A Simple Model of Money and Banking

by David Andolfatto and Ed Nosal

David Andolfatto is an associate professor of economics at Simon Fraser University. Ed Nosal is an economist at the Federal Reserve Bank of Cleveland. This article represents a simplified version of their work in progress, "Money and Banking with Costly State Verification." The authors thank Peter Rupert and Bruce Smith for their helpful comments and discussion.

Introduction

This article presents a simple environment giving rise to banks that create and lend out money. We define money to be any object that circulates widely as a means of payment. In our model, this object takes the form of a fully secured and redeemable bearer bond. This monetary instrument is issued by an agent that can credibly commit to monitoring a pool of real investments; that is, this capital forms the requisite backing for a circulating private debt instrument. While direct trade in securities is feasible without money, we find that money can economize on monitoring costs, which enhances the efficiency of the exchange process.

We define a bank as an agency that simultaneously issues money and monitors investments. In reality, banks also accept deposits of money, which are then redirected to borrowers. In our model, banks do not accept deposits; we do not view this function as a defining characteristic of a bank.[1] In particular, financial markets also accept deposits of money in exchange for marketable liabilities (equity and debt instruments). We think that banks differ from financial markets in two ways. First, bank liabilities are designed to be high-velocity payment instruments (money). Second, banks specialize in screening and monitoring their investments. Banks in our model perform both of these functions.

While our framework allows private non-bank liabilities to serve as the economy's medium of exchange (as mentioned earlier, exchange is even possible without any money at all), we demonstrate that the cost-minimizing structure has a bank creating liquid funds, which are then lent to borrowers (for example, entrepreneurs) with suitable collateral (contingent claims against future output). These liquid funds constitute real bills of exchange; that is, they are backed by the issuing bank with enforceable claims against real assets (the collateral supplied by borrowers).

1 Needless to say, there are those who would disagree with this point of view. To our knowledge, Bullard and Smith (2001) have the only model featuring intermediaries that simultaneously take deposits, make loans, and issue circulating liabilities.

In reality, the vast bulk of the money supply consists of private debt instruments with contractual features similar to those embedded in the debt instruments that circulate in our model. In addition, the bulk of this money is created by banks; that is, institutions that spend considerable resources monitoring their investment portfolio. Thus, our model goes some way in addressing the questions of why private money takes the contractual form it does, as well as why private money is typically supplied by banks (as opposed to other types of private agencies). In our model, money and banking are inextricably linked.

Of course, we are not the first to explicitly model money and banking together. Some important recent contributions include Kiyotaki and Moore (2000), Bullard and Smith (2001), and Cavalcanti (2001). We view our work as complementary to these papers. In Kiyotaki and Moore (2000), banks are agents endowed with some sort of commitment technology that allows their liabilities to circulate. In Cavalcanti (2001), banks have verifiable histories but non-banks have not; as in Kiyotaki and Moore (2000), this special feature of banks allows their liabilities to circulate. In Bullard and Smith (2001), the pattern in which agents meet endows bank liabilities with relatively low transactions costs, making these instruments the preferred medium of exchange. Our setup is similar to that of Diamond (1984), which emphasizes the role of monitoring in the business of banking; banks are endowed with no special characteristics relative to other agents in the economy. Under some circumstances, it makes sense to have the economy's monitoring agencies (banks) issue the medium of exchange.

I. A Simple Model of Money

The Physical Environment

Consider an economy with four periods, indexed by $t = 0, 1, 2, 3$. Period 0 is interpreted as a "contracting period" (no consumption or production takes place) where individuals may (or may not) meet to trade in securities. In subsequent periods, goods are produced and consumed, and spot markets may (or may not) open. The economy is populated by a large number ($3N$) of individuals who have preferences defined over deterministic, time-dated consumption profiles (c_1, c_2, c_3). Individuals are specialized in the production of a nonstorable time-period good (y_1, y_2, y_3). In particular, we assume that there are three types of individuals and N individuals of each type: Person A produces y_3; person B produces y_1; and person C produces y_2. One interpretation of this setup is that type A (C) individuals are endowed with a long- (short-) term capital project. Let us also assume (for simplicity only) that people have specialized preferences: Person A wants c_1; person B wants c_2; and person C wants c_3 (assume also that each person values his own good just a little bit). The chart below describes, for each individual type, the goods that he desires, c_i, and the good that he produces (is endowed with), y_j.

	A	B	C
Good 1	c_1	y_1	
Good 2		c_2	y_2
Good 3	y_3		c_3

Note the complete lack of double-coincidence of wants: Any bilateral pairing of individuals will result in no exchange of goods.

Trade occurs at a centralized location that is accessible by all agents in each period. In period 0, the only objects available for exchange are *claims* (contracts) against y_1, y_2, and y_3. In an environment where such claims can be costlessly exchanged and enforced, a market needs to open only once (in period 0). In this period, A sells a claim to y_3 and purchases a claim to y_1, B sells a claim on y_1 and purchases a claim on y_2, and C sells a claim on y_2 and purchases a claim on y_3.

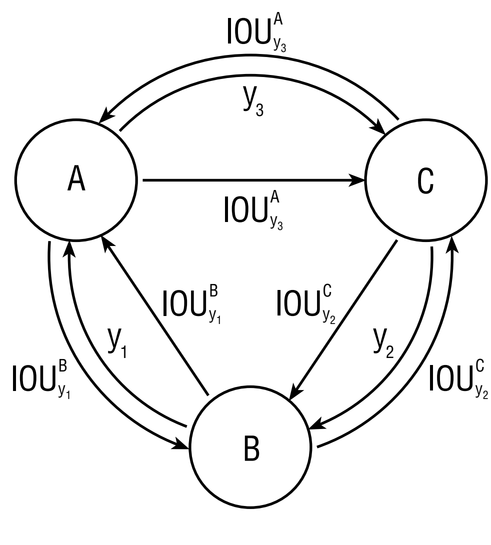

Figure 1

Figure 1 characterizes the various trades that occur in equilibrium. Period-0 trades are denoted by the inner straight-line arrows and the various time-dated trades are given by the outer curved-line arrows. We denote a claim issued by agent *j* for output produced at time *t* as $IOU^j\, y_t$. As time unfolds, previously agreed-to contracts are simply executed, that is, no further trades occur. This "Arrow–Debreu" market delivers an equilibrium allocation that is Pareto optimal without the aid of anything that one might identify as "money" in the model.

Limited Commitment and Monetary Exchange

Consider now an environment in which not all individuals can commit to keeping their promises. In particular, suppose that only type *A* agents can credibly commit to honoring claims against their anticipated earnings stream, y_3. In this case, the market value of both *B*'s and *C*'s securities as of period 0 equals zero (since these securities represent unenforceable claims against y_1 and y_2, respectively). At first blush, one might be inclined to think that financial markets could break down completely. After all, *B* (*C*) can acquire claims to y_2 (y_3) only by selling his claims to y_1 (y_2); but if these latter claims are worthless, then *B* (*C*) will be unable to purchase any claims to y_2 (y_3).

In fact, if spot markets open up after period 0, then the Pareto optimal allocation can be implemented with the following sequence of trades: In period 0, "nothing" happens. In period 1, agent *A* sells his claim y_3 to agent *B* in exchange for y_1. While agent *B* does not value y_3 directly, he is nevertheless willing to accept the security as payment, anticipating that he will be able to resell it in the future for something he does value. When the market reopens in period 2, agent *B* is in a position to purchase y_2 *directly* (instead of trying to collect on a previously negotiated *claim* to y_2; this purchase can be made with the security issued by agent *A*. Agent *C* is willing to accept the claim against agent *A*'s output because *C* values y_3 and the claim against y_3 can (by assumption) be enforced. Figure 2 summarizes the various trades.

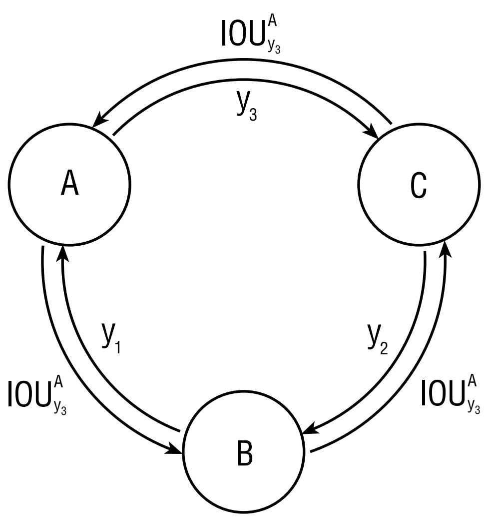

Figure 2

Notice what has happened here. In effect, agent *A* has issued a debt instrument entitling the *bearer* to the output generated in period 3 by agent *A*. Since agent *A* can commit to keeping his promises, this bearer-bond will circulate as a medium of exchange; in other words, agent *A*'s security can be properly identified as "money" in this model economy.

The money that arises in this model takes the form of a circulating private debt instrument, redeemable in some form of good or service. As such, it may appear somewhat removed from most modern-day monies, which primarily take the form of either government-issued fiat currencies or privately-issued debt instruments that are redeemable in government fiat (such as demand deposits). However, there are, in fact, several instances of privately-issued monies that are redeemable in goods or services. For example, the Canadian Tire Corporation has for many years issued small-denomination paper notes (referred to as *Canadian Tire money*) redeemable in a wide array of store products; in small communities, these notes have been known to circulate as a medium of exchange. The *Ithaca hour* is a privately-issued monetary instrument that circulates quite widely in Ithaca, New York; these notes, in various denominations, are meant to be redeemable in the labor services of local residents. As well, if one were to interpret y_3 as gold, then history offers innumerable examples of "gold-backed" monetary instruments.[2]

2 Some of the earliest forms of paper money possessed this feature. In the sixteenth century, for example, merchants would deposit their gold in vaults rented from goldsmiths. Apparently, the receipts issued by the goldsmiths (representing claims against the gold in the vault) began to circulate as a means of payment (primarily among merchants); see Smith (1936).}

The basic lesson here is that money is needed to facilitate some trades because not all individuals have the ability or willingness to commit to their promises (as John Moore has cleverly remarked, evil is the root of all money). The institutions that do arise to supply money will be those that have an ability, either endowed or manufactured, to make credible commitments. By issuing a debt instrument designed to circulate as a means of payment, the supplier of money is, in a sense, renting his commitment power to those who lack it. Specifically, even though individual B lacks commitment, he is able to purchase good y_2 by virtue of the fact that individual A can commit to promises.

One might legitimately ask where agent A's commitment power comes from and why others seem to lack it. The model of Kiyotaki and Moore (2000) provides some foundations to the structure of commitment power that depend on: 1) the existence of multiperiod investment projects; 2) the ability of initial creditors and debtors, if given the opportunity, to conspire against a third party who purchases existing debt; and 3) an institution, called a bank, that by design cannot conspire against anyone. In that environment, it is the bank's liability that circulates in the economy. Below, we provide a different foundation that is based on asymmetric information and monitoring activities, instead of asymmetric distribution of commitment power.

II. Money and Banking

In this section, we modify the physical environment described above in a few simple ways. To begin, assume that all individuals are identical in terms of their willingness and/or ability to commit to their promises and that commitment is feasible only up to what is verifiable (for example, through observation by a third-party enforcement agency).

Assume that there is now some risk associated with the endowment of each agent; in particular, for t = 1, 2, 3,

$$y_t = \begin{cases} y & \text{with probability } 1-\lambda \\ 0 & \text{with probability } \lambda. \end{cases}$$

In addition, suppose that there is no aggregate risk, so that $(1-\lambda)Ny$ represents the aggregate output in each period. We will continue to assume that individuals are risk-neutral.

The structure of information is as follows: Each person has the ability to costlessly observe the return realized on his own "project." Other agents are also in a position to observe this return, but only at a utility cost equal to μ; think of this cost as representing the effort exerted in monitoring project returns. This setup is similar to that of Diamond (1984), except that we will assume that if an agent is monitored, the information revealed becomes a matter of public record.[3]

Arrow–Debreu Securities

The type of securities that will be exchanged on this market are contracts that promise delivery of a good in the event that returns are reported to be positive. Since it will always be in the interest of the person who issues a security to report zero output (we assume that people cannot commit to tell the truth), it has to be understood that the holder of any such security will, in equilibrium, monitor project returns.

Clearly, for any kind of trade to occur, agents must have an incentive to purchase claims from other agents and then to monitor them. If the marginal utility of state-contingent consumption is constant (and equal to unity), then the parameter restriction $(1-\lambda)y > \mu$ is sufficient to guarantee that trade and monitoring will occur. As in the previous section, period 0 trades are as follows: A sells a contingent claim to y_3 and purchases one on y_1; B sells a contingent claim on y_1 and purchases one on y_2; and C sells a contingent claim on y_2 and purchases one on y_3. In period 1, agent A will monitor agent B; in period 2, agent B will monitor agent C; and in period 3, agent C will monitor agent A. Note that the total (economy-wide) monitoring costs are $3N\mu$. Figure 3 summarizes the various trades and monitoring.

3 This assumption is made primarily to simplify the exposition; it does not affect our main conclusions.

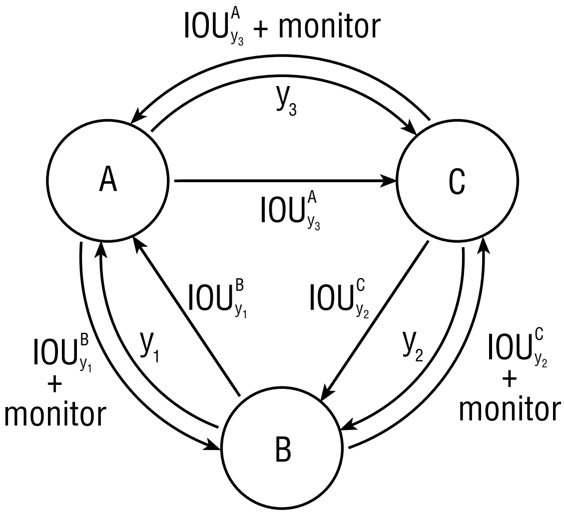

Figure 3

Monetary Exchange

Instead of trades in Arrow–Debreu securities, imagine that trades occur in a sequence of spot markets with the help of a circulating private debt instrument issued by type *A* agents. The private debt instrument issued by *A* is a contingent claim, as described earlier. Assume that project returns are realized at the beginning of each period and that a spot trade of money for goods occurs after the period's risk has been resolved.

The sequence of trades is as follows: In period 1, just after project returns are realized, type *A* agents offer their security to anyone who actually has y_1 to sell (as opposed to a contingent claim against y_1 that they would have purchased before realizing project returns). In the context of this environment, the people who are in a position to approach type *A* agents are the "successful" type *B* agents. It is important to note here that under this scenario, successful type *B* agents can costlessly reveal their success by the very act of displaying the goods they have to trade; in other words, *there is no need for type A agents to monitor.*

Successful type *B* agents are willing to accept a type *A* security as payment because they anticipate being able to use this security as payment for future goods that they desire. In particular, following the resolution of risk in period 2, a type *B* agent can purchase y_2 directly from a successful type *C* agent. Again, there is no need for monitoring. Type *C* agents willingly accept type *A* securities as payment because they represent direct claims against the goods that they desire. In period 3, each type *C* agent with a claim against y_3 will present the claim for redemption.

In order for any claim against *A* to be enforced, a monitoring expense must be incurred. Notice that while all *N* type *A* agents have issued securities, these securities end up being held by only $(1-\lambda)N$ type *C* agents. Consequently, each type *C* agent will hold claims for y_3 that were issued by different type *A* agents. To avoid coordination and monitoring problems, it makes sense here to appoint a "designated monitor," that is, to let one (arbitrarily chosen) type *C* agent set up a "monitoring business" that agrees to monitor a type *A* agent in exchange for some fraction ϕ of the project returns.[4] Since there are *N* projects that require monitoring (assume that projects cannot be monitored sequentially), the monitor incurs a total cost $N\mu$. Assuming free entry into the monitoring business, the equilibrium monitoring fee ϕ must adjust to ensure zero net returns to monitoring; thus $(1-\lambda)N\phi y = N\mu$ or

$$\phi^* = \frac{\mu}{(1-\lambda)y}$$

Under this "monetary regime," the expected utility payoff for agents *A* and *B* is equal to $(1-\lambda)y$, which clearly exceeds the payoff they would have generated under the Arrow–Debreu market structure: $(1-\lambda)y-\mu>0$. For each type *C* agent (including the monitor), the expected payoff is identical to what he would have generated under the Arrow–Debreu market structure. Consequently, we see that monetary exchange dominates trade in state-contingent securities by economizing on aggregate monitoring costs; that is, $N\mu < 3N\mu$. Figure 4 summarizes the various trades and monitoring. The *C* agent who does all of the monitoring is denoted by C^M. Note that in Figure 4, C^M receives ϕ^* from another *C* agent only if that *C*

■ **4** For values of λ less than half, the typical successful type *C* agent will hold one plus some fraction of type *A* IOUs. A type *C* agent will not have an incentive to monitor the type *A* agent for which he holds a fraction of an IOU if the fraction is sufficiently small: In this situation, the monitoring cost exceeds its expected benefit. If all type *C* agents who hold a fraction of an IOU issued by the same type *A* agent cannot somehow co-ordinate their monitoring activities, then this type *A* agent will not be monitored. But if a type *A* agent is not monitored, there will be a misallocation of resources. A designated monitor can overcome these coordination problems.

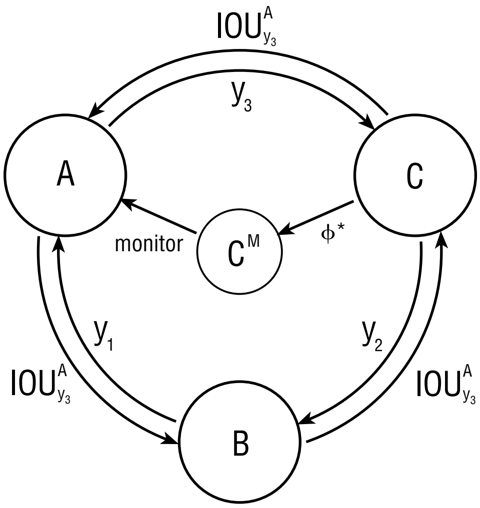

Figure 4

agent's claim pays off. In an attempt to reduce its complexity, figure 4 does not depict the potential exchange of y_3 for a claim against y_3 between agent C^M and an A agent.

Banking

The type of monitoring activity described above is commonly regarded as an important part of the business of banking. But banks are also associated with the business of creating liquidity (nowadays in the form of transaction deposits, but historically also in the form of paper money), which is injected into the economy by way of money loans (as well as wage and dividend payments). In the model described above, the money creation and monitoring activities are undertaken by separate sets of agents; in reality, these activities appear to be bundled. What might account for this bundling?

The first thing to note is that our model is not necessarily inconsistent with the fact that money creation and monitoring activities are bundled (although the model does not necessarily point to bundling as the unique organizational form either). We could, for example, imagine that the monitoring agent also decides to take on the responsibility of creating the economy's monetary instrument. In this case, such an agent would more closely resemble what is commonly called a bank. So let us consider what happens when the monitoring agent also issues money (for example, banknotes).

Trading activity proceeds as follows: In period 0, A agents approach the bank for a money loan. Suppose that each type A agent borrows (M/N) banknotes, which he promises to pay back in period 3 (if his project return is positive) with interest equal to R; that is, $(1+R)(M/N)$ represents the principal and interest that is to be paid back in the form of banknotes. Agent A then takes these banknotes and uses them to purchase y_1 from type B agents with output to sell. Since M "dollars" are exchanged for $(1-\lambda)Ny$ units of output, the first-period price level is

$$P_1^* = \left[\frac{1}{(1-\lambda)y}\right]\frac{M}{N}.$$

In period 2, each B agent with money purchases the output displayed for sale by successful type C agents; the equilibrium price level remains the same, that is, $P_2^* = P_1^*$.

Now, in period 3, type C agents who hold banknotes will want to purchase the output produced by successful type A agents. The question here is whether a type A agent is willing to give up a good that he values (slightly) in exchange for paper that he does not value at all. To give an A agent the incentive to behave "properly," the initial money-loan contract must contain a clause that transfers property rights over project returns from A to the bank in the event of default. In effect, the money loan is collateralized with securities that constitute contingent claims against y_3. As before, all type A agents must be monitored. Consequently, it will do no good for a successful type A agent to claim failure. At this stage, the agent has the choice of either selling his output for the banknotes that he needs to pay back the money loan; or of having his output "seized" by the bank (which owns an enforceable contingent claim against it). Either of these options leaves a type A agent with the same payoff, so the agent might rationally choose either one. A third possibility is that an A agent might renegotiate the terms of the loan contract, leaving both the bank and himself better off (at the expense of type C agents). To prevent either outright default or renegotiation (avoiding both would be necessary for the banknote to circulate in the first place), the monetary instrument must be include a redemption clause: The note-bearer must have the right to redeem the banknote for output.[5]

5 A similar redemption clause appears to be embedded within modern-day private monetary instruments (for example, deposits in checking accounts are typically redeemable in government cash).

Finally, if $R > 0$ (as must be the case), then it appears that there are not enough banknotes in circulation: How can A agents acquire the money needed to pay off a debt equal to $(1+R)M$ when there are only M dollars in circulation? It turns out that some "new money" (RM dollars) must be injected into the economy in period 3 by the bank itself. That is, the bank simply prints up RM dollars of new money, which can be used to purchase some period-3 output to compensate the bank for its monitoring services. With free entry in the banking business, the interest rate charged by the bank must result in zero profits; so

$$R^*M = P_3^* N\mu .$$

What prevents a bank from "overissuing" money at this stage? We have to assume that the supply of banknotes is verifiable (that is, the bank's balance sheet can be observed by a court and cannot be falsified). Consequently, the bank will be bound to charge a maximum interest of R^* and will not be in a legal position to inject more than R^*M of new money into the economy. Now, in order to derive R^*, we need an expression for the price level in period 3. Since $(1+R)M$ dollars are exchanged for $(1-\lambda)Ny$ units of output, the price level must satisfy the condition

$$P_3^* = \frac{(1+R^*)M}{(1-\lambda)Ny} .$$

Combining these latter two expressions, we can solve for the equilibrium interest rate and price level:

$$R^* = \frac{\mu}{(1-\lambda)y - \mu} > 0;$$

$$P_3^* = \left[\frac{1}{(1-\lambda)y - \mu}\right] \frac{M}{N} > P_1^* = P_2^* .$$

Notice that period-3 inflation (which is fully expected) has served to diminish the purchasing power of C's money holdings; but in equilibrium, this loss is exactly equal to the amount of purchasing power the type C agent would willingly have transferred to a professional monitor (as demonstrated in the earlier scenario). The various trades and monitoring are depicted in figure 5. The C agent who is the bank is denoted as C^B.

Transaction Costs

The equilibrium allocation associated with this "banking regime" corresponds to the allocation that resulted when money creation and monitoring activities were performed by separate sets of agents. Strictly speaking, the model here is unable to pin down the banking regime as a unique organizational form.

However, suppose that we follow Bullard and Smith (2001) and extend the model slightly by assuming that every time a good "changes hands," a small *fraction* of it, $0 < \varepsilon < 1$, disappears. One can think of ε as being a small transaction cost. Let us now compare the total transactions costs when money creation and monitoring are bundled relative to when they are not.

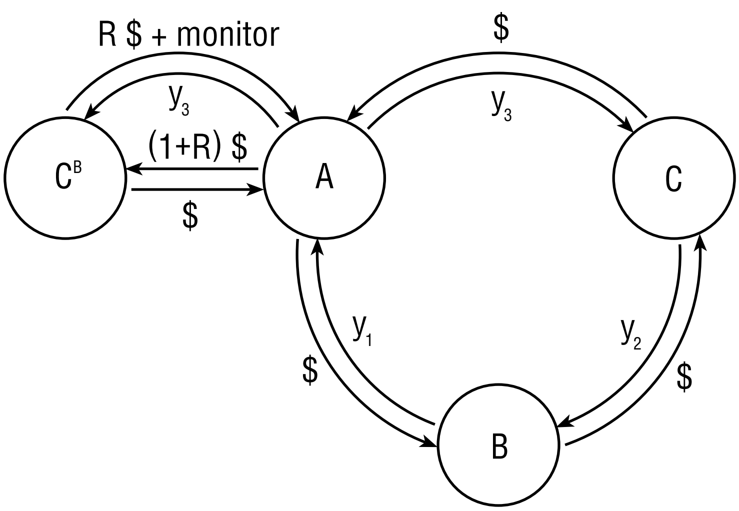

Figure 5

When money creation and monitoring are unbundled, each type A agent issues a security that circulates which is ultimately monitored by one of the type C agents. Goods produced in periods 1 and 2 change hands once, so the total transactions costs in these periods is $2(1-\lambda)N\varepsilon y$. At date 3, however, some of the goods produced will change hands twice: There is a set of transactions between C agents who possess A's security and A agents who produced output; there is also a set of transactions between C agents who now possess goods and the monitor who requires payment for his services. The transactions costs incurred in period 3 are $(1-\lambda)N(1+\phi^*)\,\varepsilon y$ if the monitor did not possess any securities issued by A agents and $(1-\lambda)N(1+\phi^*)\,\varepsilon y - \phi^*\varepsilon y$ if he did. Hence, total transactions costs for the economy are $3(1-\lambda)N\varepsilon y + N(1+\phi^*)\,\varepsilon y - \delta\,\phi^*\,\varepsilon y$, where $\delta = 0$ if the monitor did not possess any securities issued by A agents and $\delta = 1$ if he did (recall that whether the monitor ends up holding A's security depends on whether his project is successful).

In contrast, when money creation and monitoring are bundled, goods change hands only once at each date. In particular, at date 3 the bank's compensation for its monitoring costs takes the form of printing money and *directly* purchasing goods from type A agents. Hence, under this regime, total transaction costs are $3(1-\lambda)N\varepsilon y$, which is strictly less than the total transaction costs associated with the latter arrangement. Hence, when there are transaction costs associated with the exchange of goods, a banking regime—an institutional setup in which monitoring and money creation are undertaken by the same agent—will Pareto dominate an environment in which money creation and monitoring activities are performed by separate agents.

Conclusions

We have constructed an environment in which something that looks like a bank emerges as an efficient exchange mechanism. A bank monitors projects and issues money that circulates as a medium of exchange. When individual transactions are costly, the banking institution turns out to be an efficient trading mechanism because goods are only exchanged between the initial seller and the final consumer. An economy that has private (nonbank) securities circulating will have some fraction of final output exchanging hands more than once; as a result, its transaction costs will be higher than those of a banking economy.

References

Bullard, J., and B. Smith. "Intermediaries and Payments Instruments," unpublished manuscript, 2001.

Cavalcanti, R. "A Monetary Mechanism for Sharing Capital: Diamond and Dybvig Meet Kiyotaki and Wright," unpublished manuscript, 2001.

Diamond, D. "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, vol. 51 (1984), pp. 393–414.

Kiyotaki, N., and J. Moore. *"Inside Money and Liquidity,"* unpublished manuscript, London School of Economics, 2000.

Smith, V.C. *The Rationale of Central Banking.* Westminster, England: P.S. King and Son, Ltd., 1936.