Daniel L. Thornton is a vice president and economist at the Federal Reserve Bank of St. Louis. The author thanks John Duffy, Kevin Hoover, David Laidler, Alvin Marty, and Bennett McCallum for helpful comments. Jonathon Ahlbrecht and Stephen Majesky provided research assistance.

Money in a Theory of Exchange

Daniel L. Thornton

"In primitive traffic the economic man is awaking but very gradually to an understanding of the economic advantages to be gained by exploitation of existing opportunities of exchange...Consider how seldom it is the case, that a commodity owned by somebody is of less value in use than another commodity owned by someone else! And for the latter just the opposite relation is the case. But how much more seldom does it happen that these two bodies meet!...Even in the relatively simple and so often recurring case, where an economic unit, A, requires a commodity possessed by B, and B requires one possessed by C, while C wants one that is owned by A—even here, under a rule of mere barter, the exchange of the goods in question would as a rule be of necessity left undone."
— Carl Menger, "On the Origin of Money," The Economic Journal *(June 1892), p. 242.*

"Money, I consider, is a device which facilitates the working of markets." Sir John Hicks, A Market Theory of Money *(1989), p. 2.*

A major problem in monetary economics has been to introduce money into the economy in a way that: (1) explains how money arises endogenously, (2) explains why money is preferred to other methods of exchange, and (3) identifies the welfare gains associated with money's use. Money has been introduced by including it as an argument in consumers' utility functions or producers' production functions, assuming the existence of a welfare-reducing cash-in-advance constraint, assuming that it is a vehicle for making intergeneration transfers with no role in exchange, or simply assuming that money exists—although it is given no specific role to play.[1]

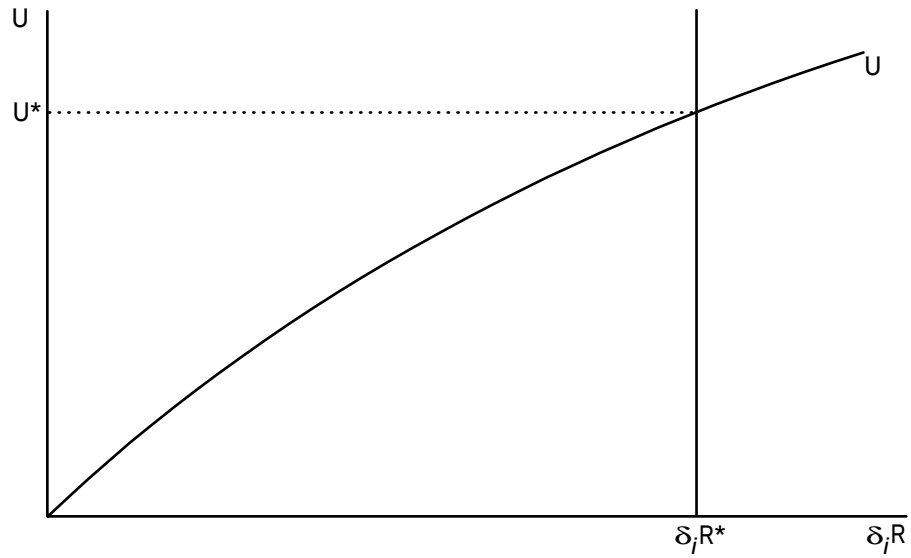
This paper develops a framework for assessing money's role and the welfare gains associated with its use. This framework shows how money reduces the resources necessary for exchange, thereby increasing both consumption and leisure. The welfare gains from trade are enhanced because the use of money promotes further trade and greater specialization. For expository purposes the analysis is linear; however, it is more correct to think of trade, money, and specialization as essentially evolving simultaneously, with the development of each reinforcing the development of the others. Nevertheless, I argue if there were no trade, there would be no need for money. To understand the role of money in an exchange economy it is not necessary to know the precipitous event that set off this evolutionary chain.

By showing how the use of money economizes on scarce resources, expands trade and promotes specialization, the analysis makes explicit Laidler's claim (1990, p. 47) that "...there is something of the nature of a public good about money" so that "...we should be very wary of treating the sum of its private products as its social product." In so doing, I show that to determine the welfare gains from money it is necessary to compare a monetary economy with economies that use alternative methods of exchange. The analysis has micro foundations, but contrary to the trend in macroeconomics, no formal, explicit general-equilibrium model of money is developed. Indeed, the analysis suggests that the obstacles to the formulation of such a model of money are considerable.

The analysis is presented in four sections. The first section develops a model of an autarkic economy and derives a measure of social welfare as a function of economic resources. In the second section, individuals are permitted to trade and several fundamental consequences of costly trade are derived and discussed. The third section argues that there only are three possible methods of effecting exchange: barter (simple and sequential), credit (simple and sequential), and money. Because economic agents have an incentive to choose the least costly method of effecting exchange, an analysis of the relative cost of each of these methods reveals why the world has been dominated by monetary, rather than barter or credit economies. The analyses in the first three sections provide the spring-

[1] See Hoover (1995) for a discussion of these and other approaches.

Figure 1



board for the fourth, a discussion of several interesting and difficult questions in monetary theory.

AN AUTARKIC ECONOMY

Many monetary analyses begin by modeling an economy with money and discuss money's implications. Since the purpose of this article is to better understand why money exists and how its use enhances welfare, it seems natural to start with an economy where there is no money. Moreover, since I argue that money's existence depends on trade, the natural starting place would seem to be a model of an autarkic economy. Consider an economy with N individuals and Q commodities. Each individual is endowed with a quantity of a non-depletable resource $\delta_i R^*$, where δ_i is the ith individual's proportionate share, $0 \leq \delta_i \leq 1$, of the total economy-wide resource, R, which is fixed and given at R^*. Individuals are self-sufficient and maximize utility, where the ith individual's utility function is

$$U_i(C_1^i, C_2^i, \ldots, C_Q^i, l^i). \tag{1}$$

C_j^i, $j = 1, 2, \ldots, Q$, denotes the quantities of the Q commodities consumed by the ith individual and l denotes the amount of time devoted to leisure. Each individual produces these commodities via the following production functions,

$$\begin{aligned} C_j^i &= f_j^i\left(\left[\delta_i R^*\right]_j, L_j^i\right) \\ j &= 1, 2, \ldots, Q, \\ \text{and } i &= 1, 2, \ldots, N, \end{aligned} \tag{2}$$

where $[\delta_i R^*]$ denotes the physical quantity of the resource devoted to the production of the jth commodity by the ith individual and L_j^i, denotes the amount of the ith individual's time devoted to the production of the jth commodity. The ith individual maximizes 1 subject to 2 and to the constraints

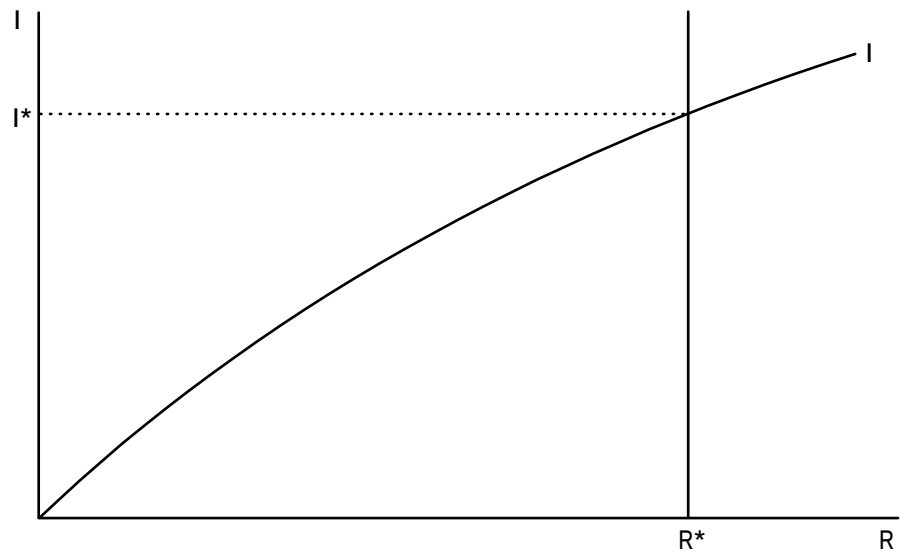
$$\sum_{j=1}^{Q} (\delta_i R^*)_j = \delta_i R^* \tag{3}$$

and

$$\sum_{j=1}^{Q} L_j^i + l^i = \Gamma. \tag{4}$$

to maximize utility, each individual must allocate resources, $\delta_i R^*$, among the production of the Q consumption goods and the total available time, Γ, among the productions of consumption goods and leisure. The solution to this optimization problem

Figure 2



results in the first-order conditions,

$$\frac{\partial f_j^i / \partial(\delta_i R^*)_j}{\partial f_k^i / \partial\left(\delta_i R^*\right)_k} = \frac{U_k^i}{U_j^i},$$
$$j,\ k,\ = 1,\ldots,Q,\ j \neq k$$

and

$$\left(\partial f_j^i / \partial L_j^i\right) = \frac{U_l^i}{U_j^i}, \quad j = 1,\ldots,Q.$$

These conditions are familiar. The first set requires individuals to allocate resources, $\delta_i R^*$, between the production of the goods that they consume by equating the ratio of the marginal utilities with the marginal rate of technical substitution for each pair of commodities consumed. The second set requires individuals to allocate time between the production of the goods that they consume and leisure by equating the marginal product of each good with respect to the labor devoted to its production to the ratio of the marginal utility of leisure to the marginal utility of that good.

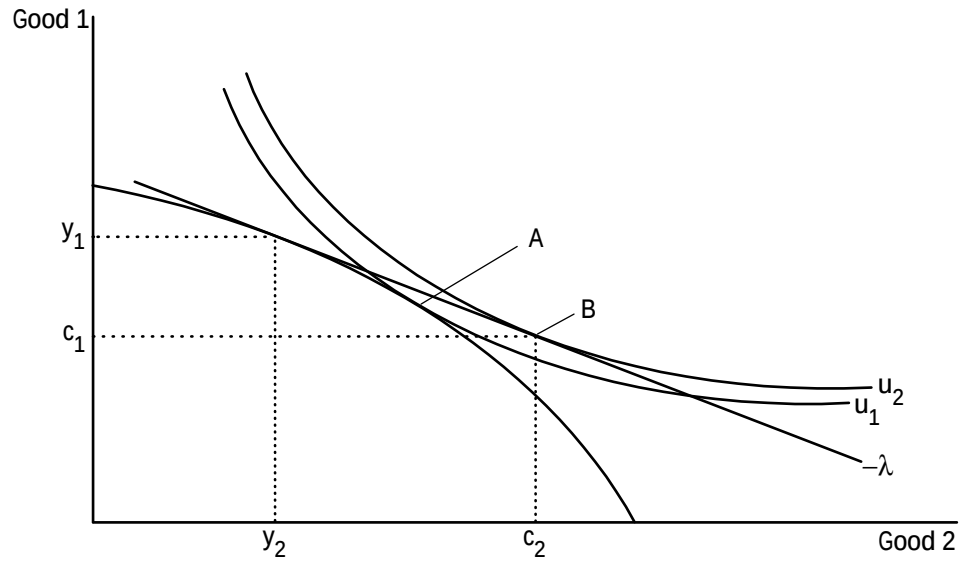
Let U_j^i, (C_1^*, C_2^*, ..., C_Q^*, l^*) denote the solution to the optimization problem for the ith individual. Assume that utility is measured ordinally, i.e., each individual assigns a real number, n, to a particular level of utility such that $U' > U$, if $n' > n$. Under the usual assumptions about preferences and production, the ith individual's maximum utility can be put into a monotonic relationship with that individual's resources, as illustrated in Figure 1. The point of interest is where U intersects the vertical line at $\delta_i R^*$. This is the point where the ith individual maximizes his well-being given the state of technology, the available resources and time.

The simple aggregation of the ordinal utility measures over the N individuals yields an index of maximum utility for society as a whole, I, shown in Figure 2. Since this index is a linear combination of monotonically increasing functions in R, for a given distribution of resources, i.e., a given set of δ's, it is a monotonically increasing function in R as well. Society's well being is maximized given the state of technology, the available resources and time at the point where I intersects society's resource constraint, R^*. This is the point of maximum social welfare for the autarkic economy.

AN EXCHANGE ECONOMY

The purpose of this section is to illustrate the effects of trade and to show how transactions costs

Figure 3



reduce the benefits from trade, thereby limiting the extent of trade. The intent is not to develop a complete model of costly trade, explain the degree of specialization that one observes, or to specify the precise benefits from trade. For simplicity, leisure is omitted as an argument in the utility function and exchange and production require only time, not additional resources, R.

Figure 3 illustrates the autarkic optimum and the gains from trade. The point A is the autarkic, no-trade optimum for individual i. At point A the slope of the production frontier $-f_2'/f_1'$ is equal to the slope of the indifference curve, $-U_1/U_2$. The point marked B is the trade optimum, given the exchange ratio, λ. By producing more of good 1 and less of good 2, relative to autarky, the individual is able to reach a level of consumption that was infeasible without trade, thereby, increasing utility from u_1 to u_2. With trade the individual produces y_1 units of good 1 and y_2 of good 2 and consumes C_1 and C_2 units of good 1 and good 2, respectively.

Now consider the effects of costly trade. The analysis is kept simple by assuming that individual 1 wishes to maximize the utility function $U(C_1,C_2)$ and can produce these goods with the following technologies:

$$y_1 = f_1(L_1)$$

and

$$y_2 = f_2(L_2),$$

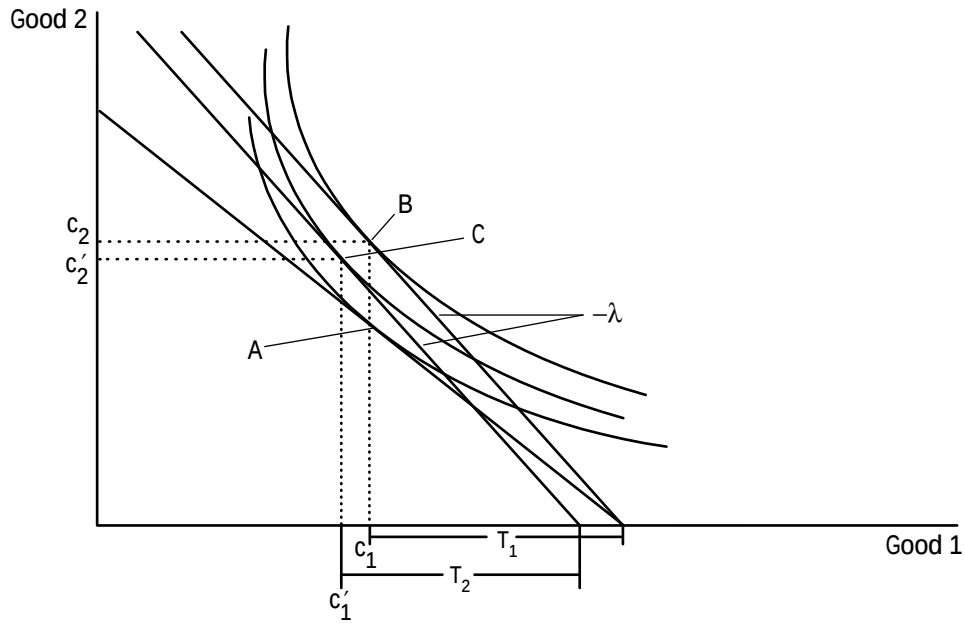
where y_1 and y_2 are the quantities of the two goods produced and L_1 and L_2 are the amounts of labor time devoted to the production of each of the two goods. Assume that the production technologies are linear, so that the marginal rate of technical substitution, f_2'/f_1', is constant. Further assume that individual 1 specialized in the production of good 1, which can be traded for good 2. Note that if exchange is costless, trade will be advantageous for any exchange ratio $\lambda > f_2'/f_1'$.

Assume that the transaction cost, i.e., the amount of time needed to trade, is fixed at Ω and does not vary with the volume of trade. Given this assumption and the others, the Lagrangian, L, can be written as

$$L = U(C_1, C_2) + \mu\big(C_2 - \lambda(f(\Gamma - \Omega) - C_1)\big). \tag{5}$$

Differentiating and solving the usual first-order

Figure 4



conditions, yields

$$\frac{U_1}{U_2} = \lambda . \tag{6}$$

Equation 6 is the familiar condition that the marginal rate of substitution equals the exchange ratio.

The effect of costly trade on optimal consumption is illustrated in Figure 4. The autarkic, no-trade optimum is denoted by *A*. The costless trade optimum, for a given exchange ratio, λ, is denoted by *B*. The fixed-cost trade optimum for the same exchange ratio is denoted by *C*. Costly trade reduces the welfare gains from trade, as the consumption of both goods is smaller when trade is costly. Note that the volume of trade, T_2, is smaller when trade is costly than when trade is costless, T_1. The reason is that trade draws resources, in this specific example time, away from production. Indeed, if the cost of trade, Ω, is large enough, no trade will take place—the autarkic optimum will dominate the trade optimum.

The most important thing to notice, however, is that the exchange ratio that is required to achieve the costless trade outcome is larger when trade is costly. This is seen by noting that to a first-order approximation,

$$\lambda \geq \frac{f_2'}{f_1'} + (f_1(\Gamma) - C_1^A)^{-1}\Omega ,$$

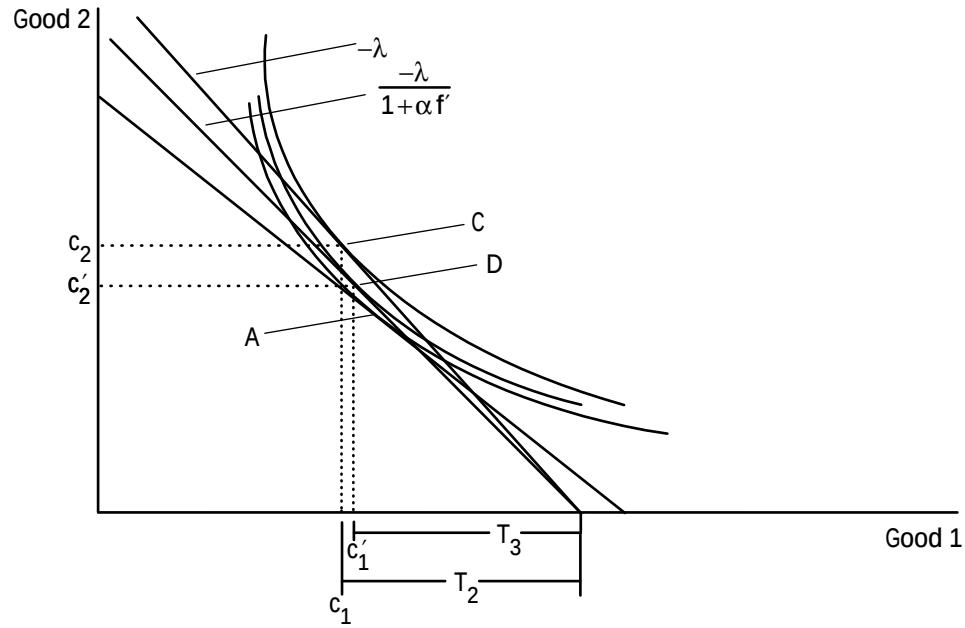
where C_1^A,is the optimum consumption of good 1 in the autarkic economy (see the appendix for details). When trade is costly ($\Omega > 0$), the individual must be compensated for the cost per unit of y_1 that must be given up to make the trade.

This effect of costly trade is more apparent when exchange costs vary with the volume of trade. The exact outcome depends on the assumption made about the nature of these trading costs, however, the basic effect of costly trade will be invariant to their nature. Hence, for simplicity, assume that the exchange cost, L_e, the time that is required to trade, is proportionate to the volume of trade, i.e.,

$$L_e = \alpha(f(L_1) - C_1), \quad \alpha > 0.$$ [2]

[2] Note that given this specification, the marginal exchange costs of using one more unit of time for exchange rather than production, i.e., $dL_e/dL_1 = \alpha f'$, depends on the marginal product of labor. The greater the marginal product of labor, the larger is the marginal cost of exchange. This fact may help explain why some individuals specialize in production and others specialize in marketing or exchange. For example, the stereotypical western *storekeeper* is someone who cannot hunt, rope, ride, or steal.

Figure 5



Again, the individual is assumed to specialize in the production of good 1. Moreover, the total amount of time, Γ, is used either in the production of good 1, L_1, or exchange, L_e, i.e., $\Gamma = L_1 + L_e$. The LaGrangian for this maximization problem is

$$L = U\big(C_1, \lambda(f(L_1) - C_1)\big) + \mu(\Gamma - L_1 - \alpha(f(L_1) - C_1)). \tag{7}$$

The first-order conditions are:

$$U_1 - \lambda U_2 + \mu\alpha = 0$$

$$\lambda U_2 f' - \mu - \alpha\mu f' = 0$$

$$\Gamma - L_1 - \alpha(C_1 - f(L_1)) = 0$$

Solving the first-order condition yields,

$$\frac{U_1}{U_2} = \frac{\lambda}{(1 + \alpha f')}. \tag{8}$$

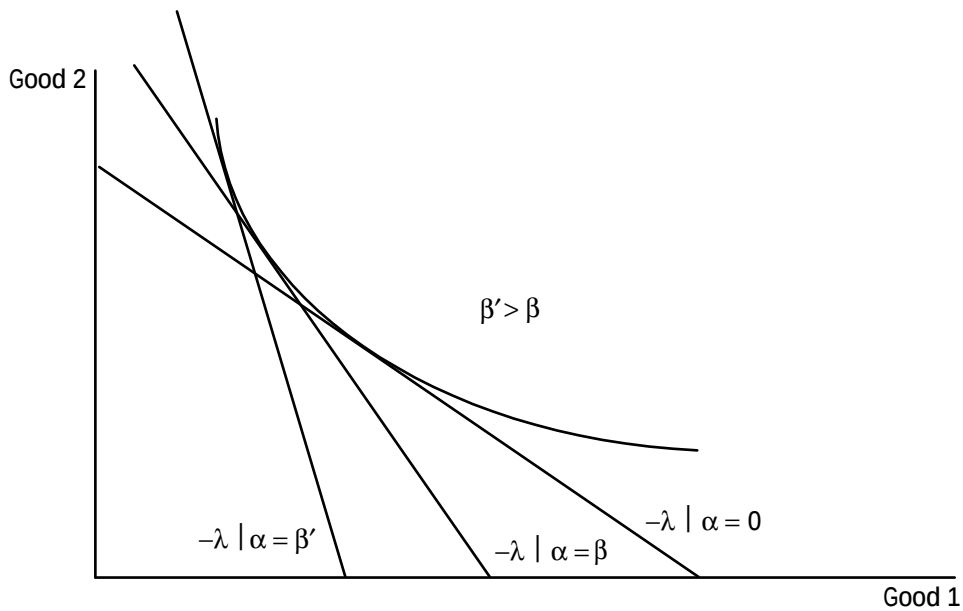
If $\alpha = 0$, this condition reduces to the previous one. When there are no exchange costs, an individual who specializes in the production of good 1 is better off trading whenever the exchange ratio, λ, is greater than or equal to the individual's fixed marginal rate of technical substitution, i.e., $\lambda > f_2'/f_1'$. When exchange costs vary with the volume of trade, however, the marginal condition for trade becomes $\lambda > (1+\alpha f_1')(f_2'/f_1')$.[3]

The effect of marginal exchange costs is illustrated in Figure 5. The point marked C is the same as that in Figure 4; namely, it is the optimal point assuming that the exchange costs are fixed. D denotes the optimum when exchange costs vary with the volume of trade. The effect of variable trade costs is to reduce the effective exchange rate for a given exchange ratio, λ. The gains from trade are smaller than when exchange costs are fixed and there is a corresponding reduction in the volume of trade.[4] All other things being the same, the volume of trade falls from T_2 to T_3. The trade optimum, for a given λ, is pushed closer to the autarkic no-trade optimum, which again is denoted by A. If the exchange costs are sufficiently high,

[3] The strict inequality is due to the fact that an individual must be compensated for the total amount of y_1 that must be given up to make the trade.

[4] Figure 5 is drawn on the assumption that the total cost of trade in this example is exactly equal to the fixed costs of trade in the previous one, i.e., $\alpha(f(L_1)-C_1) = \Omega$.

Figure 6



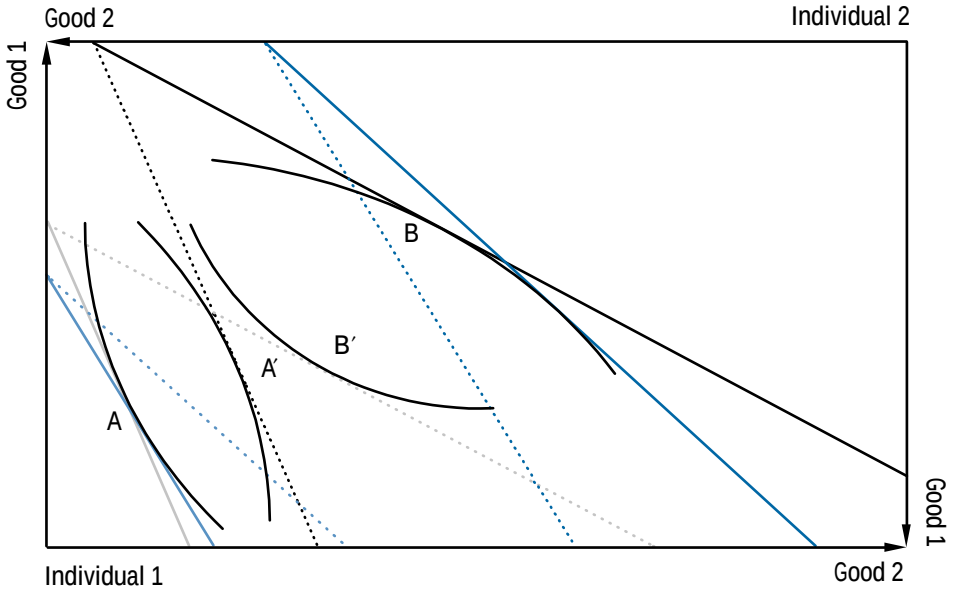
the autarkic optimum will dominate the trade optimum for a given λ.

Costly trade not only reduces the gains from trade, but more importantly, it increases the minimum exchange ratio that is required for the individual to benefit from trade. Define the *reservation exchange ratio, rer*, to be the minimum exchange ratio required for an individual to acquire the same level of utility that would be acquired under autarky. If exchange is costless, *rer* is simply an individual's marginal rate of technical substitution. Figure 6 shows *rer* for three assumptions about exchange costs: There is no cost of exchange, i.e., $\alpha = 0$, and two cases where exchange costs are positive, $\alpha = \beta > 0$ and $\alpha = \beta'$, $\beta' > \beta$. Figure 6 illustrates that the higher the exchange cost, the larger is *rer*, i.e., the more of good 2 that an individual who specializes in the production of good 1 must get to compensate for the cost of trade. Trade is advantageous only when the terms of trade are sufficiently favorable, i.e., λ is sufficiently large to compensate for the cost of trade.

All of the above conclusions were predicated on the assumption that the individual specializes in the production of good 1. Hence, it is important to see how costly trade affects the potential exchange between individuals where the comparative advantage is well defined. Assume that individuals 1 and 2 are able to produce both goods 1 and 2. Again, the production technologies are assumed to be linear and the solid gray and black lines, respectively, in Figure 7 denote their production frontiers. Individual 1 has a comparative advantage in the production of good 1; individual 2 has a comparative advantage in the production of good 2. The exchange ratio at which trade can take place, λ, must be between the slopes of the solid gray and black lines which, in the case of costless trade, represent individual 1's and 2's *rer*, respectively. The points *A* and *B* are optimal if each individual is self-sufficient. Point *A′* denotes a trading possibility where all of the gains from trade accrue to individual 2, while *B′* denotes a trade possibility where all of the benefits from trade accrue to individual 1.

Now assume that both individuals have marginal exchange costs, i.e., $\alpha_1 f_1'$ and $\alpha_2 f_2'$, which are positive but not necessarily equal. The effect of costly exchange is to raise the *rers* for both individuals. The solid light blue and dark blue lines represent the *rers* for individuals 1 and 2, respectively, when trade is costly. The dashed dark blue line is parallel to the solid light blue line and, hence, denotes the maximum benefits from trade when trade is costly if all the benefits from trade accrue to individual 2. Likewise, the dashed light blue line is parallel to the

Figure 7



solid dark blue line and, hence, denotes the maximum benefits from trade when trade is costly if all the benefits accrue to individual 1. The maximum benefits from trade are clearly less when trade is costly and diminish as the cost of trade increase.

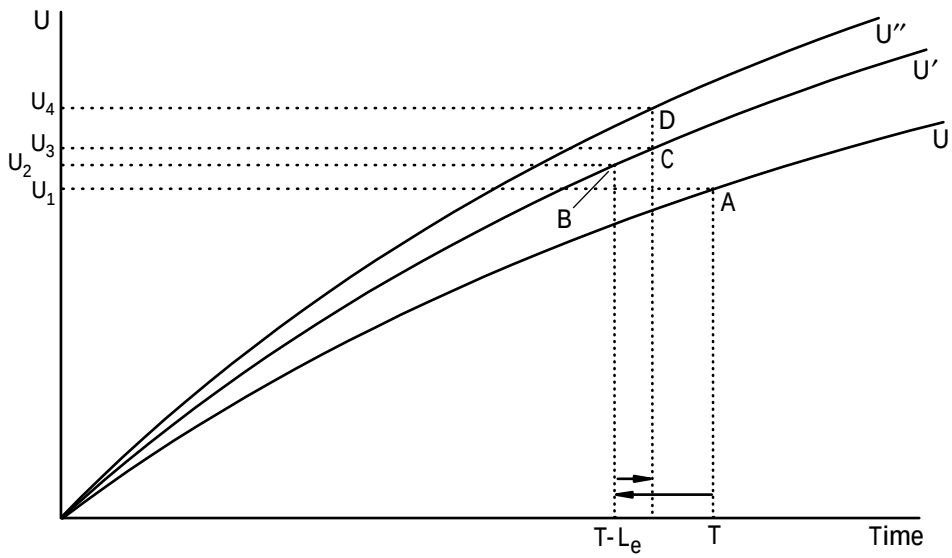
The important thing to note is that while the precise gains from trade for the two individuals depend on the respective size of the transactions costs, costly trade reduces the range of exchange ratios where trade is mutually advantageous and, therefore, the volume of trade. Moreover, the larger the exchange costs, the smaller the region where trade is mutually beneficial. Indeed, if the slopes of the solid light blue and dark blue lines were sufficiently large for either individual, no exchange ratio would exist where trade would be mutually advantageous—no trade would take place.[5] Costly trade reduces the feasible set of opportunities where trade is mutually advantageous.

The exchange ratio at which individuals trade and how much each benefits from trade depends on the relative costs of trade for both individuals, which in turn depends on strategic considerations that go well beyond the scope of this inquiry. For example, the above analysis assumed that the costs were borne by both traders and that there were no social arrangements for sharing the costs. Moreover, there is nothing in this analysis that ensures these individuals trade or that more trade takes place if the exchange costs are reduced. The conclusion that trade increases is inferred from noting that individuals have an incentive to engage in mutually advantageous trade up to the point where the marginal resource cost-of-trade equals the utility gains from trade. Anything that reduces exchange costs gives rise to this potential by expanding the feasible set of mutually beneficial trades.

The existence of exchange costs and the desire to reduce them has implications for the development of markets and, more generally, for their structure. Anything that reduces transactions cost encourages greater trade and specialization. At the same time, the benefits from specialization and trade encourage the use of the most efficient method of exchange. Of course, the catalyst for all of this is the heterogeneity that makes trade mutually advantageous. Reducing the cost of trade enhances welfare by (a) reducing the amount of resources that

[5] Since trading costs are positively related to the real volume of the goods traded, the *rers* will vary with the level of trade. This makes determining the exact amount of trade in costly trade environments very difficult. Moreover, it has been assumed that the trade costs are proportional to the volume of trade, but this need not be the case. Trading costs also may vary across individuals or goods. Regardless of how exchange costs are treated, the same fundamental conclusion emerges: The larger the exchange costs the less trade will take place.

Figure 8



must be devoted to exchange, freeing up resources for production (and/or leisure) and by (b) increasing the amount of trade that takes place, i.e., increasing the extent of the market.

A host of mechanisms have evolved to reduce the cost of exchange: bazaars, trading posts, retail establishments, brokers, agents, dealers, and other specialists and, most especially, money. Some of these have given way to more efficient methods of exchange. Others have not—at least not yet.

The effect of innovations, such as money, which reduce the cost of exchange, is illustrated in Figure 8. Point *A* denotes the level of utility for an individual who is self-sufficient. All time is spent in production or leisure. Point *B* denotes the level of utility associated with costly trade. The utility level is higher than for autarky despite the fact that some time, L_e, is used for exchange. Money reduces exchange costs, so that fewer resources are devoted to exchange, and there is more time for production (or leisure). This gain can be seen by comparing points *B* and *C*. In addition, the use of money increases welfare by expanding the set of feasible transactions and, thereby, increasing the volume of trade. This is illustrated by the difference between points *C* and *D*. The total gains from reducing exchange cost are illustrated by the difference in utility levels associated with points *B* and *D*. Applying the same aggregation analysis as before yields the implication that any reduction in exchange cost is welfare enhancing for society as a whole. This analysis makes it clear that by reducing the transaction cost, money expands the set of exchange ratios where trade is mutually advantageous. To this extent, the use of money expands the feasible set of transactions.[6]

The Implications of Costly Trade

The objective of the above analysis was to illustrate a role of money and the benefits from money's use, not to construct a general theory of trade. Indeed, the analysis says nothing about who trades with whom or how much trade takes place. Rather, it illustrates that trade is welfare enhancing and that costly trade reduces welfare (relative to costless trade), both by drawing resources from production or leisure and by limiting the extent of the markets. Money increases economic welfare by mitigating some of these costs. Moreover, welfare is enhanced even though money does not appear in the utility or production functions or is a prerequisite for trade through a cash-in-advance

[6] For the view that money does not increase the set of feasible transactions, see Ostroy, (1973), pp. 608-9.

constraint; nor are there legal restrictions requiring the use of money. Furthermore, there is nothing to rule out the possibility that some transactions are achieved using barter or credit. The analysis confirms Brunner and Meltzer's argument (1971, p. 804) that "...the private and social productivity of money are a direct consequence of the saving in resources that the use of money permits and of the extension of the market system that occurs because of the reduction in the cost of making exchanges."[7]

The conclusion that money enhances economic welfare by reducing exchange costs is independent of the market structure, so long as there is some degree of decentralization.[8] Of course, the exact nature and extent of the benefits from innovations that reduce exchange costs depend on such factors. Consequently, any attempt to quantify the benefits associated with innovations that reduce exchange costs is necessarily stylized: Specific results will depend on the assumptions made about the structure of markets, production technologies, the nature and extent of the exchange costs, and who bears them. Since my purpose is to gain insight into how money ameliorates exchange costs, it is essential to deal with these costs very generally. Before turning attention to the issue of exchange cost, however, several other implications of the above analysis for money are noted.

First, innovations that reduce the exchange cost of one individual can benefit others. This implication is clear from Figure 7. Instead of both individuals experiencing a reduction in exchange cost, assume that only one individual does. The effect still is to increase the feasible region of exchange ratios where exchange could take place. Consequently, an innovation that reduces the exchange cost of one individual can raise the utility of both.

Second, individuals have an incentive to use the least-costly method of exchange. The fact that some form of money has evolved in every society suggests that money is efficient relative to other methods of effecting trade. Moreover, that similar assets have functioned as money in very different societies suggests that certain assets seem to have a distinct advantage in reducing exchange costs.

Third, exchange costs limit the extent of trade and, hence, specialization and the use of money reduces these costs. Consequently, it is not surprising that the division and specialization of labor appear to have evolved simultaneously with the use of money.

Finally, and perhaps most importantly, the welfare gains from money can be obtained only by comparing a monetary economy with an economy that uses an alternative method of exchange. Moreover, the welfare gains from money will change as markets develop and economies become increasingly specialized. Consequently, the more highly specialized the economy and the greater the extent of trade, the larger the likely effects of disruptions to the supply of money will be. As we will see, this point has implications for the welfare costs of inflation.

It has been recognized for some time that economies with a medium of exchange are better off than if no such medium of exchange exists. Indeed, McCallum (1983a, p. 24) uses "the *traditional presumption* that an economy with a medium of exchange is more productive than it would be if no medium of exchange existed" to show that overlapping generations models of money (e.g., Wallace, 1983) have no role for money as a medium of exchange. The above analysis refines McCallum's point by explicitly showing how money necessarily enhances welfare by facilitating trade. Models that do not explicitly recognize this role of money are unlikely to capture money's essential feature.[9]

---

[7] Despite the large amount of resources devoted to market activity (bringing buyers and sellers together) economic analyses have focused on production and consumption. For an exception, see Hirshleifer (1973).

[8] Ostroy (1973) was one of the first to observe that the Walrasian general equilibrium market had no role for money because no trade takes place until the equilibrium set of *accounting prices* Patinkin (1965), is determined. This is why Meltzer (1995), Hicks (1989), and others argue that such models may be of limited use in understanding the role of money in the economy.

[9] Even models that explicitly capture the medium-of-exchange function of money do not necessarily capture the welfare enhancing properties of money noted here. For example, the shopping-time model of McCallum and Goodfriend (1987) or the money-in-exchange models of Dornbusch and Frenkel (1973), Benhabib and Bull (1983), and Fischer (1986) explicitly recognize the medium-of-exchange function. In the latter models, however, the metric for measuring aggregate welfare provides no motive for exchange. In these models welfare is measured by aggregate consumption, which is given by,

$$C = f(k)\left[1 - v(m)\right]; \quad 0 \le v < 1,\ v' < 0,$$

where $v(m)$ is the proportion of total output, $f(k)$, that is used in exchange. Since C for $v > 0$ is always less than C for $v = 0$, exchange does not enhance economic welfare. Hence, there is no motive for exchange and, consequently, no motive for money as a medium of exchange. The problem is analogous to that of Tobin (1965) and Mundell (1971), where per-capita output is maximized when money holdings are zero.

This is not true of the search-theoretic models of money (e.g., Jones, 1976; Oh, 1989; Kiyotaki and Wright, 1989, 1991, 1993; Trejos and Wright, 1993; Johri, 1994) that explicitly model money's role as a medium of exchange. In these models, money facilitates exchange by ameliorating the search costs associated with the *double coincidence of wants* essential for barter. Implications of some of these models have been born out experimentally (e.g., Duffy, 1998).

Before discussing why money exists and why money dominates barter and credit as a medium of exchange, one final point should be made. Specifically, the welfare gains from money discussed above are those associated with the real stock of money, i.e., money's purchasing power. If money is held solely to facilitate transactions, a larger real money stock means that more transactions are facilitated and, hence, the welfare gains are larger relative to the next best method of exchange; therefore, society's welfare should increase with the equilibrium stock of real money. Changes in the nominal stock of money, however, do not necessarily result in an increase in the equilibrium stock of real money. Given classical neutrality and Archibald and Lipsey's (1958) invariance principle, *ceteris paribus*, increases in the stock of nominal money may have no significant effect on economic welfare.[10]

EXCHANGE COSTS WITH ALTERNATIVE METHODS OF EXCHANGE

The previous section showed why anything that reduces the exchange cost is welfare enhancing. Such innovations enhance welfare by reducing the quantity of society's scarce resources devoted to exchange, freeing up resources, time for production, or time for leisure, and by enabling society to achieve a greater extent of specialization and trade. I inferred that money is one such innovation without carefully defining what money is. This section takes up this issue. Specifically, money is defined as a commodity that is a generally acceptable medium of exchange.

The essential feature of an exchange economy is that individuals trade the commodity that they have, commodity *j*, for one that they want, commodity *k*. The essential point is that there are only three possible methods of exchange: barter, money, or credit. In an exchange economy, trade must take place with one of these methods.[11] Which of these methods is used depends on their relative costs in effecting exchange. Hence, the analysis of money necessarily requires an analysis of the relative costs of barter, credit, and money in exchange.

In discussing the relative costs of these alternative methods of exchange, it is important to distinguish between simple barter (trading commodity *j* for commodity *k*) and multistage or sequential barter (trading commodity *j* for commodity *h* and trading commodity *h* for commodity *k*). Money always entails a sequential transaction (trading commodity *j* for *m* and trading *m* for commodity *k*). Hence, a monetary transaction can be thought of as a sequential barter transaction that involves a particular commodity, *m*. When *m* becomes generally acceptable, it is money.[12]

It also is important to distinguish between simple and sequential credit transactions. A simple credit transaction involves trading commodity *j* for a promise of some commodity (*k*, *j*, *m* or some other commodity) in the future. Hence, a simple credit transaction is just an intertemporal barter transaction.

A sequential credit transaction involves trading commodity *j* for an IOU and trading the IOU for commodity *k* or perhaps another IOU.[13] According to this definition, a sequential credit transaction is a particular form of a sequential barter transaction where the intermediate commodity is an IOU. If a particular IOU were *generally acceptable*, it would be *money*. Simple credit transactions have been used to effect exchange for a long time, e.g., trade credit. IOUs have even circulated as a form of local currency for relatively short periods of time. There are a number of reasons, however, why money initially took the form of tangible commodities and not IOUs. Indeed, I will argue later that the use of credit for the purpose of facilitating trade is due to the existence of money. Hence, while credit can supplement money in effecting exchange, it will not supplant it. In this section, however, I only will consider the question of how well credit can facilitate exchange. For this purpose, sequential credit is required.

It is important to distinguish between costs that are independent of the method of exchange and those

[10] See Bullard (1999) for a survey of the evidence for monetary *neutrality* and *superneutrality*.

[11] McCallum (1985) has also made this point.

[12] In the search literature on money, *general acceptability* is achieved in various ways. For example, in Oh's (1989) model where individuals search randomly, a dominant medium of exchange emerges due to the assumption that one commodity has the largest subjective probability of trade. He shows that if traders try to minimize the number of encounters that result in the desired trade, e.g., commodity *j* for commodity *k*, this commodity will emerge as the dominant medium of exchange. Money and barter coexist because barter occurs when an individual who has commodity *j* and wants commodity *k* just happens to meet an individual with commodity *k* and who wants commodity *j*.

[13] Brunner and Meltzer (1971) consider what they termed a barter-credit economy. In their discussions, credit is simply bartered for goods. That is, one person gives the other an IOU for the goods that the former wants. But this implicitly assumes that the latter person wants the IOU and not some other commodity that is desired for current consumption. The case where credit is used in barter transactions is discussed later.

that vary with the method of exchange. That is, it is essential to differentiate between costs that money can ameliorate and those that it cannot.[14] To this end, exchange costs, i.e., all costs associated with making the exchanges of two or more commodities between two or more individuals, are categorized as either information costs or noninformation costs.

Noninformation costs are the packaging, handling, and other assorted costs associated with getting a commodity from the point of final production to the point of final consumption. Such costs are independent of whether the transaction is achieved with barter, credit, or money. Hence, while such costs are essential for determining the extent of and the benefits from trade, they are immaterial for the broader question of why money is used and for determining the welfare gains from its use.[15]

Information costs vary with the method of exchange. Information costs are divided into *assurance* costs and *shopping* costs. Comparisons of the assurance costs associated with money, barter, and credit explain why money dominates sequential barter or sequential credit in exchange. Comparisons of the shopping costs associated with money, barter, and credit add to this explanation.

Assurance Costs

Consider first the case of multistage transactions. This is when individuals trade a commodity they have for one that they currently do not. This process will continue until they obtain the desired commodity. Such multistage transactions require that individuals obtain assurance that they will be able to obtain the desired commodity, *k*. Broadly speaking there are two distinct, although not mutually exclusive, categories of costs associated with obtaining this assurance. The first of these I term *verification costs*. Verification costs, which are discussed extensively by Brunner and Meltzer (1971) and Alchian (1977), are the costs of verifying the characteristics and attributes of the good received. These costs include the costs of inspection, measuring, perfecting property rights, etc.[16]

The second category of costs I call *value-determination costs*. Value-determination costs are the costs associated with determining the value or worth of the commodity received. The value of the commodity is the number of units of it that must be traded for *x* units of the desired commodity, *k*. Clearly, a poor-quality commodity is less valuable than a high-quality commodity. Nevertheless, knowing the quality of the product does not necessarily mean that you know its exchange value, hence, it is useful to treat these information costs as separate and distinct.

A sequence of barter transactions that ultimately leads to the acquisition of commodity *k*, requires verification costs at each stage in the sequence.[17] Hence, the verification costs associated with sequential barter could be considerable, especially if a large number of intermediate transactions are required. Money economizes on verification costs relative to sequential barter for two reasons. First, money has relatively low verification costs. Indeed, other things being the same, the commodity with the lowest verification cost will emerge as money (Jevons,1875).[18] Second, the use of money means at most two transactions, *j* for *m* and *m* for *k*, are required.

The verification costs of sequential credit transactions are likely to be high. If one person exchanges commodity *j* for an IOU of Mr. Smith, he has no difficulty in verifying that this is the IOU of Mr. Smith; after all, he watched Mr. Smith write it.[19] When he attempts to trade Mr. Smith's IOU for commodity *k*, however, the verification costs for the next person are likely to be significant. There may be considerable difficulty in verifying that this IOU is the promise of a particular Mr. Smith.[20] Moreover, even if

[14] In this context, it is somewhat arbitrary to assume where production ends and exchange begins; however, Hirshleifer (1973) suggests treating transportation costs as part of production costs.

[15] Of course, there could be second-order effects. Specifically, an innovation to the method of exchange could significantly increase the extent of the markets, which may in turn reduce transportation costs if there were economies of scale in transportation. Such innovations also could foster innovations in the transportation industry.

[16] Alchian (1977, p. 134) argues that it is the low verification costs alone that make a commodity money.

[17] See Jones (1976), Oh (1989), Kiyotaki and Wright (1993) and Trejos and Wright (1993).

[18] Jevons' observation is explicitly modeled by Jones (1976) and Oh (1989), who assume that one good is more *in demand* than other goods. Specifically, they take the "subjective expected transaction costs"—the time spent searching for complimentary trading partners—to be the number of encounters one anticipates before completing a single trade.

[19] Search-theoretic models abstract from the problem of credit by assuming that chance meetings of individuals have a Poisson distribution, so that the probability of the same individuals meeting twice is infinitesimally small relative to the probability of meeting once.

[20] Of course, methods have been developed to deal with such verification costs, but the costs still may be rather high relative to the verification costs of money.

one knew it was the IOU of a particular Mr. Smith, one might not know Mr. Smith's willingness and ability to honor the obligation. Even if the IOU is collateralized, all subsequent holders of the IOU must evaluate and perfect their interest in the collateral. Because of the costs associated with such activities, it seems likely that the verification costs of sequential barter in IOUs will be higher than the verification costs of sequential barter in commodities. Consequently, if money has lower verification costs than sequential barter, it must have lower verification costs than sequential credit as well.

The individual also must determine the value of the commodity, i.e., the number of units of the commodity k (or h) that they can get for x-units of the commodity j. The value-determination costs of sequential barter are high because it requires that the individual know up to $Q(Q-1)/2$ relative prices. The problem associated with a multitude of prices is exacerbated if credit is used to effect trade. Credit instruments can be denominated in any one of the Q commodities for any one of the N individuals. Hence, the credit price of a commodity can vary across goods and individuals. Credit prices also can differ in other dimensions, such as the maturity of the contract, whether there is collateral and the nature and extent of the collateral. Given problems associated with asymmetric information, it seems that the cost of determining the value of IOUs is so high that it eliminates the possibility that a credit instrument—denominated in a nonmoney asset—could serve as an effective, generally acceptable medium of exchange, i.e., it could serve as money.[21] Money has significantly lower value-determination costs than either sequential barter or sequential credit because traders are required to know at most $Q-1$ money prices.

Value-determination costs also depend on the variability of the value of money. All other things being the same, money serves best as a medium of exchange when its value remains relatively stable. Because of the difficulty in determining the relative value of commodities, it would be surprising to find that a commodity whose value fluctuates considerably—relative to other commodities—serves as a generally accepted medium of exchange, i.e., serves as money.

Maintaining the stability of money's value over long periods of time is important for what Jevons (1875, pp. 5-6) called the *standard of value* function of money. Because of money's role as a medium of exchange, and hence a source of generalized purchasing power, it is convenient to denominate credit contracts in terms of money. That is, the existence of money facilitates the use of credit. Indeed, as Hoover (1988) has noted, and I will argue later in more detail, money appears to be essential for credit. Variation in the value of money can have a significant, detrimental effect on money's standard of value function.[22] I will argue later that this effect may be larger than the effect of variation in the value of money on money's function as a medium of exchange.

Shopping Costs

If an individual who has commodity j and wants commodity k runs into an individual who has commodity k and wants commodity j, is the result necessarily barter? Some insights into the answer to this question come from considering shopping costs. Shopping costs are of particular interest because these are the costs that money is ordinarily thought to ameliorate relative to simple barter. Shopping costs encompass a wide variety of costs, including costs associated with locating a seller or buyer, haggling over price, budgeting, and planning expenditures. Shopping costs that are of particular interest are those associated with the lack of a *double coincidence of wants* and those associated with not having a common *unit of account*. While the medium of exchange and unit of account functions are separable, the fact that most often the same good has performed both functions suggests that this arrangement is efficient.

Trade is a planned activity. Sellers seek buyers and buyers seek sellers. Trade that occurs as the result of chance meetings is rare.[23] Indeed, special-

[21] Note that if credit were to supplant money, credit contracts would have to be denominated in something other than money. Moreover, it should be clear from this discussion that credit is less likely to arise as a general method of trading goods in primitive economies, where the costs of acquiring information are relatively high. Credit is more likely to exist as the primary method of exchange in advanced societies where information costs are relatively low.

[22] Jevons (1875, p. 6 and p. 12).

[23] In the search-theoretic literature, e.g., Jones (1976), Oh (1989), Kiyotaki and Wright (1993), and Trejos and Wright (1993), trade results from chance encounters. In such settings, money arises because it minimizes the costs of transactions due to a double coincidence of wants. Traders bump into each other randomly and engage in simple barter if there is a double coincidence of wants, they trade goods for money if one of the traders has money, or they do nothing. Recently, search models of money have allowed for the development of longer-term relationships, see Corbae and Ritter (1998).

ists who are particularly efficient at verifying and determining the value of particular products arise. In some cases, these specialists make markets for specific commodities by dealing in them. Matching *buy* and *sell* orders and managing inventories is difficult if a whole host of commodities is traded for the commodity(s) that the specialist deals in. Hence, the specialist's function is more efficient if transactions are carried out in a single commodity. Efficiency is further enhanced if *bid* (offers to buy) and *ask* (offers to sell) prices are quoted in this same commodity.

The efficiency of the market is increased if all participants agree to use the same medium of exchange and if this commodity also serves as the unit of account.[24] The elimination of the double coincidence of wants reduces the time for buyers and sellers to locate each other. Money also reduces the time spent haggling over price if everyone agrees to quote prices in the same commodity and if that commodity is accepted generally in exchange. For analogous reasons, planning and budgeting are made easier if a single commodity is used as both the medium of exchange and the unit of account.

Let us now reconsider the intriguing question that began this section: *If an individual who has commodity j and wants commodity k runs into an individual who has commodity k and wants commodity j, is the result necessarily barter?* In the search-theoretic literature of money (e.g., Jones, 1976; Oh, 1989; Kiyotaki and Wright, 1993; and Trejos and Wright, 1993), the answer is unequivocal. Yes! In these models, money and barter coexist because chance encounters sometimes result in a double coincidence of wants. The scarcity of barter is related inversely to the probability of such encounters. In economies where monetary exchange is well established, however, barter is scarce because seeking a double coincidence of wants becomes increasingly inefficient with the increased use of money and specialization (Jevons, 1875, p. 3). Individuals who possibly could barter might bump into each other and never know it. Furthermore, they would not care. In highly specialized monetary economies, barter is motivated more by tax considerations or thin markets for peculiar goods (such as second-hand goods—goods become more heterogeneous as they get used—e.g., trading a used car in on the purchase of a new car) rather than by chance encounters.

Why Is Money Held?

The above analysis explains why money is the most efficient method of effecting transactions, it does not explain why money is held. The classical explanation of why money is held deals with the lack of synchronization of receipts and expenditures. Brunner and Meltzer (1971, p. 785, fn. 4) challenged this view, arguing

> It is easy to see why a "lack of synchronization" does not imply that money is used and held. Consider an economy that has neither a medium of exchange nor money. If there are no costs of acquiring information, differences in the timing of receipts and payments are adjusted by issuing verbal promises in exchange for goods and, later, delivering goods. More generally, in a barter-credit economy, commitments or promises to pay bridge the gap between receipts and payments.

Brunner and Meltzer are correct that the lack of synchronization does not explain why money is held and they also are correct in suggesting that costly information explains money's dominance in exchange. Money is held for only one reason—by its very nature the process of exchange takes time so that anything that functions as a medium of exchange must be held. This is so obvious that Brunner and Meltzer (1971, p. 804) refer to it as "trivial." Trivial though it is, this fact alone is sufficient to explain why money is held. While economists have dealt with time in a number of ways, it is convention to refer to things that are measured as a point in time as *stocks* and things that are measured through time as *flows*. Out of necessity, money is a stock. If the stock of money facilitates the flow of trade, it must exist before and after the trade takes place. In a barter economy, the goods that are traded in the interval from t to $t + 1$ must exist at time t. If, hypothetically, trade and consumption are permitted to occur simultaneously, consumption goods bartered at time t do not exist at $t + 1$. In the case of money (or sequential barter or credit),

[24] Niehans (1978) has argued that the medium-of-exchange and unit-of-account functions are wedded because money cannot ameliorate the pricing problem noted above unless money also is the unit of account. The wedding of the medium-of-exchange and unit-of-account functions enhances market efficiency also has been noted by White (1984, p. 711).

however, this is not the case. The nominal quantity of money held at $t + 1$ must be the same as that held at t. Hence, although money is continuously changing hands, it always is being held by someone—it is never consumed. This is not solely a characteristic of money; it is true of any asset.[25] Assets traded at time t exist at both t and at $t + 1$.

What then distinguishes money from any other asset? *Ipso facto* every asset is a store of wealth. To say money is a store of wealth is tautological! Being a store of wealth is not a defining characteristic of money or any other asset.[26] Money can be distinguished only from other assets by applying another criterion.[27] The important criterion for separating money from other assets is that money is an asset (or group of assets) that is generally acceptable as the means of trading goods—other assets are not. This characteristic distinguishes money from other assets and has a long tradition in classical monetary economics (e.g., Menger, 1892; Jevons, 1875; Brunner and Meltzer, 1971; and McCallum, 1983a, b, 1985). At different times and under different circumstances various assets have served as money. Nevertheless, some assets appear to have characteristics that have made them the predominant forms of money.

Money is distinguished from other assets by the function it performs; this is illustrated by a simple story from Jevons (1875):

> When Mr. Wallace was traveling in the Malay Archipelago, he seems to have suffered rather from the scarcity than the superabundance of provisions. In his most interesting account of his travels, he tells us that in some of the islands, where there was no proper currency, he could not procure supplies for dinner without special bargain, and much chaffering upon each occasion. If the vendor of fish or other coveted eatables did not meet with the sort of exchange desired, he would pass on, and Mr. Wallace and his party had to go without their dinner. It therefore became very desirable to keep on hand a supply of articles, such as knives, pieces of cloth, arrack, or sago cakes, to multiply the chance that one or other article would suit the itinerant merchant. (pp. 2-3).

The Wallace party undoubtedly kept a cache of articles that were most highly demanded by the natives. These articles are clearly stores of wealth, but the party's members would normally not have chosen to store their wealth in this form. Rather, these particular articles were held because they facilitated trade, i.e., because they reduced transaction costs. The Wallace party used these articles as a form of local currency. Whether an asset can be used as money depends solely on whether it generally is *held* to facilitate exchange.

Finally, it is worth emphasizing that money is unique among assets not solely because it facilitates the exchange of goods for consumption. Money also facilitates the exchange of other nonmoney assets as well. Individuals typically do not exchange shares of stock for acres of land even when the person who has land wants stock and vice versa. Rather, land is traded for money and the money for stock. More importantly, it is generally convenient to denominate credit contracts in units of money rather than bushels of wheat, acres of land, or other commodities or assets. Generally speaking, assets are held for the myriad of reasons that individuals accumulate wealth. In contrast, money is primarily held because of its low cost in effecting transactions. Money is distinguished from other assets in that it is the only asset that is a generally held medium of exchange. Because of this, it also is the standard of value.

ISSUES IN MONETARY THEORY

The above analysis has implications for several interesting issues in monetary theory such as the origins of fiat money, the asset demand for money, the relationship between money and credit, the buffer-stock notion of money demand, the welfare benefits of money, and the welfare costs of inflation. I will now discuss each of these issues in turn.

Fiat Money

Until now, money implicitly has been a tangible real commodity, i.e., a commodity money, or a claim to such. In modern monetary economies, however, money is typically paper currency with no intrinsic value. A question that has troubled monetary econ-

[25] Assuming, of course, that assets do not depreciate.

[26] Hicks (1989, p.42) also has made this point.

[27] There are several criteria for differentiating among assets, e.g., whether they are tangible or intangible, financial or real, liquid or illiquid, etc. These are not defining characteristics of money, however.

omists is why do people hold an asset that is valueless except in exchange? The answer suggested here is that money is the only asset that provides exchange services that other assets cannot provide.[28]

A better question is why have all monetary economies evolved into fiat money economies? The answer to this question arises naturally from the framework presented here. In the analysis presented above, I implicitly assumed that commodity money is costless to produce and/or to maintain. This is not the case. Commodity monies require that resources be used in their production and to maintain the stock. In the case of commodity monies like precious metals, most of the costs are production costs (the maintenance costs, i.e., depreciation, are fairly low). In the case of more abstract commodity monies, like checkable deposits, the costs are on going, and are related to the extent of their use.[29] Indeed, even paper currency is not costless to produce and maintain.[30]

Because the production/maintenance of money requires the use of economic resources, the welfare gains associated with the use of resource-using money are necessarily smaller than if money were costless. If money production requires resources, augmenting the stock means that resources will be drawn, at least temporarily, from other uses. The nominal money stock (e.g., tons of gold or silver) will increase as long as the marginal exchange value of the last unit produced exceeds its marginal production cost. If there are maintenance costs, (e.g., the rate of physical depreciation is positive), then resource-using money will be less welfare enhancing because of higher maintenance costs. Because the verification costs are likely to be higher the more rapidly and less predictably an asset depreciates, there is an incentive to choose as money commodities that have a low, perhaps negligible, rate of depreciation. That is, all other things being the same, the asset with the lowest maintenance cost will serve as money.

For these reasons, society has an incentive to replace high-cost commodity money with lower cost money. One step in this evolution was to replace commodity money with lower cost *representative money*. A further step is to replace representative money with even lower cost fiat money. Because fiat money requires fewer resources for production and maintenance, its use is welfare enhancing. In addition, if fiat money were to further reduce transactions costs, social welfare could be further enhanced if specialization and trade were encouraged.

Despite its advantages over resource-using money, fiat money evolved slowly over a considerable period of time. Money's evolution was undoubtedly affected by wars, other political events, and difficulties associated with regulating the supply of various commodity monies. Nevertheless, the fact that fiat money increases society's welfare relative to commodity money suggests the evolution to fiat money is the result of economic forces rather than the happenstance of a number of noneconomic events (Russell, 1991).

Governments issue fiat money because private fiat money issuers have an incentive to issue money as long as the marginal value of the last nominal unit issued is greater than its production cost. Consequently, it would be difficult for private money issuers to make a credible commitment not to over issue fiat money, so as to make it "worthless" (e.g., Ritter, 1995).[31] Moreover, only the government can credibly commit to distribute the seigniorage revenue from money's creation. Seigniorage arises when the exchange value of the money issued exceeds money's production cost.

The existence of a money whose value in exchange exceeds its production cost has given rise to the notion that society's wealth exceeds the stock of tangible assets by the real value of fiat money held. As Tobin (1965, p. 676) put it:

> ...as viewed by the inhabitants of the nation individually, wealth exceeds the tangible capital stock by the size of what we might call the fiduciary issue. This is an illusion, but only one of the many fallacies of composition which are basic to any economy or any society.

Many economists accept Tobin's claim that the "wealth" associated with the real stock of fiat money is illusionary, but this proposition is erroneous. Tobin's error comes from viewing a monetary econ-

[28] See Tobin (1992, p. 774) for the traditional answer to this question.

[29] Thornton (1983) shows that the relevant issue for determining whether "inside money" is part of a society's stock of net wealth is whether there are resource costs involved in its production and maintenance.

[30] The United States is issuing another in a series of dollar coins. The purpose of these coins is to reduce the cost of maintaining the stock of currency, since coins depreciate less rapidly than paper money.

[31] Goodhart (1998) suggests that gold's role as a medium of exchange was greatly enhanced by government's use of gold to pay tributes or tariffs to avoid feuds.

omy as simply a barter economy with money. In so doing, he fails to recognize the private and public benefits that accrue from money's use. The benefits from the use of money naturally accrue to fiat money when society shifts from using a more costly commodity money to a less costly (or, ideally, completely costless) fiat money. Hypothetically, if resource-using money were replaced unit-for-unit with fiat money, the real value of the stock of fiat money would reflect the welfare benefits associated with the previously held stock of commodity money. Hence, the benefits of commodity money are embodied fully in the same real quantity of fiat money.[32]

Furthermore, the fact that all of the benefits from the previous stock of money would be obtained at lower cost guarantees that welfare is enhanced by the switch, even if there is no further reduction in the marginal transaction cost and, consequently, no further increase in trade and specialization.

Because of the existence of positive externalities associated with money's use, it is inappropriate to equate the welfare benefits of money with the real value of the money stock, i.e., M/P, as is frequently done. Nevertheless, it is clear that including the real value of the "fiduciary issue" as part of society's net wealth is not an illusion. Some time ago, Clower (1967) pointed out the dangers of treating monetary economies as if they were analytically equivalent to barter economies. In a similar vein, Coase (1960) argues that when a comparison of economies with alternative social arrangements is made, it is essential to consider the total effect. The true benefits of fiat money only can be obtained by comparing a fiat money economy with a commodity money economy or with barter or credit economies. The conclusion that the wealth associated with the real quantity of fiat money is illusionary emerges from a naive comparison of a fiat money economy with an economy where all of a sudden no one uses or holds money but nothing else changes.

Money and Credit as Media of Exchange

The world is dominated by monetary economies; however, this does not mean that transactions are not carried out using barter or credit. In monetary economies, all three methods of effecting exchange are used. Indeed, money may not be used to initiate most transactions. For example, when one considers every extension of trade credit or the transfer of goods by credit card, it is arguably the case that more transactions are carried out initially with credit than with cash or checks. Given the large and increasing use of credit in effecting transactions, how can one reasonably argue that this is a *monetary economy* and not a *credit economy*?

Let us begin this discussion by trying to answer the intriguing question: Could there be a pure credit economy with no medium of exchange? A pure credit economy may have been what Brunner and Meltzer (1971) had in mind when they argued that the problem of synchronizing payments and receipts could be achieved by making verbal promises. To see what such a world might look like, I will assume that not only is there perfect information, but that all individuals' promises are fully credible, i.e., no person makes a promise that cannot be kept.[33] In such a world, individual *A* could give individual *B* commodity *j* in exchange for a promise to receive commodity *j* or some other commodity at a later date. This world would be very complicated. For example, assume that individual *A* sells his labor services to Firm *F* for the promise from *F* to pay a certain quantity of commodity *j* at week's end. Individual *A* then buys the goods that he needs by promising to deliver *j* or some other commodity at some point in the future or by transferring part of Firm *F*'s promise to deliver commodity *j*. Of course, it is not necessary that these promises change hands *per se*, it could be that some centralized accountant keeps track of all promises made to and from all parties, or everyone could simply have a perfect memory.[34]

If promises were denominated in all possible commodities, quantities, and future dates, the problem of calculating the prices in this economy would be extremely difficult. The pricing problem could be significantly reduced (and the accounting simplified) if individuals agree to denominate all credit

[32] This point was initially made by Johnson (1969), p. 38, who recognized that there were utility or output gains associated with the use of money.

[33] There are a number of similarities between the no-money world I am about to describe and that described by Fama (1980, 1983). Others who have suggested that transactions could be carried out without the use of money are Black (1970), and Greenfield and Yeager (1983). See McCallum (1985) and White (1984) for analyses of these models.

[34] Kocherlakota (1998) suggests that fiat money is "merely a physical way of maintaining this balance sheet." Hence, he suggests that money is merely memory. He even suggests that his approach "represents an advance over the usual justifications for the existence of money: Money is a store of value, money is a medium of exchange and/or money is a unit account...After all, money does not allow society to transfer resources over time. Money does not reduce the cost of transferring resources from one person to another."

contracts in the same commodity. This would give rise to this commodity being a medium of exchange, however. For example, if all credits are denominated in *m*, it must be the case that a credit instrument worth *z* units of commodity *m* today must trade for *z* units of *m* itself. This means that individuals with *m* could simply trade it for the commodities they desire just as well as they could trade credit instruments denominated in *m*. If a credit instrument denominated in *m* facilitates trade, then so too must commodity *m*—*m* would be money.[35]

It could be that *m* is bulky, like a barrel of oil, so that it could not circulate hand-to-hand.[36] The promises themselves would be inconvenient, however, because longer-term contracts would have to be discounted relative to shorter-term contracts. This difficulty could be overcome by issuing non-interest-bearing sight drafts denominated in the common unit of account, i.e., currency.[37] In this case, a credit economy would give rise to money.

It could be, however, that the commodity is completely abstract, like a quark. Hence, we would have a pure credit, nonmonetary, exchange economy where all credit contracts are denominated in a unit of account, whose only function is to determine the price level (Fama, 1983). People, however, only would accept promises denominated in something abstract or something that they did care to hold if they were certain that they would be able to exchange these promises for the commodities they desire. Hence, that would demand that credit contracts be denominated in things that they value or are certain that they would be easily converted into other commodities.

Note the similarity between the world I have just described and the one that exists today. Our money is called the *dollar*. Congress adopted the *dollar* (and the decimal system) as our unit of currency in 1785. Alexander Hamilton's coinage recommendation establishing the U.S. dollar as 270 grains, 11/12 fine of gold or 416 grains, 0.89242 fine of silver was not adopted until April 1792.[38] Because of the inconvenience of carrying gold or silver, sight drafts were issued in convenient denominations. These claims on the U.S. stocks of gold and silver circulated in lieu of the commodities themselves. Over the years the dollar has been redefined. U.S. currency now is just a claim on the same quantity of U.S. currency. That is, we now have a pure paper currency standard. People are willing to hold intrinsically useless pieces of paper and claims that are denominated in intrinsically useless pieces of paper because they are certain that other individuals will accept the same. Collectively, the people agree to maintain the paper's value by limiting its issuance and to share the seigniorage.[39]

The above analysis also reinforces why it is efficient to have credit contracts denominated in the same commodity, and better still if this commodity is money. Jevons (1875) termed this the *standard of value* of money. The point to emphasize is that money facilitates the use of credit just as it facilitates the trade of consumable commodities and tangible assets, (e.g., savings deposits are exchanged for dollars that are used to purchase bonds). Consequently, while credit figures prominently in many transactions, the analysis presented above makes it clear why credit almost never is used sequentially for other transactions and why the adoption of a commodity medium of exchange has tended to precede credit arrangements, and not the other way around.[40]

Can credit instruments function as money? The answer is yes. They can and they have. Checkable deposits (or electronic transfers of funds) are the liabilities of the entities who hold the balances.[41] As such, they are promises to pay dollars upon demand. Such balances are included in measures of transactions money not only because they facilitate exchange, but also because financial institutions are committed to exchanging these deposits for cash immediately and at a fixed one-to-one ratio. This is what Pesek and Saving (1967) termed the

[35]Hoover (1988) has made this point in a similar fashion to argue against Fama's (1980) "new monetary economics."

[36]Fama (1980, 1983) eliminated the possibility that what he termed the "numeraire—unit of account" would circulate as money by assuming that it was a "barrel of oil." Note that this was not a necessary consequence of his model, but assumed. Hence, Fama did not establish that there would be no commodity that would circulate as a medium of exchange, rather he assumed it.

[37]In addition, there may be a problem with the denominations of such contracts. Indeed, Russell (1991) notes that both of these problems were drawbacks to bills of exchange circulating as currency during the seventeenth and eighteenth centuries in England.

[38]The mint began to coin silver in October 1794 and gold in July 1795, but a mistake by the first mint director resulted in coins of 9/10 fine. See Studenski and Krooss (1952) for more details.

[39]See McCallum (1985) for other ways of achieving price-level determinacy under a currency standard.

[40]Bagehot (1873) makes this point with respect to the origins of bank credit.

[41]See Goodfriend (1991) for a good discussion of the evolution of bank money.

instant repurchase clause. As long as the commitment is fully credible, such deposits and currency substitute perfectly. In this case, it is sensible to add such commitments to the stock of cash and call the sum the stock of money. Indeed, this is what is done.

Finally, it is worth noting that the increased use and availability of credit might mitigate the effects of disruptions to the supply of money, at least in the short run. In economies where the credit market is not well developed, a negative shock to the money supply may have a more immediate effect on output and/or prices than in an economy where individuals and businesses can not only readily borrow against their future income but can make transactions without having money immediately available. In addition, as more transactions are initiated with the use of credit, the stock of money necessary to support a given level of commodity transactions could diminish, i.e., the velocity of money could rise. It should be remembered, however, that financial transactions also require the use of money. In any event, it is reasonable to speculate that the relationship between money and output and money and prices is likely to change as financial markets develop and mature.

The Asset Demand for Money

The asset demand for money has been associated with two literatures. The first deals with demand for money as an asset and focuses on the interest elasticity of the demand for money. The second focuses on whether money should be defined to include non-medium-of-exchange assets. Money's essential function is to facilitate transactions. Hence, while it is appropriate to consider the effect of close substitutes for money on its demand, it is inappropriate to define money to include such non-medium-of-exchange assets.[42]

The asset demand for money focused attention on holding money for asset purposes, just like you hold any other asset. I will argue, however, that the asset demand for money is inconsequential.[43] The asset demand for money has its origins with Lavington (1968), but was most influentially advanced by Keynes. If money was held primarily as an asset, its demand should be quite sensitive to changes in interest rates, because the nominal return to holding money is zero.

If money is primarily a medium of exchange, however, the interest elasticity of money demand might be quite low. To see why, I note that Brunner and Meltzer (1971) begin their seminal work on money by noting that money remains in circulation even during periods of high and accelerating inflation. They argue that this fact "calls into question the relevance of treating money as an asset that provides little or no return."[44] The analysis of why individuals continue to use money during periods of high and accelerating inflation presented here is complementary with theirs. Money continues to function as a medium of exchange even under conditions of severe or hyperinflation because it enjoys a significant cost advantage over both barter and credit as a medium of exchange. Indeed, this advantage is likely to be so large that it would take an extreme increase in the holding cost to induce individuals to shift to the widespread use of either barter or credit to facilitate exchange. Moreover, the cost advantage of money increases as economies become increasingly specialized and dependent on exchange. The advantage also increases as payments practices become increasingly institutionalized.[45]

The point is that a large discontinuity exists between money and the next best alternative for exchange. Economists normally think of continuous functions where small changes induce individuals to switch from one alternative to another. No such continuum of media of exchange exists, however. Money so dominates barter and credit as a medium of exchange that it continues to serve as a medium of exchange despite very large increases in the cost of holding it. Jevons (1875, p. 6) was aware of this, stating:

> ...even if the medium of exchange varied considerably in value, people would go on making their payments in terms of it, as if there had been no variation, some gaining at the expense of others.

[42] See Mason (1976) for an excellent critique of this approach to defining money.

[43] McCallum and Goodfriend (1987) also have suggested the asset demand for money should be relatively inconsequential, stating that money "will also serve as a store of value, of course, but may be of minor importance to the economy in that capacity." They do not elaborate on why this should be so, however.

[44] Brunner and Meltzer (1971, p. 784).

[45] Wallace (1983) has emphasized one of these institutional features; namely, the legal restriction that currency is *legal tender*. Overlapping generations models focus on the store of value function of money, i.e., money's function as an asset (e.g., McCallum, 1983; and McCallum and Goodfriend, 1987).

The discontinuity between money and other means of exchange suggests that the demand for the medium of exchange may be rather insensitive to changes in its holding cost.

The key observation is the degree of substitutability between money and other assets is that the substitution is *unidirectional*: While money is an asset (or group of assets) that provides a particular function that other assets do not provide, at times, money may be held for the same reasons that other assets are held—the asset that normally serves as money also is now being held as a store of wealth. It is never the case, however, that other assets are held for the reason that money is primarily held. This means that when rapid and accelerating inflation significantly increases the cost of holding money it will not be a simple matter for other assets to substitute for it, i.e., become money. The most individuals can do is to economize on their money holdings along the lines suggested by Baumol (1952) and Tobin (1956). Since other assets dominate money in their ability to transfer wealth through time, however, individuals have a strong incentive to economize on their holdings of money for transactions purposes even when the returns to other stores of wealth are low.

Nevertheless, it is possible to envision circumstances where the return on real assets is so low that some individuals choose to hold money for the same reason they normally hold other assets. Indeed, classical economists, including Keynes, were concerned about the consequences of hoarding money. Given the observed stickiness of prices, they argued that hoarding money would have significant consequences for the real economy.

Hoarding money by individuals seems more likely, however, in economies with relatively poorly developed financial markets.[46] If few alternatives to holding wealth are readily available, more individuals may opt to hoard money, especially during times of economic or financial uncertainty. The more sophisticated and well developed the financial system becomes, however, the less likely it is that individuals will choose to hold money as an asset, even when nominal interest rates are extremely low.[47]

Keynes' notion of the asset demand for money focused the attention of monetary economists on the interest sensitivity of money demand. The interest sensitivity of money demand has been extensively investigated, with a wide array of results (e.g., Goldfeld and Sichel, 1990; and Laidler, 1993).

The amount of money held for transactions purposes depends on the planned volume of transactions. This, in turn, depends on the timing of receipts and payments, which are affected by the degree of specialization and the structure of the markets, as well as the size, extent, and activity in credit markets, etc. Changes in the opportunity cost of holding money will induce individuals to economize on their holdings of money balances, but the degree to which they do this depends on the size of the gain relative to the marginal cost of the economizing activity. Given that money holdings are typically a small part of an individuals' wealth and that individuals have a strong incentive to minimize their holding of money at any nonzero nominal interest rate, it would not be surprising to find a relatively low interest responsiveness of money demand. Indeed, empirical investigations of currency demand (e.g., Hess, 1971; and Dotsey, 1988), which has a zero nominal return and is held primarily for transactions purposes, suggest that the interest elasticity of currency demand is zero.[48] Other mediums of exchange that pay an implicit or explicit interest may be held, in part, for the same reasons individuals hold other assets, so that the demand for them is likely to be more sensitive to changes in their relative holding cost.

The Buffer-Stock Notion of the Demand for Money

The idea that there is no close substitute for money as a medium of exchange is complementary with the buffer-stock notion of money demand. In the buffer-stock theory (Laidler, 1984, 1987), holdings of real balances substitute for costly information and uncertainty. Individuals absorb shocks to their real money holdings due to a shock to their nominal money balances. Over time, nominal money holdings are adjusted to a level more consistent with individuals' demand for real money balances, given the level of nominal interest rates, the level

[46] Unfortunately, Keynes attempted to rationalize hoarding at a time when financial markets were well developed. Hence, it was difficult to explain why individuals held money when there were assets that had all of the same risk characteristics of money but yielded a positive rate of return (Barro and Fischer, 1976).

[47] It is usually assumed that zero is a lower bound for the nominal interest rate because individuals could simply hold money that bears a zero nominal return. This analysis too ignores the costs of acquiring and storing money. See Thornton (1999).

[48] Furthermore, most studies find a remarkably low substitutability between currency and transactions deposits, suggesting that these alternative media of exchange are held for quite different reasons.

and pattern of current income and expenditures and expectations of future nominal interest rates, income and expenditures, etc. The buffer-stock notion implies that individuals will not change their holdings of real money balances immediately when nominal interest rates, real income, or prices change.

Because other assets cannot perform money's function as a medium of exchange, I speculate that individuals respond more quickly to reductions in the real money balances due to negative nominal money shocks (or positive price-level shocks) than they do to increases caused by positive money shocks (or negative price-level shocks). For example, when there is a positive aggregate nominal money shock, individuals may hold these balances temporarily rather than spending them for goods and services or purchasing other assets. If this were to happen, there would be no immediate adjustment of output, employment, prices, or interest rates. On the other hand, since individuals cannot substitute for money, negative aggregate shocks may affect economic behavior directly and more quickly.

The Welfare Benefits of Money

The usual approach to assessing the welfare benefits of money is to assume that money is like other assets; for instance, shoes or cars. In the case of these assets, the benefits accrue only to the consumer so the welfare gains can be obtained by simply summing up the so-called *Harberger triangles*. It is well known that this approach fails when there are significant social externalities. Since I have argued that there are significant social benefits from money—because of the role it plays in expanding the size and extent of the markets for goods and credit, and the degree of specialization—this approach cannot possibly work. Indeed, it seems reasonable to speculate that the social benefits of money could eclipse its private benefits.

Unlike many innovations, it is virtually impossible to internalize the benefits from using money. This further enhances the idea that there is significant social benefit to money. Indeed, once the usefulness of money is recognized, the one who recognizes it has an incentive to share the insight with others, as my *parable of the trader* illustrates.

The Parable of the Trader

There was a producer who once every period loaded some of his produce on a wagon and went to a destination where he and other producers would meet to trade their wares. One day, the producer noticed that there was one good, *g*, that nearly everyone wanted and would exchange goods for *g*. Realizing that he can buy virtually any good he desired using *g*, he offers to take *g* for the goods he was trying to trade. Initially he does this only when the double coincidence of wants necessary for barter is lacking. He soon discovers, however, that trading in *g* is much faster and easier than searching out barter opportunities, so he stops seeking barter opportunities and his barter transactions become increasingly infrequent. By trading his wares for *g*, and *g* for the goods that he desires, this producer discovers that he can accomplish the desired trading in a fraction of the time that he had previously spent.

Now he could attempt to internalize the gain from his private knowledge (no one else has made this observation yet) by offering to tell others how they could save trading time for a fee. He realizes, however, that no one would pay for this information because all they have to do is observe him and they, too, would know the secret. More important, he realizes that he could further shorten his trading time if the others behaved as he. Hence, rather than keeping this information private and attempting to internalize the benefit from his superior information, the trader has an incentive to make the information public. In so doing, however, not only does he gain by shortening the transactions time, but others do as well.

As Laidler (1990, p. 48) puts it, "one agent's cash balances produce services not just for that agent then but for all other agents with whom his market activities bring him into contact." The use of money that facilitates the trade of one agent facilitates the trade of all agents. In addition, the reduction in individuals' exchange cost associated with money's use causes markets to flourish. Increased trade promotes greater specialization, greater dependence on trade, and a greater need for and use of money, and so on, and so forth.

The synergy among trade, money, and specialization makes isolating the welfare benefits of

money extremely difficult, if not futile. The welfare benefits of money can be ascertained only by comparing monetary economies with economies that have alternative arrangements for exchange, i.e., only by comparing the total welfare of a monetary economy with that of a nonmonetary economy.[49]

The Welfare Costs of Inflation

The main implication of the discontinuity between money and barter or credit as a medium of exchange is that money will continue to be used even at very high rates of inflation. This implies that the welfare costs of inflation, which are associated with the reduced reliance on money as a medium of exchange, may be relatively small. This is particularly likely at relatively modest rates of inflation. Hence, it is not surprising that estimates suggest that the cost of inflation is large only at relatively high inflation rates (e.g., Bruno and Easterly, 1996).

Furthermore, not only is it inappropriate to estimate the welfare gains from the use of money by adding up Harberger triangles, it is equally inappropriate to measure the welfare costs of inflation this way, as is frequently done.[50] Since money will continue to circulate as a medium of exchange and since the ability to economize further on money holdings is likely to be small, so, too, is the cost of inflation from holding money balances. This is important because many discussions about inflation assume that its principal cost is the private *shoe leather* cost associated with economizing on the use of money as a medium of exchange. If the externalities associated with money are important and significant, such analyses understate the welfare costs of inflation, perhaps significantly.

Most economists would argue that if an economy were just starting, the optimal rate of inflation would be zero. Nevertheless, many argue that once inflation is underway, society is better off tolerating some inflation rather than to suffer the output loss they believe would be associated with reducing inflation to zero. This idea is called Howitt's (1990) Rule. The effects of inflation on the institutional arrangements of trade are likely to be extremely important, however, and these costs are missed completely by estimates that ignore the externalities associated with money's roles as a medium of exchange and a standard of value. Consequently, Tobin's often cited dictum that "it takes a heap of Harberger triangles to fill an Okun gap," which underlies such analyses, is simply irrelevant if there are significant social costs of inflation.[51]

The third consequence of the discontinuity between money and other methods of exchange is that it may be inflation uncertainty, rather than inflation *per se*, that produces the most significant welfare cost. Here it is important to distinguish between the *medium of exchange* and *standard of value* functions of money. An important benefit of money is that it reduces shopping costs—gathering information about relative prices, planning, budgeting, etc. Uncertainty interferes with the shopping function by distorting price signals that enhance market efficiency. Price-level uncertainty makes distinguishing between absolute and relative prices and between permanent or transitory changes in the price level difficult. Distortions to the pricing mechanism affect the efficiency of markets that affect investment (e.g., DeLong and Summers, 1991; and Barro, 1995), financial markets, and relative input prices (e.g., Easterly, 1993). Inflation also reduces efficiency by encouraging the development of alternative market structures that would not exist in a world with a stable price level. Because uncertainty about the future level of prices increases with the average rate of inflation, these costs are likely to be small at relatively low rates of inflation but increase with the rate of inflation.

It could be, however, that the most deleterious effects of inflation on economic welfare may come from the effect of inflation on the efficient function of the credit market. Both the rate of inflation and

[49] This may have implications for how money is modeled. For example, it is frequently the case that money is modeled in the context of one good economy where exchange is implied but not explicitly modeled. Given the possibility that there are large externalities associated with money, this practice may not be useful for some issues. It also may have implications for other models. For example, Lucas (1980, p. 145) states, "When we apply theories of barter economies to problems in, say, public finance or labor economics, it is not our intent to obtain results applicable only to primitive or prehistoric societies. We apply this body of theory to money-using economies such as our own because we believe that for many problems the fact that money is used in attaining equilibrium can be abstracted from, or that the theoretical barter economy is a tractable, idealized model which approximates well (is well-approximated by) the actual monetary economy. If this practice is sound, then we want monetary theories which rationalize it or at least do not radically conflict with it."

[50] For example, Bailey (1956), Friedman (1969) and Lucas (1994).

[51] For a critique of some other limitations of Howitt's Rule, see Thornton (1996). Also see Marty and Thornton (1995) for a discussion of some other arguments for the desirability of moderate inflation.

inflation uncertainty are detrimental to denominating credit contracts in terms of fixed units of money. Consequently, while high, accelerating, and especially uncertain inflation may have a relatively small effect on money's medium-of-exchange function, they may have a significant effect on financial markets. It is not easy to replace money as the standard of value. Recently, credit contracts have been denominated in variable units of money, so that the value of the contract varies with a measure of the actual inflation experience during periods of inflation uncertainty. For reasons that are not well understood, however, this practice has been relatively limited, especially at relatively moderate inflation rates. Long-term debt markets tend to dry up during periods of rapid inflation and, as a consequence, the rate of capital formation slows. While far from definitive, the evidence suggests that the covariance between inflation and the rate of economic growth is negative (e.g., Bruno and Easterly, 1996).

While inflation potentially has a significant effect on the rate of economic growth, its potential to affect the level of output may be modest. To the extent that high and accelerating inflation reduces the reliance on money as either a medium of exchange or a standard of value, resources are drawn from one use to another. The result is that the level of measured output may change relatively little between high and low inflation states, but the distribution of output may be significantly different and the level of economic welfare may be significantly lower in higher inflation environments. This may account for the fact that economists have not found a statistically significant relationship between the rate of inflation and the level of output, at least for relatively moderate rates of inflation.

SUMMARY AND CONCLUSION

I have argued that money is a social arrangement resulting from a complicated evolutionary process. Money exists because it facilitates exchange by reducing the cost of trade. Seen in this point of view, money is but one of several institutional arrangements designed to reduce the costs of exchange. By reducing the cost of exchange, money reduces the reservation relative price where trade is mutually advantageous thereby encouraging more trade and greater specialization. Because of their strategic complementary, it is not surprising that money, trade, and specialization have tended to evolve simultaneously.

I argue that there are only three methods of effecting trade: simple and sequential barter, simple and sequential credit, and money. I then explain why the information and shopping costs of sequential barter and/or sequential credit are likely to be high relative to those of money. It is not surprising that the world is populated with monetary economies and not barter or credit economies.

I also have argued that money has a significant cost advantage relative to simple barter and credit and this advantage helps explain why the same good has served most often as both the medium of exchange and the unit of account, and why the development and widespread use of money tends to make simple barter scarce.

The use of money promotes specialization and trade by reducing exchange costs. The reduction in exchange costs associated with money cannot benefit one individual without benefiting others. Indeed, it is virtually impossible to internalize the benefits from money. Consequently, there are significant externalities associated with the use of money. Money is a social arrangement whose benefits can be calculated correctly only by comparing monetary economies with barter or credit economies. I speculate that the social gains from the use of money are likely to be large relative to the private opportunity cost of holding it. Furthermore, these benefits extend to nonresource-using fiat money. Indeed, the fact that nonresource-using money frees resources for production and/or leisure necessarily implies that, other things being the same, the transition from commodity to fiat money is welfare enhancing.

I argue that money enjoys an enormous cost advantage over barter or credit as a medium of exchange. Because of this, inflation is not likely to result in a large-scale substitution away from money as a medium exchange. Hence, money continues to circulate as a medium of exchange even during periods of hyperinflation. Significant costs of inflation could be associated with the effects of inflation uncertainty on the efficiency of the goods, labor, and financial markets, most especially the efficiency of the credit market because of the deterioration of money's function as a standard of value.

The fact that there are significant externalities associated with the use of money and that inflation increases the costs of using money gives rise to the possibility that the welfare costs of inflation are significant. Because money dominates barter and credit as a medium of exchange, the welfare costs of inflation due to a reduction in money's role as a medium

of exchange are likely to be small, relative to those associated with its function as a standard of value.

REFERENCES:

Alchian, Armen A. "Why Money?" *Journal of Money, Credit and Banking* (February 1977), pp. 133-40.

Archibald, Glen C., and Robert G. Libsey. "Monetary and Value Theory: A Critique of Lange and Patinkin," *The Review of Economic Studies* (October 1958), pp. 1-22.

Bagehot, Walter. *Lombard Street*, John Murray, London, 1873.

Bailey, Martin J. "The Welfare Cost of Inflationary Finance," *Journal of Political Economy* (April 1956), pp. 93-110.

Barro, Robert J. "Inflation and Economic Growth," Bank of England *Quarterly Bulletin* (May 1995), pp. 166-76.

_____ and Stanley Fischer. "Recent Developments in Monetary Theory," *Journal of Monetary Economics* (April 1976), pp. 133-67.

Baumol, William J. "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* (November 1952), pp. 545-56.

Benhabib, J., and C. Bull. "The Optimal Quantity of Money: A Formal Treatment," *International Economic Review* (February 1983), pp. 101-11.

Black, Fischer. "Banking and Interest Rate in a World Without Money: The Effects of Uncontrolled Banking," *Journal of Bank Research* (Autumn 1970), pp. 9-20.

Braun R. Anton. "Another Attempt to Quantify the Benefits of Reducing Inflation," Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall 1994), pp. 17-25.

Brunner, Carl, and Allan H. Meltzer. "The Uses of Money: Money in the Theory of an Exchange Economy," *American Economic Review* (December 1971), pp. 784-805.

Bruno, Michael, and William Easterly. "Inflation and Growth: In Search of a Stable Relationship," this *Review* (May/June 1996), pp. 139-46.

Bullard, James. "Testing Long-Run Neutrality Propositions: Lessons from the Recent Research," this *Review* (November/December 1999) pp. 57-77.

Clower, Robert W. "A Reconsideration of the Microfoundations of Monetary Theory," *Western Economic Journal* (December 1967), pp. 1-8.

Coase, Ronald H. "The Problem of Social Cost," *The Journal of Law & Economics* (October 1960), pp. 1-44.

Corbae, Dean, and Joseph A. Ritter. "Money and Search with Enduring Relationships," unpublished manuscript, 1998.

DeLong, J.B., and Lawrence H. Summers. "Equipment Investment and Economic Growth," *The Quarterly Journal of Economics* (May 1991), pp. 445-502.

Dornbusch, R., and Jacob Frenkel. "Inflation and Growth: Alternative Approaches," *Journal of Money, Credit and Banking* (February 1973), pp. 141-56.

Dotsey, Michael. "The Demand for Currency in the United States," *Journal of Money, Credit and Banking* (February 1988), pp. 22-40.

Duffy, John. "Monetary Theory in the Laboratory," this *Review* (September/October 1998), pp. 9-26.

Easterly, William. "How Much Do Distortions Affect Growth?" *Journal of Monetary Economics* (November 1993), pp. 187-212.

Fama, Eugene F. "Banking in the Theory of Finance," *Journal of Monetary Economics* (January 1980), pp. 39-57.

_____. "Financial Intermediation and Price Level Control," *Journal of Monetary Economics* (July 1983), pp. 7-28.

Fischer, Stanley. "Monetary Rules and Commodity Money Schemes under Uncertainty," *Journal of Monetary Economics* (January 1986), pp. 21-35.

Friedman, Milton. "The Optimum Quantity of Money," in *The Optimum Quantity of Money and Other Essays*, Aldine Publishing Co., 1969.

Goldfeld, Stephen M., and Daniel E. Sichel. "The Demand for Money," in the *Handbook of Monetary Economics*, Volume I, Benjamin M. Friedman and Frank H Hahn, eds., Elsevier Science Publishers, 1990, pp. 299-356.

Goodfriend, Marvin. "Money, Credit, Banking and Payments System Policy," Federal Reserve Bank of Richmond *Economic Review* (January/February 1991), pp. 7-23.

Goodhart, Charles A. E. "The Two Concepts of Money: Implications for the Analysis of Optimal Currency Areas," *European Journal of Political Economy* (1998), pp. 407-32.

Greenfield, Robert L., and Leland B. Yeager. "A Laissez Faire Approach to Monetary Stability," *Journal of Money, Credit and Banking* (August 1983), pp. 302-15.

Hess, Alan C. "An Explanation of Short-Run Fluctuations in the Ratio of Currency to Demand Deposits," *Journal of Money, Credit and Banking* (August 1971), pp. 666-79.

Hicks, John. *A Market Theory of Money*, Clarendon Press, Oxford, 1989.

Hirshleifer, J. "Exchange Theory: The Missing Chapter," *Western Economic Journal* (June 1973), pp. 129-46.

Hoover, Kevin D. "Money, Prices and Finance in the New Monetary Economics," *Oxford Economic Papers* (March 1988), pp. 150-67.

_____. "Some Suggestions for Complicating the Theory of Money," unpublished manuscript, 1995.

Howitt, Peter. "Zero Inflation as a Long-Term Target for Monetary Policy," in *Zero Inflation: The Goal of Price Stability*, Richard G. Lipsey, ed., C. D. Howe Institute, 1990.

Jevons, William Stanley. *Money and the Mechanism of Exchange*, Twentieth Century Press, 1875.

Johnson, Harry G. "Inside Money, Outside Money, Income, Wealth, and Welfare in Monetary Theory," *Journal of Money, Credit and Banking* (February 1969), pp. 30-45.

Johri, Alok. "On the Real Effects of Fiat Money in a Search Model," unpublished manuscript, Boston University, 1994.

Jones, R. "The Origin and Development of Media of Exchange," *Journal of Political Economy* (August 1976), pp. 757-75.

Kiyotaki, Nobuhiro, and Randall Wright. "A Search-Theoretic Approach to Monetary Economics," *American Economic Review* (March 1993), pp. 63-77.

_____. "A Contribution to the Pure Theory of Money," *Journal of Economic Theory* (April 1991), pp. 215-35.

_____. "On Money as a Medium of Exchange," *Journal of Political Economy* (August 1989), pp. 927-54.

Kocherlakota, Narayanna R. "Money Is Memory," *Journal of Economic Theory* (August 1998), pp. 232-51.

Laidler, David. "The Buffer-Stock Notion in Monetary Economics," *Conference Proceeding, Supplement to the Economic Journal* (March 1984), pp. 326-34.

_____. "Buffer-Stock Money and the Transmission Mechanism," Federal Reserve Bank of Atlanta *Economic Review* (March/April 1987), pp. 11-23.

_____. *Taking Money Seriously*, Humel Hempstead: Philip Allan; Cambridge, Mass.: MIT press, 1990, pp. 1-23.

_____. *The Demand for Money*, Harper Collins College Publishers, 1993.

Lavington, F. *The English Capital Market*, Augustus M. Kelley, New York, 1968.

Lucas, Robert E. Jr. "On the Welfare Cost of Inflation," unpublished manuscript, University of Chicago, 1994.

_____. "Equilibrium in a Pure Currency Economy," in *Models of Monetary Economies*, John H. Kareken and Neil Wallace, eds., Federal Reserve Bank of Minneapolis, 1980, pp. 131-45.

Marty, Alvin L., and Daniel L. Thornton. "Is There a Case for 'Moderate' Inflation?" this *Review* (July/August 1995), pp. 27-37.

Mason, Will E. "The Empirical Definition of Money: A Critique," *Economic Inquiry* (December 1976), pp. 525-38.

McCallum, Bennett T. "The Role of Overlapping Generations Models in Monetary Economics," *Carnegie-Rochester Conference Series on Public Policy* (Spring 1983a), pp. 9-44.

_____. "A Model of Commodity Money: Comments," *Journal of Monetary Economics* (July 1983b), pp. 189-96.

_____. "Bank Deregulation, Accounting Systems of Exchange, and the Unit of Account: A Critical Review," *Carnegie-Rochester Conference Series on Public Policy* (Autumn 1985), pp. 13-45.

McCallum, Bennett T., and Marvin S. Goodfriend. "Demand for Money: Theoretical Studies," *New Pargrave: A Dictionary of Economics*, John Eatwell, Murray Mulgate and Peter Newman, eds., MacMillan Publishing, 1987, pp. 775-80.

Meltzer, Allan H. "Information, Sticky Prices, and Macroeconomic Foundations," this *Review* (May/June 1995), pp. 101-18.

Menger, Carl. "On the Origin of Money," *The Economic Journal* (June 1892).

Mundell, Robert. *Monetary Theory, Inflation, Interest, and Growth in the World Economy*, Goodyear Publishing Co., 1971.

Niehans, Jurg. "Money and Barter in General Equilibrium with Transaction Costs," *American Economic Review* (December 1971), pp. 773-83.

_____. *The Theory of Money*, Johns Hopkins University Press, 1978.

Oh, Seonghwan. "A Theory of a General Acceptable Medium of Exchange and Barter," *Journal of Monetary Economics* (January 1989), pp. 101-19.

Ostroy, Joseph M. "The Informational Efficiency of Monetary Exchange," *American Economic Review* (September 1973), pp. 597-610.

Patinkin, Don. *Money, Interest, and Prices: An Integration of Monetary and Value Theory*, Harper and Row, second edition, 1965.

Pesek, Boris P., and Thomas R. Saving. *Money, Wealth and Economic Activity*, MacMillan Company, 1967.

Ritter, Joseph A. "The Transition from Barter to Fiat Money," *American Economic Review* (March 1995), pp. 134-49.

Russell, Steven. "The U.S. Currency System: A Historical Perspective," this *Review* (September/October 1991), pp. 34-61.

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*, Edwin Cannan, ed. Random House, 1937.

Studenski, Paul, and Herman E. Krooss. *Financial History of the United States*, McGraw-Hill Book Company, 1952.

Thornton, Daniel L. "Bank Money, Net Wealth and the Real Balance Effect," *Journal of Macroeconomics* (Winter 1983), pp. 105-17.

_____. "The Costs and Benefits of Price Stability: An Assessment of Howitt's Rule," this *Review* (March/April 1996), pp. 23-38.

_____. "Nominal Interest Rates: Less Than Zero?" *Monetary Trends* (January 1999) p. 1.

Tobin, James. "Money," in *The New Palgrave Dictionary of Money and Finance*, by Peter Newman, Murray Milgate and John Eatwell, eds., MacMillian Press, 1992, pp. 770-79.

_____. "Money and Economic Growth," *Econometrica* (October 1965), pp. 671-84.

_____. "The Interest-Elasticity of the Transactions Demand for Cash," *Review of Economics and Statistics* (August 1956), pp. 241-47.

Trejos, Alberto, and Randall Wright. "Search, Bargaining, Money, and Prices: Recent Results and Policy Implications," *Journal of Money, Credit and Banking* (August 1993, Part 2), pp. 558-76.

Wallace, Neil. "A Legal Restrictions Theory of the Demand for 'Money' and the Role of Monetary Policy," Federal Reserve Bank of Minneapolis *Quarterly Review* (Winter 1983), pp. 1-7.

White, Lawrence H. "Competitive Payments Systems and the Unit of Account," *American Economic Review* (September 1984), pp. 699-712.

Appendix

The Effect of Exchange Costs of the Exchange Ratio Necessary for Trade

Robert D. Dittmar and Daniel L. Thornton

The text argues that costly trade reduces the feasible range of exchange ratios where trade is mutually advantageous. The question that arises is by how much must the exchange ratio change to compensate an individual for the costs of exchange if there are fixed exchange costs? The question is not answered easily because the relevant comparison is the utility levels obtained with costly trade and with autarky. What must the terms of trade be to compensate an individual for fixed transactions costs? Some intuition about this can be obtained by considering the effect of a small change in fixed costs in the borderline case when the individual is indifferent between the autarkic consumption bundle and the consumption bundle obtained by paying a fixed cost and trading.

Let (c_1^A, c_2^A) denote the individual's consumption bundle under autarky, and (c_1^T, c_2^T) denote the individual's consumption bundle with fixed exchange costs. Indifference implies that

$$U(c_1^T, c_2^T) = U(c_1^A, c_2^A). \tag{A.1}$$

When production technologies are linear and the terms of trade, λ, are such that $\lambda \neq f_2'/f_1'$, individuals specialize in the production of one of the goods, good 1 or good 2. The budget constraint

$$\lambda c_1^T + c_2^T = \lambda f_1(\Gamma) - \Omega \tag{A.2}$$

is satisfied for individuals specializing in good 1, and

$$\lambda c_1^T + c_2^T = f_2(\Gamma) - \Omega \tag{A.3}$$

for individuals specializing in good 2.

The optimality of (c_1^T, c_2^T) with trade implies that the individual equates the ratio of the marginal utilities of the two goods to the exchange ratio, λ, so that

$$U_1(c_1^T, c_2^T) = \lambda U_2(c_1^T, c_2^T) \tag{A.4}$$

is satisfied.

Equation A.4, equation A.1, and one of the two budget constraints above implicitly determine the consumption bundle in the case of specialization and the terms of trade that are necessary to compensate the individual for trading when there is a fixed transaction cost, Ω.

In principle these equations can be solved to determine the effect of Ω on λ when an individual specializes in the production of either of the two goods. A closed-form solution cannot be obtained, however, without making explicit assumptions about functional forms. Linear approximations to these functions that will be accurate predictors of the effects of small transaction costs can be made, however.

These linear approximations are obtained by implicitly differentiating the equations and evaluating the resulting expressions at $\Omega = 0$. Note that if $\Omega = 0$, $\lambda = f_2'/f_1'$ if an individual is to be indifferent between trading and autarky. Furthermore, at these terms of trade, the individual must be indifferent between specializing in the production of good 1 or good 2. Consequently, either of the budget constraints above can be used as the starting point of the approximation. Finally, note that optimization requires the individual to equate the ratio of the marginal utilities of the two goods to terms of trade. In the absence of exchange costs, or under autarky, the condition is

$$\frac{U_1(c_1^A, c_2^A)}{U_2(c_1^A, c_2^A)} = \frac{f_2'}{f_1'}\ . \tag{A.5}$$

To economize on notation, we use an overdot to represent differentiation with respect to the fixed cost of exchange, Ω, i.e., $\dot{\lambda} = d\lambda/d\Omega$. Regardless of whether the individual specializes in good 1 or 2, we can implicitly differentiate Equation A.4 to obtain,

$$(A.6) \qquad U_1(c_1^T, c_2^T)\dot{c}_1^T + U_2(c_1^T, c_2^T)\dot{c}_2^T = 0.$$

Evaluating this expression at $\Omega = 0$ yields

$$(A.7) \qquad \frac{f_2'}{f_1'}\dot{c}_1^T(0) + \dot{c}_2^T(0) = 0.$$

Implicitly differentiating Equation A.2, on the assumption the individual specializes in good 1, yields

$$(A.8) \qquad \dot{\lambda}c_1^T + \lambda\dot{c}_1^T + \dot{c}_2^T = \dot{\lambda}f_1(\Gamma) - 1.$$

Evaluating this expression at $\Omega = 0$, yields

$$(A.9) \qquad \dot{\lambda}(0)c_1^A + \frac{f_2'}{f_1'}\dot{c}_1^T(0) + \dot{c}_2^T(0) = \dot{\lambda}(0)\, f_1(\Gamma) - 1.$$

Combining Equations A.7 and A.9 and solving for $\dot{\lambda}(0)$ yields

$$(A.10) \qquad \dot{\lambda}(0) = \frac{1}{f_1(\Gamma) - c_1^A}.$$

Equation A.10 shows the effect on λ of a small fixed transaction cost, evaluated at the point $\Omega = 0$. Hence, an individual will be induced to specialize in good 1 and trade only if the terms of trade are approximately

$$(A.11) \qquad \lambda(\Omega) \approx \frac{f_2'}{f_1'} + \left(\frac{1}{f_1(\Gamma) - c_1^A}\right)\Omega$$

or larger. This result is intuitive. The quantity $f_1(\Gamma) - c_1^A$ is approximately the amount of good 1 that the individual must give up to trade. Hence, $\Omega/(f_1(\Gamma) - c_1^A)$ is the extra amount of good 2, per unit of good 1 exchanged, that the individual must obtain to be compensated for the fixed cost of entering the market.

A similar analysis applies to individuals that specialize in the production of good 2. In this case, Equation A.3 is differentiated to obtain

$$(A.12) \qquad \dot{\lambda}c_1^T + \lambda\dot{c}_1^T + \dot{c}_2^T = -1.$$

Evaluating Equation A.12 at $\Omega = 0$, as before, yields,

$$(A.13) \qquad \dot{\lambda}(0)c_1^A + \frac{f_2'}{f_1'}\dot{c}_1^T(0) + \dot{c}_2^T(0) = -1.$$

Combining Equations A.7 and A.13, yields

$$(A.14) \qquad \dot{\lambda}(0) = -\frac{1}{c_1^A}.$$

Hence, an individual will be induced to specialize in good 2 and trade only if the terms of trade are

$$(A.15) \qquad \lambda(\Omega) \approx \frac{f_2'}{f_1'} - \frac{1}{c_1^A}\Omega$$

or smaller. The quantity $-\Omega/c_1^A$ is the discount per unit of good 1 purchased required to compensate the individual for the fixed cost of trade.