

The Optimum Quantity of Money and the Interrelationship Between Financial Markets and Intermediaries *

Discussion of the optimum quantity of money by Clower, Friedman, Johnson, Samuelson, Stein and Tobin has led to a reexamination of the fundamental role of money in an economy. The key to much of this attempted integration of money into the theory of economic behavior has centered on transactions costs. With explicit costs of exchange introduced into the economy, the fundamental role of money in reducing such exchange costs is stressed (see Brunner and Meltzer, Feige and Parkin, Niehans, Hirshleifer and Perlman).

This note is concerned with two related consequences of achieving the optimum quantity of money: the impact on brokerage costs and the impact on the existence of securities markets. This is followed by a discussion of the alternative specifications of an optimum quantity of money policy. We then turn to analyze the relationship between optimum money balances, financial markets and financial intermediaries.

Consequences of the Optimum Quantity of Money

The basic argument in the literature on the optimum quantity of money is that since the social cost of producing money is zero, individuals should hold money balances to the point of satiety, i.e.,

* Earlier versions of this paper were sent to a number of those who have contributed to the published literature on the optimum quantity of money. Comments ranged from "interesting and agree" to "intuitively obvious" to "How can this be?". In light of such variance, I decided to leave the paper substantially unaltered and let the reader judge for himself. I would like to thank Milton Friedman and Dwight Jaffee for substantive suggestions.

to the point at which the marginal utility of money is zero. This can be achieved if interest is paid on money balances (Feige-Parkin, p. 336). Samuelson (p. 10) suggests that such a satiety policy would result in "fewer trips to the bank and to the broker... small printing costs and other costs of transactions." Earlier in his paper (p. 5) Samuelson identifies costs of transactions: "...brokerage charges, needs for journal entries, spread between bid and ask when earning assets are converted into and out of cash, etc..."

The implicit assumption underlying these conclusions is that the liquidity services provided by secondary markets for securities are substitutes for money balances. Secondary markets permit security holders immediate conversion of an asset into cash. The cost of this liquidity service has been identified with the spread between the bid and ask quotations on the security by the market-maker (Demsetz). When the public is induced to hold larger money balances, the demand for liquidity services declines. But it is not correct to conclude that this results in a decline in the bid-ask spread on securities (as might be inferred from Samuelson, *op. cit.*). A distinction must be made between cost per transaction and the total cost of all such transactions. The analysis must be put into a more general context at least to the extent of introducing a supply function for liquidity services.

Both Demsetz and Tinic have discussed the economics of liquidity services from the market-maker's standpoint. They have shown that bid-ask spreads in securities markets are negatively related to the volume of transactions. In other words, there are economies of scale in supplying liquidity services. The appropriate scenario as the economy approaches the optimum quantity of money is as follows: individuals hold larger real cash balances, reduce their demand for liquidity services, hence they make fewer bond transactions at existing bid-ask spreads. Given the fixed costs associated with providing market-making services, the bid-ask spread required to keep the market in existence rises.

The decline in secondary market transactions will drive some resources away from providing market-making services. This is the saving that arises from approaching the optimum quantity of money (see Tobin p. 846). In the limit, when the marginal utility of money balances reaches zero and the *demand* for liquidity services falls towards zero, all security transactions into and out of cash should be zero and all resources should be diverted away from market-

making services. The dynamic approach to this limiting condition is characterized by *rising* costs per transaction and a decline in total transactions (as well as a decline in total revenues accruing to the liquidity services industry).¹

Do Securities and Their Secondary Markets Really Disappear?

Secondary markets permit holders of financial assets to dispose of a security at their own discretion, rather than depending on the redemption schedule of the issuer. An apparent consequence of the optimum quantity of money is the disappearance of the need for such liquidity services provided by secondary markets. Secondary markets for securities should, therefore, disappear. An even more startling conclusion has been reached by Tsiang (p. 267) who argues that securities themselves would not exist if the optimum quantity of money were reached. The argument is quite simple: if risk-free money balances were rewarded with the same rate of interest as that on securities, the former would dominate in investor portfolios. All securities would disappear.

At first glance, Tsiang's conclusion does not seem to be correct. Securities should exist even in a world with the optimum quantity of money. There should be a distinction between secondary market services and an initial transfer of purchasing power between surplus and deficit units which gives rise to newly-issued securities. In a multiperiod context, some economic units may want to spend more than their current income and accumulated wealth permit. The flotation of securities should, therefore, take place.² The key question is: Why should individuals hold money *and* securities in the same portfolio if money holding is rewarded with the same rate of interest as that on securities?

The answer to whether securities exist in a world with the optimum quantity of money must be given in two parts. The first stage treats securities as no different from the underlying capital stock. The issue then reduces to whether optimum money balances

¹ FEIGE and PARKIN (p. 346) also show that the *number* of bond transactions declines as society achieves the optimum quantity of money. They also err, however, by assuming that cost *per* transaction is constant.

² In STEIN's review of Friedman's essay (1970, p. 41, footnote 14) the question of whether bonds exist is raised but no answer is given.

can coexist with the capital stock. The second part of the answer introduces other characteristics of securities which distinguish physical capital from "claims to physical capital" (either bonds or equities). The latter is the essence of the financial sector.

The stationary state envisioned in Friedman's essay does not rule out uncertainty (1969, p. 2). Stein (1970, p. 398) confirms that individuals are faced with uncertainty and that investment in real capital is subject to stochastic returns. This suggests that if risk-free money balances receive a rate of interest equal to the expected return on physical capital, then a risk-averse public should be unwilling to hold the latter in its portfolio.

The question can be put more concretely as follows: In the presence of uncertainty, which interest rate should be paid on money balances such that the public holds the "optimum" quantity of money? Since the optimum quantity of money is presented as a policy issue as well as a theoretical construct, the question takes on even greater significance in a world with a multitude of yields attached to securities with different risk characteristics.³

The term "optimum" quantity of money is practically a tautology in this context. First we decide which set of real costs can be eliminated by holding larger money balances and then set the interest rate on money equal to the level at which the use of such scarce resources disappears. The discussion by Tobin (p. 846) concentrates on the resources devoted to transactions between money and bonds aimed at economizing on the "working balances" held by individuals and firms. This is the same set of market-making services discussed by Samuelson (pp. 5 and 10). This argument implies that the yield on money balances should be such as to eliminate secondary market services. As a practical matter, this suggests setting the yield on money balances equal to the rate on risk-free assets. Short-term government debt would, therefore, be replaced by money.

This optimality condition still forces economic units to issue securities when their *expected* income and total accumulated wealth falls short of their planned spending during a particular time period. Can this risk be eliminated as well? Put somewhat differently, the question reduces to whether the optimum quantity of money can be defined such that it eliminates the uncertainty associated with the stochastic returns on real investment? Under such conditions any one

³ JAFFEE (p. 22) and TSIANG (p. 267) raise this question explicitly.

economic unit could never be caught short of purchasing power. If the returns on investments were, for example, normally distributed, the required size of money balances would appear to be infinite. This condition could theoretically be *approached* if the rate of interest on risk-free money balances were set equal to the expected rate of return on capital. As a practical matter, however, there would always have to be some finite yield differential to induce the public to hold both risky capital and less-than-infinite risk-free money balances in its portfolio.⁴

There are two specifications of an optimum quantity of money policy as just outlined. One concentrates on eliminating resources devoted to transactions into and out of relatively riskless assets and the other extends the optimum quantity towards removing the risk associated with holding real capital. From a policy standpoint, the question is how far should we push the optimum quantity of money? In view of the fact that risk can never be eliminated completely, it would seem appropriate to set the yield on money equal to the riskless rate and let financial institutions and markets develop to some particular level.⁵

It seems appropriate at this point to describe a number of additional characteristics of securities themselves which suggest that they will succeed in reducing risk below the level induced by an optimum quantity of money policy. In this case securities should coexist with money and physical capital in a world with the optimum quantity of money.

It is well-known that total portfolio risk, defined as variance of returns, can be reduced through diversification, i.e., by holding a portfolio of different securities whose stochastic outcomes are more or less independent of each other. Without securities, such diversification to reduce risk could be undertaken by a firm's choice

⁴ The recent article by NIEHANS (1975) argues that debt will not exist in a world with the optimum quantity of money. The reason is simple, Niehans sets the rate of interest on money equal to the marginal rate of time preference. If capital were included in his model (yielding the marginal rate of time preference) he would find that it too would not enter people's portfolios.

⁵ In a private communication, Friedman indicates that he concurs with the policy prescription of setting the rate earned on money balances equal to the risk-free rate. Thus, the Tobin and Friedman worlds of the optimum quantity of money turn out to be quite the same. Note that I have made no mention of the instability problems described by Stein and Tsiang under the optimum quantity of money policy. Since the instability occurs only when the optimum quantity of money is achieved via deflation (but not when interest is paid on money balances) it did not appear crucial to our analysis.

of real investment projects. But such "direct diversification" is less efficient in lowering risk compared with financial diversification. If a firm combines, say, two investments with negative covariance of returns, all individuals are then forced to hold these two investments in the same *proportion*. With diversity of preferences among economic units, this shrinkage of the opportunity set is inefficient.⁶ Financial claims, on the other hand, can be held by individuals in different proportions. A second disadvantage of direct diversification stems from conflicting criteria that may emerge from factors determining the optimum size of a productive organization (see Alchian and Demsetz). The financial sector provides the additional *option* of separating the "size of firm" decision from risk-diversification considerations.

We can conclude, therefore, that as long as risk is not driven to zero by an optimum quantity of money policy, the introduction of securities will tend to lower risk. We leave open the question of whether such securities take the form of equities or bonds or various combinations thereof.

With uncertainty remaining in the optimum quantity of money economy, secondary markets for securities are also saved from extinction. Secondary markets not only provide liquidity services, so that asset holders can convert securities into and out of cash, but they also permit portfolio exchanges of individual securities. Changes in the relative value of particular securities will lead to potential purchases and sales of existing securities — even if no additional money balances are desired. Of course, in such circumstances, there need not be any *actual* securities transactions in secondary markets. Market tenders would be sufficient to change relative securities prices as the underlying risks change. Secondary markets in existing securities will exist, but with a substantially different orientation. Indeed, they would more closely resemble "Walrasian markets" than currently existing ones.

Money, Markets and Intermediaries

We have discussed the relationship between market-making or liquidity services of secondary markets for securities and the optimum quantity of money. Our analysis was based on the assumption that

⁶ This point was suggested in discussions with Jan Mossin.

the marginal utility of liquidity services is lower the greater the level of real money balances. This is nothing more than a generalization of diminishing marginal utility of liquidity and the recognition that one aspect of secondary markets is aimed at increasing the liquidity of securities. Larger real money balances are substitutes (from society's viewpoint) for additional liquidity services provided by market-makers. A similar substitute relationship exists between financial markets and financial intermediaries.

Financial intermediaries buy securities issued by deficit units and issue their own liabilities that are held by surplus units. The ability of financial intermediaries to diversify more efficiently than individuals and therefore reduce the subjective risk borne by the economy was stressed by Gurley and Shaw. Intermediaries can, therefore, issue liabilities which are less risky than the primary securities issued by deficit units. In addition, intermediaries are also able to take advantage of the offsetting needs for funds by different economic units. Financial intermediaries can, therefore, also issue liabilities with greater liquidity than the primary securities purchased. That is, they can promise to redeem their liabilities for cash on short notice even without resorting to secondary markets. In other words, intermediaries perform some of the same functions as secondary markets. There is, therefore, a substitute relationship between intermediaries and markets. In our context, the marginal utility of liquidity services in the form of new secondary markets is lower in the presence of financial intermediaries. This means that individuals would be willing to pay less for secondary market transactions when intermediaries are in operation.

More generally, we can say that a substitute relationship exists between three key elements of the financial sector: real money balances, secondary market services and financial intermediaries. The latter two provide liquidity services whenever real money balances are less than the satiety level. When real balances reach the point of satiety, both secondary markets for financial assets and financial intermediaries will cease to be used as sources of liquidity. Neither is likely to disappear since each provides other services besides liquidity: intermediaries aid in diversification of risk (because of indivisibilities and returns to expertise) and secondary markets provide "valuation" services.

The evolution and development of the financial sector of the economy provides a more fruitful framework for these interrelationships than the optimum quantity of money literature. The

incentive to innovate new financial intermediaries and new financial markets depends upon the resource costs incurred in operating such institutions and the benefits that accrue from their establishment. The former is determined by the available technology (such as in communications and information storage) and the latter is determined (in part) by the initial level of liquidity services. An example will serve to indicate the nature of the development process.

We start with an economy where money, securities and real goods exist. Assume that money balances are at the satiety level because the price level is declining at the optimal rate. The marginal utility of money is zero, hence no one is willing to pay for the liquidity services that a broker would provide by operating a secondary market in securities. No bonds have to be sold before maturity because money balances are sufficiently large to meet all payments for real goods during the current period.

Assume that there is an exogenous shift in either the rate of inflation or the rate of interest so that individuals now economize on their cash balances — holding less than the satiety level. The demand for liquidity services (the ability to turn financial assets into cash immediately) now emerges and there will be an incentive for resources to be allocated to: (1) market making in securities; (2) running a financial institution which promises to turn assets into cash on short notice; (3) instituting an interest payment mechanism on money; or even (4) improving commodity markets so that the cost of exchange for real goods declines (and less liquidity is needed). Exactly *which* of these innovations occur or in what combinations they emerge depends upon relative costs.

This analysis provides insights into a number of aspects of the historical development of the financial sector of an economy: (1) innovation of financial intermediaries, secondary market services and other liquidity-related institutions is stimulated by conditions which cause economic units to economize on cash balances. These conditions include (when interest payments on money balances are prohibited) rising rates of interest on securities and increases in the rate of inflation. (2) When cash balances are held at less than satiety levels, improvement in the technology of communication and information induces an increase in secondary market services. (3) The establishment of financial intermediaries in underdeveloped countries as part of a government development program will reduce the incentive for the private sector to devote resources to running secondary markets.

Conclusions

We summarize by noting that the resources which would be saved by pursuing an optimum quantity of money policy extend from those utilized in providing secondary market services and the required volume of securities flotations to financial intermediaries. We have also noted, however, that it is unlikely that these institutions would disappear completely; nor should an optimum quantity of money policy aim at such an objective. Our discussion also indicates a broad range of historical developments in financial sector evolution which can be analyzed within this framework.

New York

WILLIAM L. SILBER

REFERENCES

- A. ALGHIAN and H. DEMSETZ, "Production, Information Costs and Economic Organization," *American Economic Review* (December 1972).
- K. BRUNNER and ALLAN MELTZER, "The Uses of Money: Money in a Theory of an Exchange Economy," *American Economic Review* (December 1971).
- R. CLOWER, "What Traditional Monetary Theory Really Wasn't," *Canadian Journal of Economics* (May 1969).
- H. DEMSETZ, "The Cost of Transacting," *Quarterly Journal of Economics* (February 1968).
- E. FEIGE and M. PARKIN, "The Optimal Quantity of Money, Bonds, Commodity Inventories and Capital," *American Economic Review* (June 1971).
- M. FRIEDMAN, *The Optimum Quantity of Money*, Chicago (1969).
- J. GURLEY and E. SHAW, *Money in a Theory of Finance* (Brookings Institution, 1969).
- J. HIRSHLEIFER, "Exchange Theory: The Missing Chapter," *Western Economic Journal* (June 1973).
- D. JAFFEE, "Transactions Cost and the (Optimal) Supply of Money," mimeo (1971).
- H. JOHNSON, "Inside Money, Outside Money, Income, Wealth and Welfare in Monetary Theory," *Journal of Money, Credit and Banking* (February 1969).
- J. NIEHANS, "Interest and Credit in General Equilibrium with Transactions Costs," *American Economic Review* (December 1971).
- J. NIEHANS, "Interest and Credit in General Equilibrium with Transactions Costs," *American Economic Review* (September 1975).
- M. PERLMAN, "The Roles of Money in an Economy and the Optimum Quantity of Money," *Economica* (August 1971).
- P. SAMUELSON, "What Classical and Neoclassical Monetary Theory Really Was," *Canadian Journal of Economics* (February 1968).
- J. STEIN, "The Optimum Quantity of Money," *Journal of Money, Credit and Banking* (November 1970).
- S. TINIC, "The Economics of Liquidity Services," *Quarterly Journal of Economics* (February 1972).
- J. TOBIN, "Notes on Optimal Monetary Growth," *Journal of Political Economy* (July-August 1968, Part. II).
- S. C. TSIANG, "A Critical Note on the Optimum Quantity of Money," *Journal of Money, Credit and Banking* (May 1969).