A Search-Theoretic Approach to Monetary Economics

By Nobuhiro Kiyotaki and Randall Wright*

The essential function of money is its role as a medium of exchange. We formalize this idea using a search-theoretic equilibrium model of the exchange process that captures the "double coincidence of wants problem" with pure barter. One advantage of the framework described here is that it is very tractable. We also show that the model can be used to address some substantive issues in monetary economics, including the potential welfare-enhancing role of money, the interaction between specialization and monetary exchange, and the possibility of equilibria with multiple fiat currencies. (JEL E00, D83)

Since the earliest writings of the classical economists it has been understood that the essential function of money is its role as a medium of exchange. The use of monetary exchange helps to overcome the difficulty associated with pure barter in economies where trade is not centralized through some perfect and frictionless market. Many attempts have been made in the literature to formalize this, with varying degrees of success. In this paper, we present a search-

* Kiyotaki: Federal Reserve Bank of Minneapolis and Department of Economics, University of Minnesota, Minneapolis, MN 55455; Wright: Federal Reserve Bank of Minneapolis and Department of Economics, University of Pennsylvania, Philadelphia, PA 19104. This is a much revised version of our earlier working paper, "Search for a Theory of Money." Part of the work on that project was accomplished while both authors were at the London School of Economics in the spring of 1990, and we are grateful for that institution's hospitality. The National Science Foundation, the University of Pennsylvania Research Foundation, and the Hoover Institution all provided financial support. We also thank many friends, colleagues, students, and seminar participants for their input. Peter Diamond, Robert Lucas, Dale Mortensen, David Romer, Neil Wallace, and two anonymous referees made some especially useful suggestions, although they should not be held accountable for the final product. The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 There is a voluminous literature on the foundations of monetary theory, and rather than attempting to cite all of the relevant work, we refer the reader to the survey by Joseph M. Ostroy and Ross M. Starr (1990). We would, however, like to mention the contribution of Robert A. Jones (1976) and the extensions by Seongwhan Oh (1989) and Katsuhito Iwai (1988). Although there are many technical differences, that model is definitely related in spirit to the search-theoretic approach we describe here. In particular, there are heterogeneous agents and commodities, and in equilibrium certain commodities are chosen as media of exchange in order to reduce search costs.
money. In Kiyotaki and Wright (1991), we used an alternative search-based model to illustrate the robustness of monetary equilibria; that is, fiat money can be valued as a medium of exchange even if it has intrinsic properties, like its rate of return, that are inferior to other available assets. We also constructed an example in that model to show how the use of fiat money can affect welfare.

However, due to the generality of the specification in that paper, we were not able to say much about the features of monetary equilibria, other than that they exist and are robust, and our characterization of welfare did not proceed much beyond a numerical example. The model to be presented in this paper can be thought of as a simplified version of Kiyotaki and Wright (1991). Our first objective is to demonstrate that this class of models is actually very tractable. Our second objective is to convince the reader that search-based models can be used not just to determine which objects serve as media of exchange or to prove the existence of valued fiat money, but to address some more applied issues in monetary economics as well. In particular, we use the model to discuss the potential welfare-enhancing role of money, the interaction between specialization and monetary exchange, and the possibility of equilibria with multiple currencies.

The rest of the paper is organized as follows. In Section I we describe the basic model. In Section II we characterize the welfare effects of money. Among other things, the model implies that equilibria where fiat money is universally acceptable are generally superior to nonmonetary equilibria and to equilibria where it is only partially acceptable. In Section III we introduce specialization by producers, by assuming that they face a trade-off between productivity and the marketability of their output. The model implies that use of money, by making exchange easier, leads to more specialized and, therefore, more efficient production. In Section IV we discuss a version of the model that allows for multiple fiat currencies. In Section V we conclude.

I. The Basic Model

The economy is populated by a large number of infinite-lived agents, with total population normalized to unity. There is also a large number of consumption goods. These consumption goods are indivisible and come in units of size one. We refer to them as real commodities, to distinguish them from fiat money, which is an object that no one ever consumes and can be thought of as a collection of pieces of paper or certain types of seashells, for example, with no intrinsic value. A crucial feature of the model is that there is an exogenous parameter \( x \), with \( 0 < x < 1 \), that captures the extent to which real commodities and tastes are differentiated. In particular, \( x \) equals the proportion of commodities that can be consumed by any given agent, and \( x \) also equals the proportion of agents that can consume any given commodity.\(^2\) If a commodity is one of those that can be consumed by an agent, then we say that it is one of his consumption goods. Consuming one of his consumption goods yields utility \( U > 0 \), while consuming other commodities (or money) yields zero utility.

Initially, a fraction \( M \) of the agents are each endowed with money while \( 1 - M \) are each endowed with one real commodity, where \( 0 < M < 1 \). Money may or may not have value. If it does, then it is convenient to assume that agents who are initially endowed with money are endowed with exactly one unit of real balances, so that in order to buy a real commodity they must spend all of their cash. There are two ways to guarantee that this is the case. First, and

\(^2\) For example, suppose there are \( K \) distinct goods and each agent consumes \( k \) of them; then \( x = k / K \). Alternatively, suppose there is a continuum of goods indexed by points around a circle of circumference 1 and each agent consumes goods corresponding to points in a fixed arc; then, \( x \) is the length of that arc.
most straightforwardly, we can simply assume that the monetary object is indivisible, like the real commodities in the model. Then if money trades at all it must trade one-for-one against a real commodity, and each agent with one indivisible unit of money will have one unit of real balances. Alternatively, we can assume that money is divisible, determine the price level endogenously for a given stock of nominal currency, and endow some agents at the initial date with exactly enough nominal currency to constitute a single unit of real balances. We begin with the former approach of assuming that money is indivisible and take up the latter, slightly more complicated, approach later.

Money and commodities are costlessly storable. Money cannot be produced by any private agent, while real commodities can be produced according to the following technology. One unit of output requires two inputs: a consumption good and a random amount of time. That is, once an agent consumes he enters a production process that yields one unit of one real commodity, drawn randomly from the set of all commodities, according to a continuous-time Poisson process with arrival rate $\alpha > 0$. Thus, $\alpha$ measures productivity in the sense of average output per unit time. Note that agents who have not consumed cannot produce. Furthermore, as is standard in the equilibrium search literature, we assume that agents cannot consume their own output (see e.g., Peter A. Diamond, 1982, 1984; Kiyotaki and Wright, 1991). This assumption helps to simplify the presentation and to facilitate comparison with earlier models, but as we show in Appendix A it is otherwise completely unnecessary.

An agent who has just produced enters an exchange sector where he looks for other agents with whom to trade. Traders in the exchange sector meet pairwise and at random according to a Poisson process with constant arrival rate $\beta > 0$. When two traders meet, exchange takes place if and only if it is mutually agreeable, that is, if and only if both agents are at least as well off after the trade. Because there is a large number of anonymous agents, all trade is quid pro quo (there can be no IOU’s or other forms of private credit). We also assume that there is a transaction cost $\epsilon$ in terms of disutility, where $0 < \epsilon < U$, that must be paid by the receiver whenever any real commodity is accepted in trade. This transaction cost implies that a trader who is indifferent between holding two real commodities will never trade one for the other. For simplicity, we assume for most of the presentation that the transaction cost of accepting fiat money is zero; this simplifies the presentation considerably but, as we show in Appendix B, it is possible (and interesting) to relax this assumption.

Since exchange takes place if and only if mutually agreeable, an agent with either one unit of real balances or one real commodity cannot acquire additional money or another commodity except by giving up his entire inventory. Furthermore, no agent in the exchange sector can produce anything until he trades for one of his consumption goods and consumes, given the specified technology. These observations have the following implication: if each trader starts at the initial date with either one unit of real balances or one real commodity, then in equilibrium all traders will always have either one unit of real balances or one real commodity.  

$^3$The assumption that the arrival rate $\beta$ is constant (and independent of the number of traders) is equivalent to the assumption of a constant-returns-to-scale (CRS) meeting technology. That is, a CRS meeting technology implies that the total number of meetings per unit time is proportional to the number of traders, and so the arrival rate for a representative trader (which is just the number of meetings divided by the number of traders) is a fixed constant. We ignore degenerate outcomes in which there are no agents in the exchange sector, and hence, the arrival rate for an individual should be enter this sector would be 0.

Note that there are no physical restrictions in the model against storing more than one commodity, storing arbitrary quantities of money, or storing money and commodities simultaneously. Rather, these results are due to the assumption that consumption is a necessary
are referred to as commodity traders, while agents with fiat money are referred to as money traders. Let \( \mu \) denote the fraction of traders who are money traders, so that a trader located at random has money with probability \( \mu \) and a real commodity with probability \( 1 - \mu \).

Individuals choose strategies for deciding when to accept various commodities and fiat money in order to maximize their expected discounted utility from consumption net of transaction costs, taking as given the strategies of others. We look for Nash equilibria. We restrict attention for the most part to symmetric equilibria, where all agents and all real commodities are treated the same, and to steady-state equilibria, where strategies and all aggregate variables are constant over time. To construct the set of such equilibria, we describe some basic properties that they must satisfy, use these properties to describe an individual trader’s best-response correspondence, and determine its fixed points.

The first thing to note is that an agent always accepts a real commodity if it is one of his consumption goods, whereupon he immediately consumes it and enters the production process. Also, we claim that a commodity trader will never accept a commodity that is not one of his consumption goods. This is due to the fact that in a symmetric equilibrium no real commodities are treated as special, and therefore, the probability of a trade offer being accepted by the next agent one meets is independent of the type of commodity one has. Hence, there is no advantage to trading one real commodity for another, and since there is a transaction cost \( \varepsilon \), unless a commodity is going to be consumed it will never be accepted. This means that \( x \) is the probability that a commodity trader located at random is willing to accept any given commodity, and therefore \( x^2 \) is the probability that two commodity traders consummate a barter transaction. This is precisely William Stanley Jevons’s (1875) “double coincidence of wants problem” with direct barter: not only do you have to meet someone with something that you want, this someone also has to want what you have.5

The next thing to determine is whether individuals accept money. Let \( \Pi \) denote the probability that a random commodity trader accepts money and let \( \pi \) be the best response of a representative individual. We will solve the best-response problem using dynamic programming. Let \( V_j \) denote the payoff or value function for the individual in state \( j \), where \( j = 0, 1, \) or \( m \) indicates that he is a producer, a commodity trader, or a money trader, respectively. Then, if \( r > 0 \) is the rate of time preference, Bellman’s equations are given by

\[
(1) \quad rV_0 = \alpha (V_1 - V_0)
\]

\[
(2) \quad rV_1 = \beta (1 - \mu) x^2 (U - \varepsilon + V_0 - V_1) + \beta \mu \max_\pi (V_m - V_1)
\]

\[
(3) \quad rV_m = \beta (1 - \mu) \Pi x (U - \varepsilon + V_0 - V_m).
\]

5The result that traders never accept commodities that are not their consumption goods means that there is no commodity money in a symmetric equilibrium. This is not to say that the model cannot have nonsymmetric equilibria, in which some real commodities do become media of exchange, but only that we restrict attention to symmetric outcomes here. Commodity money is analyzed in a related model in Kiyotaki and Wright (1989). Although a small transaction cost \( \varepsilon \) guarantees that there will be a double-coincidence problem in a symmetric equilibrium, the double-coincidence problem arises without transaction costs in the asymmetric equilibria studied in Kiyotaki and Wright (1989) and Aiyagari and Wallace (1991). Another way to guarantee that there is a double-coincidence problem is to assume that a real commodity can only be stored by its producer, as in Kiminori Matsuyama et al. (1993), which seems natural if we interpret these commodities as services rather than goods. Under this assumption, we can dispense with the transaction cost entirely.
Equations like these are standard in search theory (formal derivations for a closely related model can be found in Kiyotaki and Wright [1991], for example). They have the following interpretation. According to (1), the flow return to a producer, \( rV_0 \), equals the rate at which output is produced, \( \alpha \), times the gain from switching from production to exchange, \( V_1 - V_0 \). According to (2), the flow return to a commodity trader equals the sum of two terms. The first term is the rate at which he meets other commodity traders, \( \beta(1-\mu) \), times the probability that both want to trade, \( x^2 \), times the gain from trading, consuming, and switching back to production, \( U - \epsilon + V_0 - V_1 \). The second term is the rate at which he meets money traders, \( \beta \mu \), times the probability that a money trader wants to trade, \( x \), times the gain from accepting money with probability \( \pi \), where \( \pi \) is chosen optimally. According to (3), the flow return to a money trader equals the rate at which he meets commodity traders, \( \beta(1-\mu) \), times the probability that both want to trade, \( \Pi x \), times the gain from trading, consuming, and switching to production, \( U - \epsilon + V_0 - V_m \).

The above dynamic program depends not only on the strategies of others, as represented by \( \Pi \), but also on \( \mu \), the proportion of traders holding money. However, \( \mu \) can be determined as a function of \( \Pi \) and the initial endowment of money, \( M \). Begin by letting \( N_0 \), \( N_1 \), and \( N_m \) denote the proportions of the population who are producers, commodity traders, and money traders. Then the model has a dynamic structure with transitions as illustrated in Figure 1. To determine its steady state, we equate the flow out of and into production:

\[
4. \quad \alpha N_0 = \beta(1-\mu) x^2 N_1 + \beta(1-\mu) \Pi x N_m.
\]

If we use the fact that the \( N_j \)'s sum to 1 and

\[\text{the fact that } N_m = M \text{ (the number of money traders equals the number of agents endowed with money), (4) can be reduced to}
\]

\[
5. \quad M = \frac{\alpha \mu}{(\alpha + \phi)}
\]

where \( \phi = \phi(\mu, \Pi) \) is defined by\footnote{Note that \( \phi \) can be interpreted as consumption per trader per unit time: it is the rate at which a representative trader meets commodity traders, \( \beta(1-\mu) \), times the probability that a deal is consummated, which is the probability our representative trader has money and they trade, \( \mu \Pi \), plus the probability that our representative trader has a real commodity and they trade, \( (1-\mu) x^2 \).}

\[
6. \quad \phi = \beta(1-\mu) \left[ \mu x \Pi + (1-\mu) x^2 \right].
\]

Equation (5) is a quadratic in \( \mu \), and for any \( M \in [0,1] \) and \( \Pi \in [0,1] \) there will exist a unique value of \( \mu = \mu(M, \Pi) \) in \([0,1]\) satisfying this equation. Furthermore, one can show that \( \mu(0, \Pi) = 0 \), \( \mu(1, \Pi) = 1 \), \( \delta \mu / \delta M > 0 \), and \( \delta \mu / \delta \Pi > 0 \). Given \( \mu = \mu(M, \Pi) \), the unique steady state is fully described by

\[
7. \quad N_0 = \frac{\phi}{(\alpha + \phi)}
\]

\[
N_1 = (1-\mu) x / (\alpha + \phi).
\]

However, for the purpose of analyzing the above dynamic program, \( \mu \) summarizes all the agent needs to know about the steady state. If we insert \( \mu = \mu(M, \Pi) \) into (1)–(3),

\[\text{We have implicitly assumed that a money trader never accepts a commodity that is not one of his consumption goods, but one can show that this is always true in equilibrium. That is, one can show that the only time an agent would want to exchange money for a commodity that is not one of his consumption goods is when money is valueless, in which case he cannot.}\]
then, given $M$, this dynamic program defines a correspondence from $\Pi$ to best responses, $\pi$. The set of equilibria is the set of fixed points of this correspondence.

To characterize this set, first note that if $\Pi < x$ then (1)–(3) imply that $V_m < V_1$, which implies that the best response is $\pi = 0$. Intuitively, if money is being accepted with a lower probability than a barter offer, then it is harder to trade using money than barter, and so the best response is never to exchange a real commodity for money. Second, if $\Pi > x$, then (1)–(3) imply that $V_m > V_1$, which implies $\pi = 1$. If money is being accepted with a greater probability than a barter offer, then it is easier to trade using money, and so the best response is to exchange a real commodity for money whenever possible. Finally, if $\Pi = x$, then (1)–(3) imply that $V_m = V_1$, which implies that $\pi$ can be anything in $[0,1]$. If monetary exchange and barter are equally easy then traders are indifferent between having money and real commodities, and they could accept money with any probability. Based on these results, the best-response correspondence is as shown in Figure 2, and there are exactly three equilibria: $\Pi = 0$, $\Pi = 1$, and $\Pi = x$.

The equilibrium with $\Pi = 0$ will be called the nonmonetary equilibrium. In this case, agents expect that money will be valueless, so they never accept it, and this expectation is self-fulfilling. The equilibrium with $\Pi = 1$ will be called the pure-monetary equilibrium. In this case, agents expect that money will be universally acceptable, and so they always take it, and this expectation is self-fulfilling. Finally, the equilibrium with $\Pi = x$ will be called the mixed-monetary equilibrium. In this case, agents are indifferent between accepting and rejecting money as long as other agents take it with probability $\Pi = x$, and so partial acceptability can also be self-fulfilling. Alternatively, a symmetric mixed-strategy equilibrium where all agents accept money with probability $x$ could be reinterpreted as a nonsymmetric pure-strategy equilibrium, where a fraction $x$ of agents accept money with probability 1 while the rest accept it with probability 0.8

**II. Welfare**

The first thing we want to do in this section is to compare utility across the various equilibria, for a given value of $M$. For the purpose of this comparison, we keep things tractable by restricting attention to the limiting case where $\alpha \to \infty$. In this case production is instantaneous, and so all agents are either money traders or commodity traders: $N_m = M$, $N_1 = 1 - M$, and $\mu = M$. This makes it relatively easy to solve (1)–(3) for the reduced-form payoffs:

$$rV_1 = \psi(r\alpha + \beta x \Pi[M \Pi + (1 - M)x])$$

(8)

$$rV_m = \psi(r\Pi + \beta x \Pi[M \Pi + (1 - M)x])$$

(9)

where $\psi = (U - \epsilon)\beta(1 - M)x/(r + \beta x \Pi)$. We can now substitute $\Pi = 0$, $\Pi = x$, and $\Pi = 1$ into (8) and (9) and compare utility across equilibria for commodity traders and money traders.

---

8There can exist non-steady-state equilibria in this model where the probability that money is accepted varies over time. An example of a "sunspot equilibrium," in which the probability that money is accepted fluctuates randomly over time even though the fundamentals are nonstochastic and time-invariant, is constructed in Kiyotaki and Wright (1990).
If we let the superscripts \( N, M, \) and \( P \) refer to the nonmonetary, mixed-monetary, and pure-monetary equilibria, respectively, then we have the following results. First, commodity traders are equally well off in the nonmonetary and mixed-monetary equilibria and strictly better off in the pure-monetary equilibrium: \( V_N^V = V_M^V < V_P^V. \) Second, money traders are strictly better off in the pure-monetary equilibrium than in the mixed-monetary equilibrium and strictly better off in the mixed-monetary equilibrium than in the nonmonetary equilibrium: \( V_N^V < V_M^V < V_P^V. \) Thus, given the initial endowment of money and real commodities, all agents are at least weakly better off if money is acceptable than if it is not, and all agents are strictly better off if money is universally acceptable than if it is only partially acceptable.9

The next thing we want to do is to examine how utility varies with \( M, \) and for the purpose of this comparison we return to the general case where \( \alpha \) need not be \( \infty. \) The experiment we consider is to increase the number of agents initially endowed with money and to reduce the number initially endowed with real output, so that we can maintain the tractability of the unit-inventory assumption. In either the nonmonetary equilibrium or the mixed-monetary equilibrium, all agents are better off the lower is \( M. \) The reason is that, in these equilibria, money does nothing to ameliorate the double-coincidence problem, and so it is better to endow everyone with real consumption goods rather than intrinsically worthless paper or seashells. The more interesting case is the pure-monetary equilibrium, where fiat currency does have a genuine role to play in facilitating exchange.

To pursue this, define the welfare criterion

\[
W = N_0V_0 + N_1V_1 + N_mV_m.
\]

This can be interpreted as the \textit{ex ante} expected utility of all agents before the initial endowment of money and output is randomly distributed among them. After some algebra, (10) can be simplified to yield

\[
rW = (U - \varepsilon)\varphi \alpha / (\alpha + \varphi)
\]

where \( \varphi = \varphi(\mu, \Pi) \) was defined in (6) above.10 Now consider maximizing \( W \) with respect to \( M. \) Since \( W \) is increasing in \( \varphi, \) we proceed by finding the value \( M^o \) that maximizes \( \varphi \) with respect to \( \mu, \) and then determine the optimal value \( M^o \) from the steady-state condition (5), \( M = \alpha \mu / (\alpha + \varphi). \)

The result is as follows: if \( x \geq \frac{1}{2} \) then \( M^o = 0, \) which implies \( M^o = 0; \) if \( x < \frac{1}{2} \) then \( M^o = (1 - 2x)/(2 - 2x), \) which implies \( M^o > 0. \) Intuitively, when \( x \geq \frac{1}{2}, \) each agent is willing to consume (and therefore accept) at least half of the commodities produced in the economy, and pure barter is not very difficult. In this case the role for a medium of exchange is not very important, and it is optimal to endow everyone with real output and no one with money. When \( x < \frac{1}{2}, \) on the other hand, pure barter is sufficiently difficult that the introduction of some fiat money improves welfare, in spite of the fact that, in the experiment under consideration, endowing some agents with money requires endowing fewer agents with real output at the initial date. We also note that \( M^o \) is decreasing in \( x, \) and that \( M^o \to \frac{1}{2} \) as \( x \to 0. \) Thus, as \( x \) shrinks and the double-coincidence problem becomes more difficult it is optimal to endow more agents with money.

We now turn to a version of the model in which money can be interpreted as being divisible rather than indivisible and investigate the welfare implications of a particular mechanism for determining the price level.

---

9 These results differ from those in Kiyotaki and Wright (1990), where we assumed that agents initially endowed with fiat currency would freely dispose of it and produce a new commodity in the nonmonetary equilibrium. This made the initial stock of real commodities differ across monetary and nonmonetary equilibria and therefore made welfare comparisons ambiguous. Following Aiyagari and Wallace (1992), we assume here that agents initially endowed with fiat money cannot produce until they consume, which keeps the initial stock of output constant across equilibria.

10 Notice that \( rW \) equals the difference \( U - \varepsilon \) times aggregate consumption, since \( \varphi \) is consumption per trader per unit time and \( \alpha / (\alpha + \varphi) = N_1 + N_m \) is the number of traders.
P. We look for a pure-monetary equilibrium in which each money trader carries P units of cash and all P units are required to purchase one real commodity. Then real balances are given by \( M = C / P \) for any exogenous stock of nominal currency C. Of course, to make P endogenous we need to impose an additional equilibrium condition. Consider the method used by Diamond (1984) in his cash-in-advance search model, which is to impose as an equilibrium condition that the gains from trade for a commodity trader and a money trader are equalized whenever an exchange is made between them:

(12) \( V_m - V_1 = U - \epsilon + V_0 - V_m \).

The left-hand side is the gain from trade for a commodity trader who accepts money, while the right-hand side is the gain for a money trader who acquires one of his consumption goods.

If a pure-monetary equilibrium satisfies condition (12), we call it a split-the-surplus equilibrium. Notice that both sides of (12) depend on \( \mu \). If \( \mu \) is large, there are many money traders and few commodity traders, so having money is not very desirable; hence, a commodity trader who acquires money gets a smaller gain than a money trader who acquires one of his consumption goods. Thus, for large \( \mu \) the right-hand side of (12) exceeds the left-hand side, and we need to reduce \( \mu \) until either (12) holds or we hit \( \mu = 0 \). Inserting the reduced-form payoff functions and simplifying, the unique value of \( \mu \) that satisfies (12) is given by

(13) \( \mu^* = \frac{(1-2x)}{(2-2x)} - \frac{r}{2\beta x(1-x)} \).

If \( r < \beta x(1-2x) \) then \( \mu^* > 0 \), which implies a unique \( M^* > 0 \) satisfying (5), and a finite equilibrium price level \( P^* = C / M^* \). If \( r \geq \beta x(1-2x) \), then (12) cannot be satisfied for any value of \( \mu > 0 \). In this case, we say that the split-the-surplus equilibrium entails \( M^* = 0 \) and \( P = \infty \), and hence there can be no monetary exchange.

Recall that the value of \( \mu \) that maximizes \( W \) is \( \mu^o = (1-2x)/(2-2x) \), and comparing this with (13) we find \( \mu^o > \mu^* \). This means the split-the-surplus equilibrium yields a lower value of \( \mu \), and hence a lower value of \( M \) and a higher value of \( P \) for any given \( C \), than that which maximizes ex ante utility. However, the split-the-surplus equilibrium is still ex post Pareto optimal. To see why, consider the value functions \( V_m \) and \( V_1 \) (\( V_0 \) is proportional to \( V_1 \) and need not be considered independently). One can show that, as functions of \( M \), both are concave, and \( V_1 \) is increasing but \( V_m \) is decreasing at \( M = M^* \) (see Fig. 3). Hence, any movement away from \( M^* \) that makes commodity traders better off makes money traders worse off, and vice versa. The split-the-surplus equilibrium is therefore ex post efficient even though it fails to maximize ex ante welfare.\(^{11}\)

\(^{11}\)Figure 3 indicates that both commodity and money traders prefer a lower value of \( M \) than that which maximizes \( W \). This ostensibly paradoxical result can be understood by noting that the number of agents of each type varies with \( M \).
III. Specialization

An insight dating back at least to Adam Smith (1776) is that specialization is limited by the extent of the market and that the use of money encourages specialization by enlarging the extent of the market. As Smith puts it:

When the division of labour has been once thoroughly established, it is but a very small part of a man’s wants which the produce of his own labour can supply. He supplies the far greater part of them by exchanging that surplus part of the produce of his own labour, which is over and above his own consumption, for such parts of the produce of other men’s labour as he has occasion for. Every man thus lives by exchanging, or becomes in some measure a merchant, and the society itself grows to be what is properly a commercial society.

But when the division of labour first began to take place, this power of exchanging must frequently have been very much clogged and embarrassed in its operations. One man, we shall suppose, has more of a certain commodity than he himself has occasion for, while another has less. ... But if this latter should chance to have nothing that the former stands in need of, no exchange can be made between them. ... In order to avoid the inconveniency of such situations, every prudent man in every period of society, after the first establishment of the division of labour, must naturally have endeavored to manage his affairs in such a manner, as to have at all times by him, besides the peculiar produce of his own industry, a certain quantity of some one commodity or other, such as he imagined few people would be likely to refuse in exchange for the produce of their industry.

Smith is suggesting that specialization, while it may have desirable consequences in terms of productivity, makes barter difficult. Whenever they can, specialized producers will therefore tend to sell their output for a more generally acceptable medium of exchange, which they will then use to buy whatever consumption goods they desire. Hence, specialization leads to a greater role for money, while at the same time the use of money affords a greater opportunity for specialization by facilitating the process of exchange. In order to formalize this, we introduce a trade-off between productivity and marketability by assuming that the arrival rate in the production process is a function of the number of agents willing to consume the output: \( a = a(x) \), where \( a' < 0 \).

The idea is that, by becoming more specialized, a producer can increase output per unit time, \( a \), but only at the cost of reducing the fraction of consumers who will accept his output in exchange, \( x \).

Before entering the production process, agents choose \( x \), taking as given the behavior of others. If money is accepted with probability \( \Pi \) and other producers’ decisions imply that a given individual can consume a fraction \( X \) of their output, his payoff if he chooses \( x \) is described by

\[
(14) \quad rV_0 = a(x)[V_1(x) - V_0]
\]

\[
(15) \quad rV_1(x) = \beta(1 - \mu)X[x(U - \varepsilon + V_0 - V_1(x))] + \beta x \mu \Pi[V_m - V_1(x)]
\]

\[
(16) \quad rV_m = \beta(1 - \mu)X\Pi(U - \varepsilon + V_0 - V_m).
\]

The choice of \( x \) will be made by a producer to maximize the right-hand side of (14), and this \( x \) is then carried over to the exchange sector as a state variable. As was the case earlier, this best-response problem also depends on \( \mu \), but \( \mu \) will be determined below.

\textsuperscript{12} One interpretation is that each consumer derives utility from a fixed set of characteristics embodied in some commodities, and a larger value of \( x \) implies that the producer’s output contains a greater number of characteristics and hence has a larger potential market. Eduardo Siandra (1990) has independently developed a very similar model. Robert King and Charles Plosser (1986) and Harold L. Cole and Alan C. Stockman (1992) provide other analyses of the interaction between money and specialization.
using the steady-state conditions as a function of strategies and $M$.

Equations (14)–(16) can be solved for $V_0$, $V_1$, and $V_m$. In particular, after simplification, we find that

$$\alpha(x)[V_1(x) - V_0] = \phi x \alpha(x)/\xi(x) = Z(x)$$

where

$$\xi(x) = [r + \beta(1 - \mu)X\Pi][r + \beta(1 - \mu)X + \beta \mu x \Pi]$$

and $\phi$ does not depend on $x$. The individual choice of $x$ can be found by maximizing $Z(x)$. If we assume an interior solution, the first-order condition $Z'(x) = 0$ can be rearranged to yield:\footnote{One can also maximize the right-hand side of (14) directly by setting $
alpha'(x)[V_1(x) - V_0] + \alpha(x)V_1'(x) = 0$
where $V_1'(x)$ is, from (15),

$$V_1'(x) = rV_1(x)/[rx + \beta x^2[(1 - \mu)X + \mu \Pi]].$$

Manipulating these equations yields the same first-order condition as in the text, equation (18).}

$$x\alpha'(x)/\alpha(x) = r + \alpha(x)$$

$$= r + \beta(1 - \mu)X + \beta \mu x \Pi.$$\footnote{Notice that the pure- and mixed-monetary equilibria must be unique, but since $\Pi = 0$ implies that both the $X = x(X, \mu, \Pi)$ curve and the $M = M(X, \mu, \Pi)$ curve are upward-sloping, they could intersect more than once, and there could be more than one nonmonetary equilibrium. Although examples with multiple nonmonetary equilibria can be constructed, we rule this out in the following discussion.}

It can be shown that the second-order condition $Z'' < 0$ holds if we assume that $\alpha'' < 0$. Then (18) completely characterizes the individual’s choice of $x$, given $X$, $\Pi$, and $\mu$. We write $x = x(X, \mu, \Pi)$.

For a symmetric equilibrium we must have $X = x$, or $X = x(X, \mu, \Pi)$. Another equilibrium condition comes from the steady-state equation $M = \alpha \mu / (\alpha + \phi)$. Since the right-hand side of this equation depends on $X$, \mu, and $\Pi$, we write it as $M = M(X, \mu, \Pi)$. Given $M$, an equilibrium is a solution to $X = x(X, \mu, \Pi)$ and $M = M(X, \mu, \Pi)$ with either $\Pi = 0$, $\Pi = X$, or $\Pi = 1$ (since, for any given $X$, the model has a nonmonetary equilibrium, a mixed-monetary equilibrium, and a pure-monetary equilibrium, exactly as in the model without endogenous specialization). In Figure 4 we draw the locus of points in $(\mu, X)$-space satisfying each of these conditions. Notice that the $M = M(X, \mu, \Pi)$ curve is upward-sloping, shifts to the right as $\Pi$ increases, and goes through $(M, 0)$. Also, the $X = x(X, \mu, \Pi)$ curve is upward-sloping, horizontal, or downward-sloping depending on whether $\Pi = 0$, $X$, or $1$, and goes through the same intercept $(0, X_0)$ in any case. The intersection of these two curves determines the equilibrium values of $\mu$ and $X$ in the nonmonetary equilibrium, the mixed-monetary equilibrium, or the pure-monetary equilibrium, depending on whether $\Pi = 0$, $\Pi = X$, or $\Pi = 1$.

Again let the superscripts $N$, $M$, and $P$ represent the nonmonetary, mixed-monetary, and pure-monetary equilibria. As can be seen from the diagram, specialization is greatest in the pure-monetary equilibrium, lower in the mixed-monetary equilibrium, and lowest in the nonmonetary equilibrium: $X_P < X_M < X_N$. The intuition behind this result is that when money circulates there is less of an advantage to having a high value of $x$, since it does not necessarily require a double coincidence of wants in order to exchange. We can also ask how specialization depends on $M$. An increase in $M$ shifts the $M = M(X, \mu, \Pi)$ curves to the right but does not affect the $X = x(X, \mu, \Pi)$ curves. As can be seen from the diagram, when $M$ increases, the result is an increase in $X$ in the nonmonetary equilibrium, a decrease in...
Figure 4. M AND X CURVES, WHOSE INTERSECTIONS DETERMINE EQUILIBRIUM VALUES OF $\mu$ AND $X$

$X$ in the pure-monetary equilibrium, and no change in the mixed-monetary equilibrium. Roughly speaking, an increase in $M$ in the pure-monetary equilibrium encourages specialization because producers can more easily market their specialized output when there is more money in circulation.\(^{15}\)

Consider now the effect of increasing the arrival rate in the exchange sector, $\beta$, which can be thought of as reducing the frictions associated with trade (or “increasing the extent of the market”). In any pure-monetary equilibrium, an increase in $\beta$ shifts the $X = x(X, \mu, 1)$ curve down and shifts the $M = M(X, \mu, 1)$ curve to the right, resulting in a decline in $X$ and therefore an increase in specialization and productivity. As $\beta \to \infty$, $x \to 0$, and specialization becomes complete. As this happens, barter becomes extremely difficult, and the ratio of the volume of barter to monetary exchange vanishes.\(^{16}\) In the limit, agents almost always sell their production goods for money and use money to buy their consumption goods; as Robert W. Clower (1965) puts it, “money buys goods and goods buy money; but goods do not buy goods.” In this model, however, there is no constraint that agents must use cash. To

\(^{15}\)The discussion of the effect of an increase in $M$ takes the real money supply to be exogenous, say, because the monetary object is indivisible. Alternatively, we can assume that money is divisible and determine the level of real balances endogenously in pure-monetary equilibrium, given nominal balances, using the split-the-surplus condition discussed above. In terms of Figure 4, we need to shift the $M = M(X, \mu, 1)$ curve until the gains from trade for commodity traders and money traders are equalized. One can show that there exists a unique split-the-surplus equilibrium, and it implies a finite price level under appropriate parameter restrictions, as in Section II.

\(^{16}\)The rate of barter exchange is $\beta(1 - \mu)^2 x^2$, while the rate of monetary exchange is $\beta \mu (1 - \mu) x$. The ratio of these two is $(1 - \mu) x / \mu$, which vanishes as $x \to 0$ (note that $\mu$ is bounded below by $M$). For any finite $\beta$, however, there will always be some direct barter in equilibrium.
the contrary, it is because the economy has settled on the use of a generally acceptable currency that specialization becomes profitable, and it is specialization that inhibits barter.

IV. Dual Currency Regimes

In this section we take up the possibility of multiple fiat monies. It is motivated by the observation that, in some economies, there seems to be more than one type of currency in simultaneous circulation. For instance, it is possible in certain locations to have both a domestic currency and a foreign currency used in exchange, although perhaps the former is generally acceptable while the latter is only partially acceptable. One example is that Canadian dollars are often accepted just across the U.S. border, and vice versa, although the foreign currencies are not always accepted by domestic residents. Furthermore, this situation can persist even if the two currencies differ in terms of rates of return or other intrinsic properties.

In order to study the phenomenon of dual-currency equilibria, we assume that there are now two colors of fiat money: red and blue. To simplify the presentation as much as possible, we only consider the case in which specialization is exogenous, and we assume that both monies are indivisible. If we endow all agents with either one unit of red money, one unit of blue money, or one real commodity at the initial date, then all agents will always hold one and only one of these objects at all future dates as well. As above, the steady-state conditions can be solved for unique values of $\alpha$ and $\beta$, given strategies and exogenous values of $\Pi_R$ and $\Pi_B$.

Our goal is to construct an equilibrium in which both monies circulate, but with different acceptabilities: $1 = \alpha_R > \Pi_B > 0$. This requires $V_R > V_1 = V_B$. Now, $V_R > V_1$ follows immediately from $\Pi_R = 1$. Furthermore, for the case in which $y_R = y_B = 0$, (20)–(22) imply that $V_B = V_1$ if and only if

$$\Pi_B = \Delta x$$

where $\Delta = (r + \beta x_\mu_C + \beta x_\mu_R)/(r + \beta x_\mu_C + \beta x_\mu_R)$. Notice that $\Delta > 1$. If $\Pi_R = 1$ and $\Pi_B = \Delta x$, we have an equilibrium in which red money is universally accepted while blue money is only partially accepted. By continuity, we can perturb $y_R$ and $y_B$ without destroying the equilibrium, as long as $|y_R|$ and $|y_B|$ are not too great. In particular, we can construct equilibria with $1 = \Pi_R > \Pi_B$ even though $y_R < y_B$. In such an equilibrium, both monies circulate, but the high-return asset is less acceptable or less liquid.

17The way we write Bellman's equations implicitly assumes that agents never trade one currency for another, which is true in equilibrium because such a trade could not possibly make both agents better off.
than the low-return asset. That is, the reason why red money is universally acceptable, even though it is dominated in rate of return, is that it has liquidity value. If the spread \( Y_R - Y_B \) becomes too big, however, this equilibrium can no longer exist.\(^\text{18}\)

**VI. Conclusion**

We have presented a model of exchange in which the difficulty of pure barter leads to a transactions role for fiat currency, and we have used the model to address several issues in monetary economics.\(^\text{19}\) Other applications can also be studied in this framework. In Matsuyama et al. (1993), an explicit two-country version of a model similar to the one presented here is considered and used to investigate some issues relating to international monetary theory. In Steve Williamson and Wright (1991), a “lemons” problem is introduced into the model and used to illustrate the role of fiat currency in helping to overcome the frictions associated with private information. Siandra (1990) considers further the relationship between specialization and monetary exchange. Victor E. Li (1991) pursues some issues relating to externalities, welfare, and policy. Ramon Marimon et al. (1990) use a related model to analyze learning. Aiyagari and Wallace (1992) consider several other applications. Although there are many unanswered questions and much work remains to be done, we think that these search-theoretic models have definitely enhanced our understanding of the exchange process, in general, and of money, in particular.

\(^\text{18}\)Similar arguments can be used to show that there are equilibria in which both monies are universally acceptable even though one has a higher rate of return, and equilibria in which one money circulates but the other does not even though one or the other has a higher rate of return. Examples with two circulating fiat currencies have also been constructed by Aiyagari and Wallace (1992).

\(^\text{19}\)In order to focus on more substantive issues, we have neglected many of the technical aspects of search-based exchange models, like the possibility of multiple dynamic equilibria, including sunspot and cyclical equilibria (see Kiyotaki and Wright, 1990; Michele Boldrin et al., 1991; Timothy J. Kehoe et al. 1991).

**Appendix A**

Here we sketch a version of the model that makes it precise why agents need to trade, without assuming that they cannot consume their own output. This is perhaps more satisfying, but it does entail an increase in notation. The implications of this version of the model are essentially the same as those described in the text. For simplicity, we consider only the case where \( \alpha = \infty \), but it should be clear how to handle the more general model.

Suppose there are \( K \) types of agents, with equal numbers of each type, and \( K \) commodities, where \( K \geq 3 \). Agents are specialists in production but generalists in consumption, in the following sense. Each agent can produce only some of the commodities—to ease the presentation, suppose each type can produce exactly one commodity—but has a need to consume different things at different points in time. In particular, after consuming one commodity a consumer realizes a taste or need for a new commodity drawn at random. That is, the probability that the new commodity will be \( j \) is \( 1/K \) for any \( j = 1, 2, \ldots, K \). A consumer with a need for commodity \( j \) gets utility \( U \) from consuming it and no utility from anything else, at least until \( j \) is consumed and a new taste shock is realized.

Consider a representative agent. After consumption, with probability \( 1/K \) he needs the commodity he can produce, consumes immediately, and draws a new taste shock, while with probability \( (K - 1)/K \) he needs something else and must attempt to acquire it through trade. Therefore, the expected value of drawing a new taste shock, say \( V_n \), will satisfy \( V_n = (U + V_n)/K + V_n(K - 1)/K \). Eventually, our representative agent needs something that he cannot produce. When he meets a potential trading partner, the probability that this partner needs the good our representative agent produces is \( 1/K \), while the probability that this partner also has the good our representative agent needs is \( 1/(K - 1) \), since he must have one of the commodities other than the one that he himself needs. Hence, the probability of a double coincidence is \( 1/K(K - 1) \). This im-
plies that the value functions for commodity and money traders satisfy

\[
rv_1 = \frac{\beta(1-M)}{K(K-1)}(U-\varepsilon + V_n - V_1)
\]

\[
+ \frac{\beta M}{K} \pi(V_m - V_1)
\]

\[
rv_m = \frac{\beta(1-M)}{K} \Pi(U-\varepsilon + V_n - V_m).
\]

The rest of the analysis is the same as the model in the text.

**Appendix B**

Here we show how to relax the assumption of a zero transaction cost on accepting money, where, for simplicity, we again consider only the case where \(\alpha = \infty\). Let \(\eta\) denote the disutility cost of accepting currency, and assume that \(0 \leq \eta < U-\varepsilon\). The value functions in this generalized model satisfy

\[
rv_1 = \beta(1-M)x^2(U-\varepsilon)
\]

\[
+ \beta M x \pi(V_m - V_1 - \eta)
\]

\[
rV_m = \beta(1-M)x \Pi(U-\varepsilon + V_n - V_m).
\]

The only difference from the model in the text is that the gain to accepting money is now \(V_m - V_1 - \eta\). As before, there is always a nonmonetary equilibrium with \(\Pi = 0\). There is a pure-monetary equilibrium if and only if \(V_m - V_1 - \eta \geq 0\) when evaluated at \(\Pi = 1\). Manipulating the above equations implies that this is true if and only if \(\eta \leq \bar{\eta}\), where

\[
\bar{\eta} = \frac{\beta(1-M)x(1-x)(U-\varepsilon)}{r + \beta x (1-M)}.
\]

Finally, there is a mixed-monetary equilibrium if and only if \(V_m - V_1 - \eta = 0\) when evaluated at \(\Pi = \Pi^M \in (0,1)\). Manipulating the above equations implies that \(V_m - V_1 - \eta = 0\) if and only if

\[
\Pi^M = \frac{\beta(1-M)x^2(U-\varepsilon) + r \eta}{\beta(1-M)x(U-\varepsilon-\eta)}.
\]

Notice that \(\Pi^M > 0\) and \(\Pi^M < 1\) if and only if \(\eta < \bar{\eta}\).

Hence, there exists the same set of three equilibria for any \(\eta\) in \((0,\bar{\eta})\). In terms of Figure 2, an increase in \(\eta\) shifts the best-response correspondence to the right. For \(\eta < \bar{\eta}\) there are still three intersections with the 45\(^\circ\) line. For \(\eta > \bar{\eta}\) there is only one, at \(\Pi = 0\). Since \(\bar{\eta} > 0\), we can accommodate a positive transaction cost on money. Furthermore, \(\bar{\eta}\) can exceed \(\varepsilon\) (for example, as long as \(U\) is sufficiently large), and we can even accommodate a transaction cost on money that exceeds the transaction cost on real commodities. We can also use the split-the-surplus condition to determine \(M^*\) and \(P^* = M^*/C\) endogenously for any given stock of nominal balances \(C\). It can be shown that \(M^* > 0\), and therefore \(P^* < \infty\), if and only if \(\beta x(1-2x) - r > \eta(U-\varepsilon)/(r + \beta x)\), which generalizes the condition for a finite price level given immediately after (13).

**REFERENCES**


Cole, Harold L. and Stockman, Alan C., “Specialization, Transactions Technologies,


