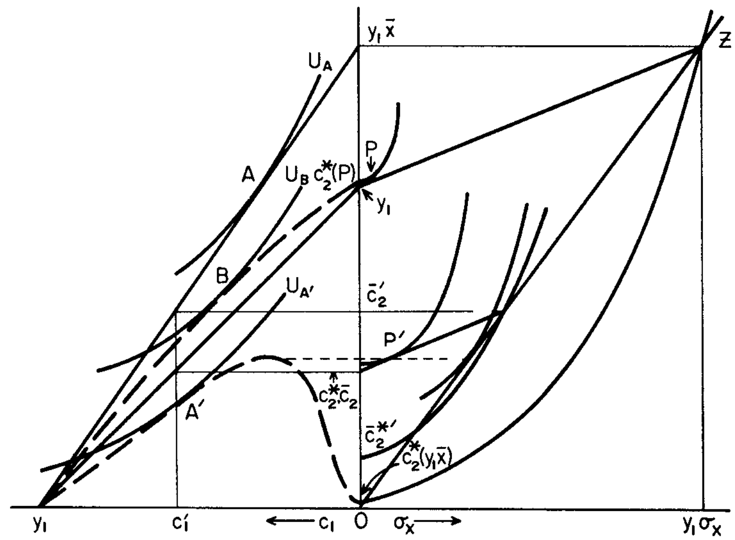

FIG. 8

45-degree line from y_1 on the c_1 axis to y_1 on the c_2 axis. Since $\bar{x}$ is assumed greater than 1, the 45-degree line lies inside the original frontier. Now, a given amount of saving can result in a variety of combinations $(\bar{c}_2, \sigma_{c_2})$ depending on how the saving is split between the two assets. For example, saving all of y_1 can provide expected consumption c_2 of y_1 with no risk, or as before $\bar{c}_2 = y_1\bar{x}$ with risk of $y_1\sigma_x$, or any linear combination of these as indicated by the line connecting the two points. The actual combination that would be chosen for highest expected utility is indicated by point P, and this has a certainty equivalent $c_2^*(P)$ Note that this certainty equivalent is higher than the equivalent of the same amount of saving in the absence of the safe asset. By repeating this procedure for other levels of saving, a new pseudo-opportunity locus can be constructed and a new equilibrium consumption-saving decision, point B, determined. The diagram shows more saving at B than at A', but less than at A. Figure 8 indicates a dramatic increase in saving, because the pseudo-opportunity locus scarcely existed in the absence of the safe asset.[5]

[5] The optimal choice of c_1 is found by setting equal to zero the derivatives with respect to c_1 and m of

$$u(c_1) + \varphi\{(y_1 - c_1)[m + (1 - m)\bar{x}], (y_1 - c_1)m\sigma_x\},$$

where m is the proportion of saving placed in the safe asset:

$$u'(c_1^*) - \varphi_1[\bar{x} + m(1 - \bar{x})] - \varphi_2 m\sigma_x = 0;$$

$$\varphi_1(y_1 - c_1^*)(1 - \bar{x}) + \varphi_2(y_1 - c_1^*)\sigma_x = 0.$$

(continued)

Thus the availability of a safe asset may very well increase saving and make the "as if" terms on which saving decisions are reached accord more closely with the opportunities actuarially available for transforming present consumption into future. But the safe asset itself absorbs saving, and it is unlikely, except in the case illustrated in Figure 8, that its availability actually increases the direct flow of risk-averse saving into capital formation. Increasing the total amount of capital formation will generally require that the savings placed in the safe asset flow indirectly, via the government or private financial intermediaries, into real investment. In the extreme it is possible to lead risk-averse investors to point A by offering them a safe asset with a guaranteed rate of return equal to the expected marginal efficiency of capital.

The same apparatus may be used to analyze need uncertainty and precautionary saving. Figure 9 has the same general format as the previous diagrams. But now the zero-saving position is indicated by points A in both diagrams. In period two expected income $\bar{y}_2$ is at the vertical level of these points; but the dispersion of this prospect is σ_{y_2}. The corresponding $c_2^*(A)$ is shown on the vertical axis. If saving can be done by acquiring an asset which will realize on average $\bar{x}$ per dollar at a risk of σ_x, further

---

These may be reduced to the following pair of equations:

$$u'(c_1^*) - \varphi_1\bar{x} = 0;$$

$$\varphi_1(1 - \bar{x}) + \varphi_2\sigma_x = 0.$$

Clearly c_1^* will be smaller, saving larger, than if m is constrained to be zero. Suppose that $\tilde{c}_1$ satisfies the equation of the previous footnote, so that

$$u'(\tilde{c}_1) = [\bar{x}\varphi_1(y_1 - \tilde{c}_1)\bar{x}, (y_1 - \tilde{c}_1)\sigma_x] + \sigma_x\varphi_2[(y_1 - \tilde{c}_1)\bar{x}, (y_1 - \tilde{c}_1)\sigma_x].$$

Now consider the marginal utility of saving $(y_1 - \tilde{c}_1)$ in the second regime:

$$\{\bar{x}\varphi_1(y_1 - \tilde{c}_1)[\bar{x} + m(1 - \bar{x})], (y_1 - \tilde{c}_1)m\sigma_x\}.$$

The second φ_1 has smaller arguments than the first and—provided again that φ_{12} is non-positive—is larger. We know also that φ_2 is negative. Therefore, the marginal utility of saving $y_1 - \tilde{c}_1$ in the second regime exceeds the marginal utility of $\tilde{c}_1$. Thus, c_1 will have to be lower than $\tilde{c}_1$ to bring the marginal utilities of consumption and saving back into equality.

Similarly, consider $\hat{c}_1$ such that $u'(\hat{c}_1) = \bar{x}\varphi_1[(y_1 - \hat{c}_1)\bar{x}, 0]$, that is, the value of c_1 that would be chosen if $\bar{x}$ were regarded as a certain outcome. Reducing the first argument of φ_1 raises its value, but increasing the second argument may work in either direction. In the "quadratic" case where φ_{12} is zero,

$$\bar{x}\varphi_1(y_1 - \hat{c}_1)[\bar{x} + m(1 - \bar{x}), (y - \hat{c}_1)m\sigma_x]$$

will definitely be greater than $u'(\hat{c}_1)$. Thus, with both uncertainty and money present, c_1 will have to be smaller than $\hat{c}$ in order to equate the marginal utilities of consumption and saving.

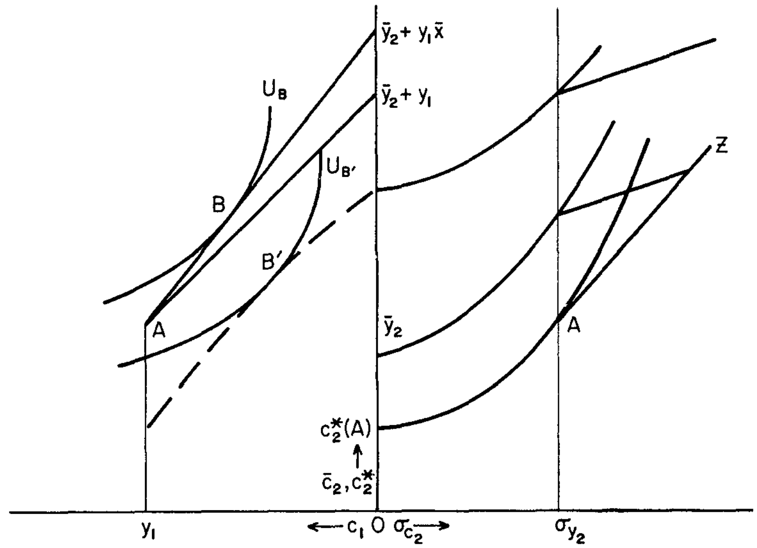

FIG. 9

possibilities are opened. Expected consumption $\bar{c}_2$ becomes

$$\bar{y}_2 + (y_1 - c_1)\bar{x},$$

as indicated by the frontier line from A in the left panel. Risk σ_{c_2} becomes

$$[\sigma_{y_2}^2 + (y_1 - c_1)^2\sigma_x^2 + 2\rho(y_1 - c_1)\sigma_{y_2}\sigma_x]^{1/2},$$

where ρ is the coefficient of correlation between second-period income and yield on saving. Conceivably, saving can reduce or even eliminate risk—provided the correlation is negative. But, in general, saving will increase expected period two consumption only by increasing its dispersion also, as illustrated by curve AZ in Figure 9. Once again the points on AZ can be converted into certainty equivalents c_2^* and the resulting pseudo-opportunity locus plotted. It is easily possible that none of these c_2^* is higher than $c_2^*(A)$, so that positive saving cannot improve on zero saving. This is the case illustrated in Figure 9.

The manner in which the availability of a safe asset modifies the pseudo-opportunity locus is the same as in Figures 7 and 8. In the case illustrated by Figure 9, introduction of a safe asset leads to a large increase in saving, all in the safe asset. Not only does this saving exceed what would be forthcoming without any safe asset available; it also exceeds the amount that would be saved if there were no uncertainty about either need or yield (point B). At the actual equilibrium point B', the individual is behaving as if the marginal reward for saving exceeded that actuarially available. Thus the departure from risk-neutral behavior may be in the opposite direction from the pure case of yield uncertainty.

Borrowing and Intermediation

It takes two kinds of people, at least, to make a market. The consumer savers of the previous section are one kind. They have a positive demand for accumulation of the less risky, monetary asset. In the market as a whole this can be satisfied only by governmental or private borrowers who create such assets. Potential private borrowers are individuals with little or no risk aversion. In the primitive diagrammatic framework introduced in the previous section, they prefer to stay on the locus *OZ*, placing all their saving in risky real investment rather than to move to the left and accumulate a diversified portfolio. The opportunity to borrow means that they can assume more risk than *OZ*. If the expected real rate of interest on their debt is smaller than that on real investment, assuming more risk also raises their expected amounts of future consumption.

The adjustment of such a borrower is illustrated in Figure 10, which follows the same format as the previous diagrams. As before, the lower dashed curve in the left panel is the pseudo-opportunity locus when only the risky asset is available. The opportunity to borrow—where $\bar{b}$, assumed greater than one but less than $\bar{x}$, is the amount to be repaid for each dollar borrowed—extends risk-return opportunities to the right of *OZ* along lines like those drawn through *P* and *P'*. The points *P* and *P'* are the preferred risk-return combinations. For example, if the individual consumes c_1', saves $y_1 - c_1'$, his real investment will be greater than his saving in the proportion $\overline{S'P'}$ to $\overline{S'Q'}$, and he will make up the difference

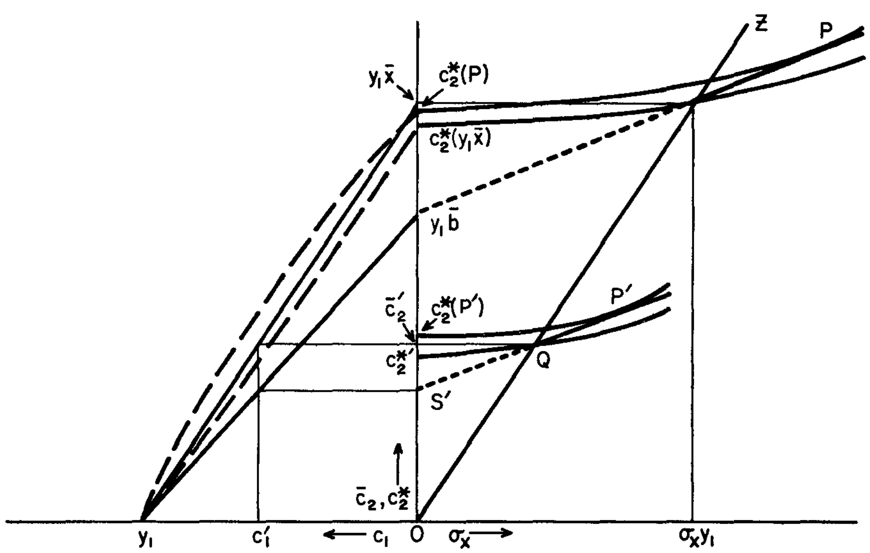

FIG. 10

by borrowing This opportunity raises the pseudo-opportunity locus in the left panel, as illustrated.

The borrower of Figure 10 is risk averse; his risk-return indifference curves are almost flat, but still upward sloping. Risk-neutral or risk-seeking borrowers would in similar circumstances have an unlimited demand for loans. They would have to be rationed by some credit line or margin requirement. After all, the borrower, under bankruptcy law, cannot lose more than his own stake; the lender is really taking the remaining risk. A credit line proportional to the borrower's commitment of his own equity would be depicted by a ray to the right of OZ and the corresponding pseudo-opportunity locus would be a straight line from y_1 steeper than the line to $y_1\bar{x}$.

The borrower then has, from a welfare standpoint, an excessive incentive to save, because the more he saves himself the more he is permitted to borrow on what he regards as profitable terms. Somehow the market must ration borrowers, and it is not clear how it can simultaneously make an efficient allocation of risk among lenders and borrowers and an efficient social choice between present and future consumption. Loan markets and financial intermediaries transfer the risks of capital ownership from conservative savers, who find such risks distasteful, to adventurous investors, who like the risks but are short of wealth. But the pattern of expected yields, interest rates, and credit lines that accomplishes this allocation does not mean the same thing for the saving-consumption choices of all the participants in the market. Whereas lenders discount the prospects of capital ownership for its risks and the prospects of lending for risks of default, borrowers limited by credit lines add to these returns on new saving the value of enlarging their capacity to borrow.

IV. Questions about Optimality

The two-period consumer can be the nucleus of a growth model with overlapping generations. Each generation works and saves one period, retires and dissaves the second period, and does no net saving over its lifetime. Alternative steady growth paths imply different sequences of first-period consumption and second-period consumption. In the absence of technical change, each path involves the same consumption sequence for every generation (Samuelson, 1958; Diamond, 1965, esp. pp. 1126–35; Cass and Yaari, 1966). The set of sustainable sequences is limited by an opportunity locus of the type ABC in Figure 11. Curve ABD is the frontier of possibilities when all saving takes the form of acquiring productive capital. Its convexity to the origin reflects the diminishing marginal productivity of capital. At point B, the golden rule point, this marginal productivity becomes equal to the rate of growth of population g. As observed above, it is inefficient to invest in capital beyond this point. Any

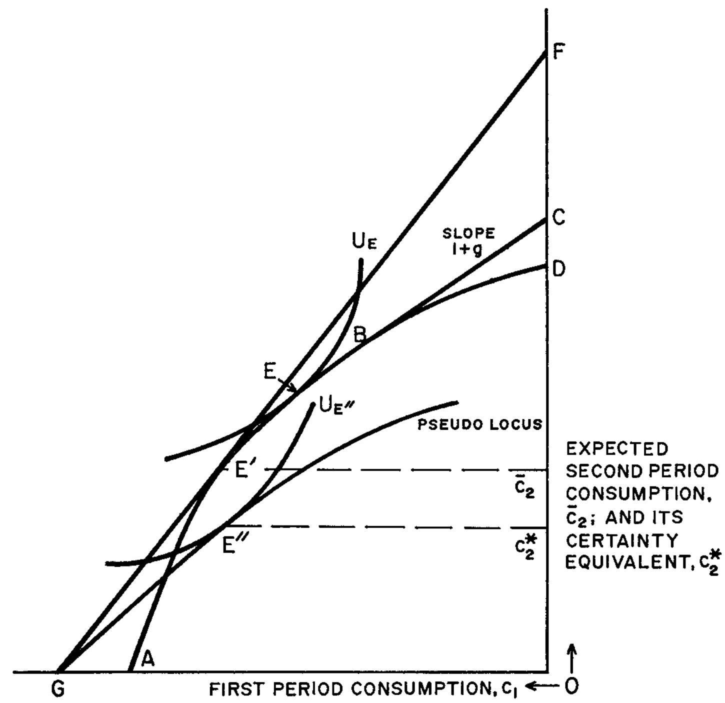

FIG. 11

further saving should take the form of paper earning a real rate of interest, g, with which each retired generation purchases some of the output of its younger contemporaries. This opportunity is represented in the diagram by the line BC with slope $1 + g$, tangent to ABD at point B.

In the absence of uncertainty, an equilibrium would be reached at a point like E, where the opportunity locus is tangent to an indifference curve of a representative consumer. The common slope would be the rate of return on saving sustained by a competitive market. This equilibrium growth path would also be the optimum. In the case illustrated, all saving is and should be acquisition of capital.

An equilibrium in which saving is influenced by uncertainty and risk aversion is exemplified by E'. Now the opportunity locus ABC refers to expected values of consumption. Each generation, when it saves, is unsure of the return its savings will earn. Although the actuarial opportunity presented by the market is described by a line $GE'F$ tangent to ABC at E', the representative individual determines his consumption and saving by reference to a psuedo-opportunity locus interior to the tangent $GE'F$. This leads him to point E'', where his first-period consumption is the same as at E'. The ordinate of E'', c_2^*, is the certainty equivalent of an expected value of second-period consumption, $\bar{c}_2$, indicated by E'.

As explained above, the availability of a safe asset bearing a lower return than physical capital brings the pseudo-opportunity locus closer to the actuarial price line. Therefore, it moves the equilibrium along *ABC* toward *E*—*provided* that all the saving goes into capital formation, that is, that the safe asset is "inside money" representing indirect investment in capital. Indeed, the equilibrium could be pushed all the way back to *E* by offering the safe asset at an interest rate equal to the expected rate of return on capital and investing in capital all saving placed in the safe asset. But it would not be possible for financial intermediation to do this if the risks that deter direct investment in capital are social as well as private risks. Not even the government can offer savers an asset that frees private saving of the risks intrinsic to the production processes of the economy as a whole.

The analysis is different for the case when the safe asset is outside money. Suppose that this bears a return of *g*, lower than the expected return of capital, and each generation of savers acquires both assets. This means—the golden rule case aside—that the society is not operating on the frontier *ABC* but inside it, along a frontier like AB^*C^* in Figure 12 where

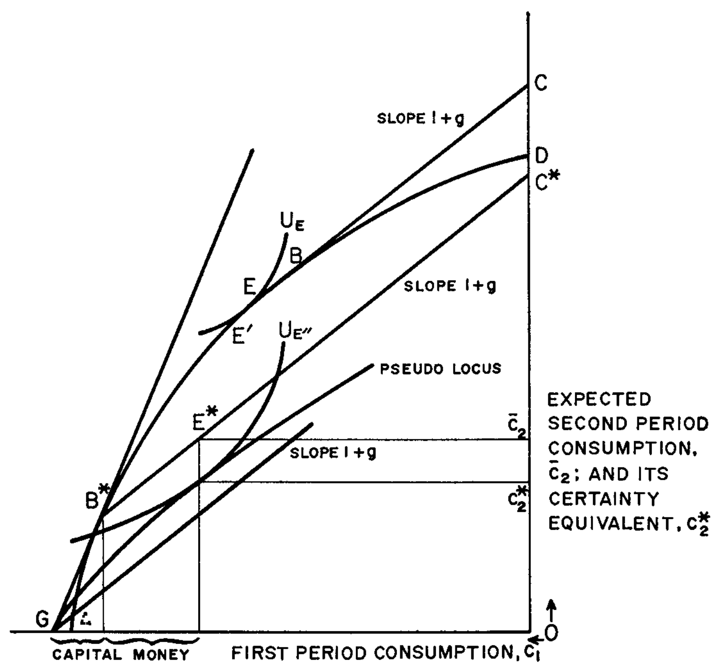

FIG. 12

B^*C^* has the slope $1 + g$ and is parallel to BC. Suppose the chosen sequence of first-period consumption and expected second-period consumption is E^*. The return on physical investment is given by the slope of the locus ABC at B^* and shown by the slope of GB^*. Saving that goes into capital formation is represented by the projection of GB^* on the horizontal axis, and saving that goes into outside money by the projection of B^*E^*. For E^* to be an equilibrium, the typical individual must wish to save the total amount indicated in the proportions indicated. In other words, E^* must lie directly above the preferred position on a pseudo-opportunity locus from point G, a locus reflecting the opportunities for combining capital bearing the return indicated by line GB^* and money bearing the interest rate g. Although the availability of the safe asset might well increase total saving, as compared with a non-monetary equilibrium with uncertainty E', it might also diminish the capitalization of the economy and even lower the expected value of future consumption.

What risks should be reflected in society's choices between present and future? There are technological and environmental uncertainties surrounding the efforts of a society to earn its livelihood from nature and from the rest of the world. If members of the society are risk averse with respect to consumption prospects, their caution should be reflected in social investment decisions. There are also private risks that are not social risks, and these should not influence one way or another the basic social choice between consumption now and consumption later. Personal hazards—for example, death, disability—should be, and frequently are, handled by insurance rather than by saving. Competition makes the risks of investment in any individual firm greater than the social risks of investment in the industry or in the economy at large. In principle, diversification should be able to cut the risks to any individual saver down to the irreducible social minimum. If so, competitive risks would not distort intertemporal choices.

However, competitive financial markets may not be able to do the whole job. Markets are not perfect; transactions costs, indivisibilities, and lack of information limit the possibilities of diversification for many savers. Private enterprise cannot guarantee the liquidity of means of payment or other financial assets against "runs" or against pervasive cyclical fluctuations in business activity or against changes in the purchasing power of currency. Only government can do these things, by making available its own obligations (possibly including purchasing-power-guaranteed obligations), by acting as a financial intermediary itself, and by guaranteeing certain private obligations like bank deposits.

The question is how the government can intervene in these essential ways while leaving to private financial markets the appraisal and accommodation of individual borrowers and the tailoring of assets to the preferences of individual lenders.

These issues are much more complicated, and also much more important, than specifying a rate at which the quantity of some narrow category of financial assets, arbitrarily designated as "money," should grow.

References

Cass, D., and Yaari, M. E. "A Re-examination of the Pure Consumption Loans Model," *J.P.E.*, LXXIV (August, 1966), 353–67.

Diamond, P. A. "National Debt in a Neoclassical Growth Model," *A.E.R.*, LV (December, 1965), 1126–50.

Johnson, H. G. *Essays in Monetary Economics*. London: Allen & Unwin, 1967.

Samuelson, P. A. "An Exact Consumption-Loan Model of Interest with or Without the Social Contrivance of Money," *J.P.E.*, LXVI (December, 1958), 467–82.

Tobin, J. "The Interest-Elasticity of Transactions Demand for Cash," *Rev. Econ. and Statis.*, XXXVIII (August, 1956), 241–47.

———. "Money and Economic Growth," *Econometrica*, XXXIII (October, 1965), 671–84.