



## optimum quantity of money

The optimum quantity of money is most famously associated with Milton Friedman (1969). The optimum is a normative policy conclusion drawn from the long-run properties of a theoretical model. Friedman posited an environment that abstracts from all exogenous shocks and nominal price and wage sluggishness. The basic logic is then straightforward. One criterion for Pareto efficiency is that the private cost of a good or service should be equated to the social cost of this good or service. The service in question is the transactions role of money. The social cost of producing fiat money is essentially zero. Since fiat money pays no interest, the private cost of using money is the nominal interest rate. Hence, one criterion for Pareto efficiency is that the nominal interest rate should equal zero. Since long-run real rates are positive, this implies that monetary policy should bring about a steady deflation in the general price level. This famous policy prescription is now commonly called the Friedman rule.

Although most closely associated with Friedman's (1969) bold statement of the policy conclusion, the basic idea of the optimum quantity can be found in Tolley (1957), who argues, on similar efficiency grounds, for paying interest on currency. Friedman (1960) credits Tolley with this suggestion, and further notes that an alternative policy would be a steady deflation. It is curious that Friedman (1960) dismisses the 'Friedman rule' deflation as not feasible for practical purposes. Finally, the optimum-quantity result is implicit, but never noted, in Bailey (1956) who examines the welfare cost of inflation but does not consider the welfare gain of deflations.

In practice, the optimum-quantity result has had remarkably little influence on monetary policy implementation. Although many central banks pursue low inflation rates with an eventual goal of price stability, no central bank has advocated a policy that would bring about a steady price deflation. There are likely several reasons, both judgemental and theoretical, that have led to this lack of influence. I will briefly review both types of objections.

One of the first theoretical objections to the optimum-quantity results was made by Phelps (1973), who argued that Friedman's first-best argument ignored the second-best fact that money growth produces seigniorage revenues for a government, and that all forms of taxation produce distortions of some kind. If 'money' or 'liquidity' is a good like any other, then familiar optimal taxation arguments would suggest that it should be taxed via a steady inflation. This argument seems all the more persuasive given empirical estimates of a fairly low money demand elasticity.

This public finance approach spawned a very large literature. Important contributions include Kimbrough (1986), Guidotti and Vegh (1993), Correia and Teles (1996; 1999), Chari, Christiano and Kehoe (1996), and Mulligan and Sala-i-Martin (1997). These analyses were much more explicit than Friedman (1969) and

considered a fully dynamic theoretical environment with no nominal rigidities. A key relationship in all these models is the transactions or shopping function. The time spent by households shopping ( $s_t$ ) is a function of the form:  $s_t = \varphi(c_t, m_t)$ , where  $c_t$  denotes real consumption and  $m_t$  denotes real cash balances. The function  $\varphi$  is assumed to be homogenous of degree  $k$ , increasing in consumption, and decreasing in real cash balances, the latter effect motivated by the transactions function of money. Money can be thought of as an intermediate good that facilitates consumption purchases. Now suppose a central government needs to finance an exogenous level of spending and can do so only with distortionary taxes on, say, labour income, or the inflation tax on money balances. In this case, is the Friedman rule still optimal?

Most of these papers were supportive of the Friedman rule, concluding that in such a second-best environment the optimal monetary policy is a zero nominal rate. Mulligan and Sala-i-Martin (1997) argued that the result was fragile as it depended on the degree of homogeneity in  $\varphi$  and the alternative tax instruments available to the government, for example, income taxes against consumption taxes. These conflicting results have been usefully explained in DeFiore and Teles (2003), who demonstrated that the reason for the divergent conclusions is an inappropriate specification of how consumption taxes are entered in the transactions cost function. They consider a more general environment in which the government has access to both consumption and income taxes. They also consider the case where money is costly to produce at a constant marginal cost of  $\alpha$ . Further, they demonstrate that if  $\varphi$  is linearly homogenous ( $k=1$ ) then the optimal interest rate is equal to  $\alpha$ . This is a modified Friedman rule in that the private cost and social cost of money are set equal to each other, and is analogous to the Diamond and Mirrlees (1971) optimal taxation result: intermediate goods should not be taxed when consumption taxes are available and the technology is constant returns to scale ( $k=1$ ). If  $\varphi$  is not linearly homogeneous, then the optimal policy involves a tax (or subsidy) on money proportional to  $\alpha$ . Since money is essentially costless to produce ( $\alpha=0$ ) the optimal nominal interest rate is zero. DeFiore and Teles (2003) thus conclude that the Friedman rule is the optimal second-best policy for all homogeneous transactions technologies. Hence, the Phelps (1973) objection appears to be settled in Friedman's favour.

A second theoretical objection to the optimum-quantity result is that, in a world with nominal rigidities, a steady general price deflation would produce unwanted relative price movements since not all nominal prices would be adjusted simultaneously. Strictly speaking this is not a theoretical objection to Friedman (1969), as he assumed a world with perfectly flexible nominal prices and wages. But if one believes that nominal rigidities are important, and that they matter even in the long run, then this is a relevant objection to the Friedman rule. For example, in the dynamic new Keynesian (DNK) class of models (for example, Woodford, 2003) the assumed nominal rigidities have permanent effects so that any departure from price stability causes permanent movements in relative prices. Hence, these models typically suggest that optimal policy is a stable price level, and that a Friedman-rule deflation would be suboptimal. These DNK models typically abstract from the