A Model in which Outside and Inside Money are Essential

David C. Mills, Jr.

A MODEL IN WHICH OUTSIDE AND INSIDE MONEY ARE ESSENTIAL

Davⁱᴅ C. Mⁱˡˡˢ, Jʀ.

Federal Reserve Board of Governors

Abstract

 I present an environment for which both outside and inside money are essential as means of payment. The key model feature is that there is imperfect monitoring of issuers of inside money. I use a random matching model of money where some agents have private trading histories and others have trading histories that can be publicly observed only after a lag. I show via an example that for lags that are neither too long nor too short, there exist allocations that use both types of money that cannot be duplicated when only one type is used. Inside money provides liquidity that increases the frequency of trades, but incentive constraints restrict the amount of output that can be traded. Outside money is immune to such constraints and can trade for higher levels of output.

Keywords: Inside and outside money

1. INTRODUCTION

The fact that credit relationships among individuals are an alternative to money as a means of exchange provides a challenge to any model of money. If credit relationships are rich enough, money is not necessary to achieve good allocations. Thus, a model of money must contain sufficient frictions to establishing rich enough credit relationships for money to be necessary as a means of exchange. Necessary frictions, as demonstrated by Ostroy (1973), Townsend (1989), and Kocherlakota (1998), include a lack of commitment to future actions by individuals and some limitations on the ability to monitor individuals' trading histories.

Kocherlakota and Wallace (1998), hereafter KW, formalize these ideas in a random-matching model of money. They model limitations on the ability to monitor trading histories by introducing a random lag into the updating of those histories. They then study how the length of the lag affects optimal allocations and demonstrate that, in general, optimal allocations involve the use of (outside) money and some form of credit when the expected lag is sufficiently long enough.

Inside or private money is a particularly interesting type of credit instrument to study because both economists and policy makers have for a long time been interested in whether and how inside money should be regulated. This goes back at least to the nineteenth-century debates between the so-called currency school, which advocated a public monopoly on money issue with strict controls, and the free-banking

3

school, which promoted a relatively laissez-faire approach to private note issue. The fundamental concern about inside money is the incentive to overissue. This incentive arises from the fact that inside money is a credit instrument that the issuer promises to redeem in the future in exchange for something that is costly to give up. The incentive to overissue is one reason that Friedman (1959) favored 100% reserve requirements, essentially advocating the elimination of inside money. The recent removal of legal impediments to inside money issue and the advent of technologies that make electronic money feasible suggest that the question of whether and how to regulate inside money remains an important issue today.

Thus, to get at the questions concerning the regulation of inside money, one needs a model of money that provides a framework to compare alternative monetary regimes, and one that can explicitly model incentives to overissue. One such model is contained in Cavalcanti and Wallace (1999a,1999b), hereafter CW(a,b). To generate a role for both money and credit, they model limitations on the ability to monitor trading histories in the following way: some people (bankers) can be perfectly monitored via a recordkeeping technology whereas others (nonbankers) cannot be monitored at all because their trading histories are private information. CW(a) show that, because their trading histories are perfectly known, it is incentive feasible for bankers to issue and redeem their own notes as inside money (i.e., bankers can engage in credit relationships), which can then circulate

among the nonbankers (who have no access to credit because of their anonymity) as a means of exchange. CW(b) compares the use of outside money with that of inside money and find that inside money is superior because, in addition to duplicating allocations that are implementable with outside money, it implements allocations that increase the frequency of trades. This is because inside money provides instant liquidity to its issuers.

Good allocations in CW(a,b) require issuers of inside money to redeem (each other's) inside money. If the KW lag is introduced for them, then the longer that lag, the greater the temptation they face to defect from such redemption. If the lag is long enough then inside money cannot work. The lag, of course, does not limit the functioning of outside money. (The lag is infinite in models that have only outside money.) The new idea is to get a role for both inside and outside money. Inside money retains its role as providing liquidity to the issuer, but the lag limits how valuable it can be and, therefore, gives a role for outside money. The limited value of inside money is something like an endogenous credit limit (see Kehoe and Levine 1993).

Both KW and CW(a,b) assume that money is indivisible and that holdings are in the set $\{0, 1\}$. I maintain that assumption, but, in order to allow a potential role for both inside and outside money, consider worlds with two distinct monies: either two inside monies, two outside monies, or one of each. I show numerically that there exist background models with a lag that is neither too long nor too

5

short and allocations that can only be implemented using both types of money. That seems to suggest that for some background models, good allocations require both inside and outside money, but so far I have not demonstrated this.

The rest of the paper proceeds as follows. In Section 2., I describe the background matching model, which is essentially borrowed from earlier work. Then, as preliminary motivation, I introduce the updating lag for money issuers when there is one inside money in Section 3.. I describe a simple class of inside-money allocations and how the lag affects what is implementable. Then in Section 4., I introduce three alternative monetary arrangements, each with two distinct monies and each corresponding to one kind of monetary system: inside money only, outside money only, and both. I then present the numerical example that demonstrates that the use of both inside and outside money can be required for implementation. Section 5. concludes and the appendix contains a formal characterization of the alternative monetary arrangements.

2. THE MODEL

2.1. Background Environment and Sequence of Events

The background environment is the familiar random-matching environment from Shi (1995) and Trejos and Wright (1995). Time is dis-

crete and the horizon is infinite. There are S distinct, divisible, and perishable types of goods at each date and there is a $[0, 1]$ continuum of each of S specialization types of agents, where $S > 2$. An agent whose specialization type is s consumes only good s and produces only good $s + 1$ (modulo S), for $s = 1, 2, \ldots, S$. Each agent maximizes expected discounted utility with a discount factor $\beta \in (0, 1)$. Utility in a period is given by $u(c) - y$ where c is the amount consumed and y is the amount produced. The function u is defined on $[0, \infty)$, increasing, and twice differentiable and satisfies $u(0) = 0$, $u'' < 0$, and $u'(0) = \infty$. Moreover, there exists $y^{\max} > 0$ such that $u(y^{\max}) = y^{\max}$.

In each period, agents are randomly matched in pairs. As is familiar, the random matching along with agent specialization in both consumption and production means that there are no meetings in which there is a double coincidence of wants.[1] Rather, meetings are either single-coincidence meetings or no-coincidence meetings. A single-coincidence meeting is a meeting that contains a type s agent (the producer) and a type $s + 1$ agent (the consumer) for some s. A no-coincidence meeting is a meeting in which neither agent produces what the other consumes.

---

[1] A lack of double coincidence of wants is another necessary feature of an economy for money to be essential as a means of exchange.

2.2. Lack of Commitment and Imperfect Monitoring

As mentioned in the Introduction, assumptions regarding a lack of commitment to future actions by agents and some limitations on the ability to monitor agents' trading histories must be present in a model of money if money is to be essential as a means of exchange. In this model, I assume that agents cannot commit to future actions and I build on the specification of imperfect monitoring in CW(b)'s model of inside money.

As in CW(b), the society is able to keep a public record of the actions of a fraction B of each specialization type of agent, where $B \in [0, 1]$. Agents whose histories are a part of the public record are called *bankers*. Society has no public record for the remaining fraction $1 - B$ of each specialization type, the *nonbankers*.

In CW(b), society's ability to monitor bankers' trading histories is perfect. In this model, I allow for imperfect monitoring of bankers' trading histories. I assume that, similar to KW, the public record is not updated immediately after every action.[2] Specifically, there is a deterministic lag of T periods, where $T \geq 0$. At the beginning of each date $t > T$, the bankers' trading histories are known up through what

---

[2]In their model, everyone's history ($B = 1$) is updated with probability ρ at each date, which produces an average updating lag of $\frac{1}{\rho}$. There, that led to a slightly simpler formulation of the payoff from defecting than a deterministic lag. Here, the more straightforward deterministic lag is simpler.

they did until the beginning of date $t - T$. All more recent actions are private information to the banker.

When two agents meet, the following is common knowledge: each trading partner's specialization type, asset holdings (described below), information type (banker or nonbanker), and the past actions of the bankers in the meeting that occurred up to $t - T$ periods ago.

2.3. Assets

There are two distinct types of indivisible and perfectly durable assets: outside monies and inside monies.

An *outside money* is neither produced nor consumed. Each banker has a technology that enables her to create as many as two distinct objects called notes.[3] Because these notes are a type of credit instrument that may circulate as a means of payment among nonbankers, they may serve as *inside money*. The notes issued by a single agent are distinguishable from those issued by another agent, so that counterfeiting is not a problem.

---

[3] Because nonbankers lack commitment and are anonymous, it is not incentive feasible for them to issue inside money, so assuming that they do not have access to a printing technology is innocuous.

2.4.   Weakly Implementable Allocations and Welfare

An allocation describes what happens in all pairwise meetings, conditional on the states of the agents in a meeting. The states of bankers represent some combination of asset holdings and trading histories. The states of nonbankers, whose histories are private, can only represent asset holdings. Initial conditions on the distribution of assets and histories make such allocations sufficient to describe the evolution of the economy.

The following two-stage game is played in each pairwise meeting. In the first stage, agents simultaneously announce states. Given the information types of the agents and their announced states from the first stage and conditional on the absence of a discovered defection, the allocation suggests actions in the meeting. If there has been a discovered defection by either agent in the meeting, then no trade is always suggested. In the second stage, agents simultaneously decide whether to agree or disagree to the suggested allocation. If both agree, then the suggested allocation is carried out. If at least one agent disagrees, then both leave the meeting without trading.

A banker is a defector if she either misrepresents her state in the first stage, or chooses not to participate in the suggested allocation in the second stage. Suppose an initial defection occurs at date t. For the $T - 1$ periods that follow it, a banker is an undiscovered defector. From period $t + T$ on, a defecting banker is a discovered defector.

10

Defecting bankers are permanently punished with autarky.

DEFINITION 1. *An allocation is weakly implementable if there is a subgame perfect Nash equilibrium in which agents truthfully announce their states in the first stage and agree to the suggested actions for their meetings.*

Weakly implementable allocations can be characterized as those that satisfy a set of feasibility and incentive constraints. The feasibility constraints pertain to agents' abilities to transfer assets to one another (i.e., whether money can be issued and redeemed or not). The incentive constraints contain both truth-telling constraints about agents' states in the first stage of a meeting, and participation constraints at the second stage. The truth-telling constraints pertain to bankers' revealing their true state (which can represent both asset holdings and histories). Because asset holdings are common knowledge in a meeting and their trading histories are unknown, nonbankers cannot misrepresent their states at the first stage. Participation constraints imply that agents must prefer to accept the suggested allocation over not accepting it in the second stage of a meeting.

I want to focus, whenever possible, on good allocations. I define a simple ex ante representative agent welfare criterion—one that treats agents as identical before they are assigned their information types and states—to be the expected discounted utility of the gains from trade over all single-coincidence meetings. The gains from trade are denoted $z(y) \equiv u(y) - y$. Welfare is maximized by the production

11

and consumption of $y^* \equiv \arg\max_y [u(y) - y]$ in every single-coincidence meeting. The limited ability of the agents to make use of credit arrangements makes this welfare level impossible to obtain. Nonetheless, such a welfare level serves as a benchmark for comparison of alternative monetary arrangements.

3. IMPERFECT MONITORING AND THE VALUE OF INSIDE MONEY

To motivate how both inside and outside money may be essential when there is imperfect monitoring of trading histories, I first introduce the updating lag for bankers and show how ex ante welfare declines with the lag when there is only one inside money and no outside money. The example has asset holdings limited to the set $\{0, 1\}$ and allocations restricted to be both stationary and symmetric.

Consider an allocation in which the same output level, $y \in (0, y^*]$, is produced in all single-coincidence meetings except when (1) a nonbanker producer has a unit of inside money and when (2) a nonbanker consumer does not have a unit of inside money. The first exception is implied by the restriction on asset holdings and the nonbanker producer's participation constraint. The second is a feature of the allocation that implies that nonbanker consumers do not receive gifts.[4]

---

[4]Because banker histories are monitored, it may be implementable for banker producers to produce for (i.e., give gifts to) nonbanker consumers who do not have a unit of inside

In addition, suppose that nonbanker consumers surrender a unit of money when they consume y and that nonbanker producers receive a unit of money when they produce y. In single-coincidence meetings between nonbanker consumers and banker producers, the banker producer redeems the unit of inside money and destroys it. In single-coincidence meetings between banker consumers and nonbanker producers, the banker issues a unit of inside money. Because of the symmetry imposed on the allocation, meetings between bankers do not involve the transfer of inside money. Banker histories are used only for the purpose of punishing defecting bankers and do not influence allocations when there are no discovered defections. This makes the first stage of the two-stage game innocuous. Finally, I assume that half of the nonbankers within each specialization type start with a unit of inside money, whereas the other half do not. Under the above scheme, it is easy to verify that the distribution of money is a steady state distribution.

I can now express the value functions for both nonbankers and bankers that are implied by the above allocation. Let v_i^n denote the no-defection expected discounted utility of a nonbanker with asset holdings $i \in \{0, 1\}$ at the start of a period. The value functions for nonbankers are

$$S(1 - \beta)v_0^n = [B + \frac{1 - B}{2}][-y + \beta(v_1^n - v_0^n)] \tag{1}$$

money. Such allocations may improve welfare.

$$S(1-\beta)v_1^n = [B + \frac{1-B}{2}][u(y) - \beta(v_1^n - v_0^n)]. \qquad (2)$$

As is well known and easy to verify, these equations have a unique solution with $\{S(1-\beta) + [B+1]\beta\}[v_1^n - v_0^n] = [B + \frac{1-B}{2}][u(y) + y] > 0$. Now let v^b denote the no-defection expected discounted utility of a banker at the start of a period. The value function for a banker is

$$S(1-\beta)v^b = [B + \frac{1-B}{2}][u(y) - y]. \qquad (3)$$

Finally, I need to express the value functions for a defecting banker. What is of interest when it comes to expressing the incentive constraints for bankers is the initial-defector expected discounted utility, which I calculate recursively. If a banker chooses to become a defector, she agrees to consume in all single-coincidence meetings in which it is possible for her to consume (and so continues to issue inside money), but she disagrees to produce whenever production is possible (and so chooses not to redeem inside money). Let $\widetilde{v}_\tau^b$ denote the expected discounted utility of an undiscovered defector at time $\tau \in \{1, T-1\}$. Then

$$\widetilde{v}_\tau^b = \frac{1}{S}[B + \frac{1-B}{2}][u(y)] + \beta\widetilde{v}_{\tau+1}^b \qquad (4)$$

with the terminal condition that

$$\widetilde{v}_T^b = 0. \qquad (5)$$

The terminal condition incorporates the fact that once discovered, a

defecting banker is punished with autarky forever. The expected discounted utility for a banker from an initial defection given that no one else defects, $\beta \widetilde{v}_1^b$, is what is relevant for banker incentive constraints. This is obtained by solving $\widetilde{v}_\tau^b$ recursively from the terminal condition $\widetilde{v}_T^b$. For this allocation, the expected discounted utility from an initial defection is

$$
\begin{aligned}
\widetilde{v}_1^b &= 0 && \text{if } T = 0 \\
&= \frac{1}{S}[B + \frac{1-B}{2}][u(y)] \sum_{\tau=0}^{T-1} \beta^\tau && \text{if } T \geq 1.
\end{aligned} \tag{6}
$$

It is immediately obvious that $\widetilde{v}_1^b$ is increasing in T.

There are two relevant incentive constraints. One is a participation constraint for nonbanker producers,

$$
-y + \beta v_1^n \geq \beta v_0^n, \tag{7}
$$

and the other is a participation constraint for banker producers,

$$
-y + \beta v^b \geq \beta \widetilde{v}_1^b. \tag{8}
$$

Welfare from the inside money example, denoted $W(I)$, is

$$
W(I) = \frac{(1+B)^2}{4S(1-\beta)}[u(y) - y] \tag{9}
$$

and is maximized by the production of y^* for all single-coincidence meetings for which production takes place.

15

CW(a,b) has shown that y^* is implementable when $T = 0$ for sufficiently high β. This is because the right-hand side of (8) is zero, so that the only relevant participation constraint to satisfy is (7), which is nonbinding at y^* for high enough β.

I now present the main result of this section.

PROPOSITION 1. $W(I)$ *is weakly decreasing in* T.

Proof. Note that $W(I)$ is increasing in y for the range $y \in (0, y^*]$. Substitute (6) into (8) to get

$$y \leq \frac{(1 - \beta)\beta(B + 1)}{2S(1 - \beta) + \beta(B + 1)}[\frac{1}{1 - \beta} - \sum_{\tau=0}^{T-1} \beta^\tau]u(y). \qquad (10)$$

Because (6) is weakly increasing in T, the maximum value of y that satisfies (10) is weakly decreasing in T.[5]

Next, I demonstrate that at some point (8) replaces (7) as the relevant constraint for implementation, and that y^* cannot satisfy (8) for all T. Given that

$$\lim_{T \to \infty} \sum_{\tau=0}^{T-1} \beta^\tau u(y) = \frac{u(y)}{1 - \beta},$$

the only y that can satisfy (10) as $T \to \infty$ is zero, which is less than the y^* that satisfies the nonbanker participation constraint (7). This combined with the facts that $W(I)$ is increasing in y and the maximum value of y that satisfies (10) is weakly decreasing in T proves that $W(I)$ is weakly decreasing in T. ∎

---

[5]The maximum value of y that satisfies (10) is the value of y at equality.

The intuition for Proposition 1 is that increasing the updating lag for banker histories tightens the banker participation constraint (8) sufficiently so that the maximum value of y that can be implemented decreases. As T increases, the short-term incentive to agree to the trades that permit banker consumption (some of which involves note issue) and to defect from the trades that require banker production (some of which involves note redemption) increases. Therefore, to implement the above type of allocation, the common level of output must be decreasing in T and must approach 0 as T goes to infinity.

Although the proof of Proposition 1 is based on a particular type of inside money allocation, the result is more general. This particular inside money allocation provides the least disutility for bankers for a given y. For example, the allocation excludes gift-giving from bankers to nonbankers (see footnote 4). If an allocation included gift-giving, it would lower the value function for a nondefecting banker by increasing the probability that a banker has to produce y in equation (3), but would not reduce the expected discounted utility from an initial defection, leaving (6) unchanged. This has the effect of tightening the participation constraint for banker producers, (8). As a result, the critical T for which y^* cannot satisfy banker participation would be higher than that for an allocation without gift-giving by bankers.

17

3.1. Comparison with Outside Money

I can compare the above result to that obtained when there is only one outside money and no inside money. This example also has asset holdings limited to the set $\{0, 1\}$ and allocations are restricted to be both stationary and symmetric.

Now consider an allocation in which the same output level, $y \in (0, y^*]$, is produced in all single-coincidence meetings in which the consumer in the meeting has a unit of outside money and the producer does not. I assume that half of all agents within each specialization type start with a unit of outside money, whereas the other half do not. As with the inside money scheme, it is easy to verify that the distribution of money is a steady state distribution.

In contrast to the example involving inside money, trade takes place in strictly fewer single-coincidence meetings. This is because banker histories are completely ignored, so that there are no credit opportunities.[6] The benefit of such an allocation, however, is that it is independent of the updating lag.

The value functions for agents without money and with money

[6]For example, banker consumers in meetings with nonbanker producers must have a unit of outside money while their trading partner must not. All that was required in the inside money example was that their trading partners not have a unit of inside money because bankers could issue new notes.

are[7]

$$S(1 - \beta)v_0 = \tfrac{1}{2}[-y + \beta(v_1 - v_0)] \tag{11}$$

$$S(1 - \beta)v_1 = \tfrac{1}{2}[u(y) - \beta(v_1 - v_0)]. \tag{12}$$

As is well known and easy to verify, these equations have a unique solution with $\{S(1 - \beta) + \beta\}[v_1 - v_0] = \tfrac{1}{2}[u(y) + y] > 0$.

There is only one relevant constraint, a participation constraint for producers:

$$-y + \beta v_1 \geq \beta v_0. \tag{13}$$

Welfare for the outside money example, denoted $W(O)$, is

$$W(O) = \frac{1}{4S(1 - \beta)}[u(y) - y] \tag{14}$$

and, as with the previous example, is maximized by the production of y^* for all single-coincidence meetings in which trade takes place.

It is well known that for β sufficiently high, (13) is not binding and y^* is implementable. Because trade occurs less frequently in the example with outside money than it does in the example with inside money, $W(I) > W(0)$ if y^* is implementable (i.e., when T is small enough). This is essentially the result of CW(b) that inside money is strictly better than outside money under perfect monitoring of bankers. Inside money gives bankers liquidity when needed that

[7]Because banker identity is ignored in this example, I do not keep track of agents' information types.

19

increases the frequency of trades. When $T = 0$, monitoring is enough incentive to prevent overissue and have y^* implementable. While outside money can duplicate the y^* when β is sufficiently high, it cannot duplicate the frequency of trades because agents cannot issue outside money.

At some T, however, y^* cannot be implemented with inside money. This reduces the maximum attainable welfare using inside money. Outside money does not depend on the lag, so y^* can be implemented using outside money, no matter how large T is. Thus, there may be a range of updating lags for which it is optimal to use both outside and inside money in order to take advantage of each type's unique feature: inside money's ability to provide liquidity and outside money's ability to maintain a higher level of consumption.

4. ESSENTIALITY OF OUTSIDE AND INSIDE MONEY

In this section, I set up three monetary arrangements—one that uses two distinct inside monies, one that uses two distinct outside monies, and one that uses one inside and one outside money—and provide an example where the use of both inside and outside money is essential to implement certain allocations. As in the previous examples, this one maintains the assumption that asset holdings are limited to the set $\{0, 1\}$ and allocations are restricted to be both stationary and

symmetric. The outside-money arrangement and the inside-money arrangement use two assets, to provide comparable alternatives to the arrangement that uses both. Aiyagari et al. (1996) show that having two distinct assets may improve welfare in environments with indivisibility of assets and the upper bound on holdings because it increases the frequency of trades by allowing agents to exchange a higher-valued asset for a lower-valued asset and production. Modeling two distinct assets in the outside-money and inside-money arrangements removes the possibility that the essentiality of both types of money is tied to this feature.

For the example, there is an implicit function that maps asset holdings and histories into states, which are members of a three-element set $A \equiv \{0, 1, 2\}$. Because nonbankers have unobservable histories, their states can only represent asset holdings, where state 0 indicates no asset holdings, state 1 indicates holdings of a unit of asset 1, and state 2 indicates holdings of a unit of asset 2.

The state interpretation for bankers, however, is contingent on the monetary arrangement. In the outside-money arrangement, a banker's state describes asset holdings. Because each element of A is needed to represent asset holdings, the states do not represent a banker's history. For the inside-money arrangement, the states do not represent asset holdings of bankers, because bankers do not hold assets, due to the symmetric treatment of notes. Thus each state in A is tied to history. For the arrangement that uses both types of money, refer to asset 1

21

as the inside money and asset 2 as the outside money. If a banker is in state 2, that implies that she has a unit of outside money. If the banker is in either state 0 or 1, then she does not have a unit of outside money and the state can carry some history dependence.

In general, allocations for each monetary arrangement can be described in the following way. Let x_i^k denote the fraction of each specialization type with information type k in state i. Let the set of information types be $\{b, n\}$, where b indicates that an agent is a banker and n indicates that he is a nonbanker. Let $y_{ij}^{kl} \in \Re_+$ be output when a producer of information type k announces state i and a consumer of information type l announces state j. Similarly, let $p_{ij}^{kl}(h) = 1$ and $q_{ij}^{kl}(g) = 1$ indicate that the next period's states for the producer and the consumer are $h \in A$ and $g \in A$, respectively. For no-coincidence meetings, let $r_{ij}^{kl}(h) = 1$ indicate that $h \in A$ is the next period's state for the agent of information type k who announces state i in a no-coincidence meeting with an agent of information type l who announces state j.

For a given list of fractions of agents in states, an allocation is outcomes in single-coincidence meetings and outcomes in no-coincidence meetings,

$$[y_{ij}^{kl}, p_{ij}^{kl}(h), q_{ij}^{kl}(h), r_{ij}^{kl}(h)],$$

for all $i, j \in A$ and $k, l \in \{b, n\}$.

Recall that the alternative monetary arrangements are differentiated by the constraints that need to be satisfied for implementability.

The intuition for this follows. A formal derivation of these differences in the general problem is in the Appendix.

An environment that uses only inside money will have the fewest feasibility constraints because the issue and redemption of inside money adds liquidity that can lead to more trading.[8] In contrast, an environment that uses only outside money will have the most feasibility constraints because no one can issue or redeem outside money. An environment that uses both inside and outside money has an intermediate number of feasibility constraints.

For incentive constraints, the ranking of monetary arrangements in terms of number of constraints from most to least restrictive is reversed. This is due to the fact that a banker's outside-money holdings are observable, but that banker states tied to histories are observed only with a lag. In the inside-money arrangement, all banker states are tied to history, whereas in the outside-money arrangement, all banker states are tied to money holdings. Thus, bankers have the most opportunities to misrepresent their states in the inside-money arrangement, and so have the most truth-telling constraints, whereas bankers have the fewest opportunities to misrepresent their states in the outside-money arrangement, and so have no truth-telling constraints. The greater number of truth-telling constraints also increases the expected utility from making an initial defection, making par-

[8]Recall that the feasibility constraints pertain to agents' abilities to issue and redeem notes.

23

ticipation constraints tightest in the inside-money arrangement and weakest in the outside-money arrangement.

A comparison of various monetary arrangements involves comparing the sets of allocations that are weakly implementable via each monetary arrangement. This comparison is explicitly set out in Definitions 3–5 in the Appendix. These permit comparison of the types of allocations implementable under each arrangement.

DEFINITION 2. *The use of both inside and outside money is essential if there exists an allocation that is weakly implementable according to Definition 1 for a monetary arrangement that uses both inside and outside money, but is not implementable under a monetary arrangement that uses only inside money or only outside money.*

The definition of the essentiality of inside and outside money is weak in that it only requires that an allocation be weakly implementable exclusively in a monetary arrangement that uses both inside and outside money. A stronger definition would require that such an allocation be a good allocation, in the sense that it achieves a higher level of welfare than can be achieved by any other allocation that can be weakly implemented by the other monetary arrangements. Such a definition is not used here because the general problem makes it difficult to demonstrate the satisfaction of this stricter requirement.

Nonetheless, the results of the previous section suggest that for some background models there exist good allocations that can only be implemented via a monetary arrangement that uses both inside and

outside money. These background models are ones in which the lag is neither too short—otherwise the incentive constraints for bankers are not much of an issue and the economy would benefit from the exclusive use of inside money as in CW(a,b)—nor too long—otherwise the incentive constraints devalue inside money sufficiently so that the economy would benefit from the exclusive use of outside money.

PROPOSITION 2. *The set of allocations that satisfy Definition 2 is nonempty.*

The proof of Proposition 2 requires an example of an allocation that satisfies Definition 2. The following section describes such an example.

4.1. Proof of Proposition 2

Consider an allocation with three output levels, $y^{\mathrm{O}}, y^{\mathrm{I}}$ and y^{S}, where $y^{\mathrm{O}} > y^{\mathrm{I}} > y^{\mathrm{S}}$ and $y^{\mathrm{O}} \in (0, y^*]$. The output level y^{O} trades in single-coincidence meetings whenever the consumer has a unit of outside money and the producer does not have a unit of either money. Also, the consumer gives the producer the unit of outside money. The output level y^{I} is produced in several different types of single-coincidence meetings. It is produced in all single-coincidence meetings between bankers (with no assets changing hands) except for meetings involving the transfer of outside money mentioned above. It is also produced whenever inside money is exchanged. This includes meetings between nonbankers where the consumer has a unit of inside

25

money but the producer does not. It also includes meetings between nonbanker producers without an asset and banker consumers without outside money (inside money is issued to the nonbanker) and meetings between banker producers regardless of asset holdings and nonbanker consumers with a unit of inside money (inside money is redeemed by the banker). The output level y^{S} is exchanged in meetings between nonbankers where the producer has a unit of inside money and the consumer has a unit of outside money. In such a meeting, the agents swap assets as well.

The suggested actions are summarized in Table 1. The first element of each triplet in the box represents output in the meeting. The second element is the end-of-period state for the producer, and the third element is the end-of-period state for the consumer. Boxes with "nt" indicate that there is no trade.

A distribution of states that satisfy the steady-state conditions contained in the Appendix is[9]

$$x_0^n = 0.36; x_1^n = 0.18; x_2^n = 0.36; x_0^b = 0.0333; x_1^b = 0.0167; x_2^b = 0.05.$$

To satisfy Definition 2 such an allocation must be implementable using inside and outside money, but must not be implementable using only inside money or only outside money. It is immediate that such an allocation cannot be implemented with only outside money because there are meetings in which (inside) money is issued and redeemed.

[9]See conditions (A.1)–(A.4).

For example, nonbanker producers in state 0 (no asset holdings) can produce for a banker consumer in state 0 and leave the meeting in state 1. This violates feasibility constraints that must be satisfied for outside money; outside money cannot be issued.

What remains to be shown is that such an allocation is weakly implementable using both types of money, but is not implementable using only inside money. I demonstrate this numerically for the model specification

$$\{T, S, B, \beta, u(x)\} = \{65, 3, 0.1, 0.99, x^{\frac{1}{2}}\}$$

with output levels

$$y^{\text{O}} = 0.25$$
$$y^{\text{I}} = 0.20$$
$$y^{\text{S}} = 0.10.$$

Table 2 provides the relevant expected discounted utilities for the example, where v_i^k denotes the no-defection expected discounted utility of an agent of information type k who is in state i at the start of a period, and $\widetilde{v}_{i1}^b(I)$ and $\widetilde{v}_{i1}^b(M)$ denote the expected utility of an initial defection by a banker in state i under the inside money arrangement, I, and the arrangement that uses both types of money, M, respectively. First, notice that the expected utilities for nonbankers are increasing in states. Asset 2, which typically trades for higher levels

of output, is valued more than asset 1, which is more valued than holding no asset. The expected utilities for bankers are weakly increasing in states. Being in state 2 is preferable to being in state 1 or 0. Consider the expected utilities for defecting bankers with the inside money arrangement. They are the same for all states, because defecting bankers have the flexibility to represent the state that provides the highest period utility. For the arrangement using both types of money, the expected utility is the same for states 0 and 1 and less than state 2 because defecting bankers have the freedom to misrepresent themselves only when their true states are 0 and 1. These expected utilities are less than that for state 2 because state 2 typically commands higher levels of consumption relative to the others.

One can now use Table 2 to verify that the example allocation satisfies all of the constraints set forth in Definition 5 in the Appendix, and so is implementable under the monetary arrangement that uses both inside and outside money. Further, one can verify that the inside-money arrangement cannot implement the allocation because certain banker-producer participation constraints are violated. Specifically, in meetings with nonbanker consumers who enter with a unit of asset 1, banker producers in states 0 and 1 are not willing to redeem inside money for the amount of output called for by the allocation. Also, in meetings with banker-consumers, bankers in states 0 and 1 are not willing to produce for other bankers in states 0 or 1.

5. CONCLUSION

I present features of an environment for which both outside and inside money are essential as means of payment. The key model feature is that there is imperfect monitoring of issuers of inside money.

In deriving the results, I make use of the assumption that agents can only hold one unit of one asset at a time. I conjecture that this assumption is not crucial. This is because what makes both types of money essential is the aforementioned trade-off between outside and inside money. This trade-off would still exist if the assumption about the unit upper bound on money holdings were dropped.

An important issue not addressed in this paper is whether both types of money are essential in the stronger sense of improving welfare. Establishing this is difficult because of the large dimensionality of an allocation (there are 36 possibly distinct single-coincidence meetings) and the large number of complicated constraints.

REFERENCES

Aiyagari, S. R., N. Wallace, and R. Wright (1996) Coexistence of money and interest-bearing securities. *Journal of Monetary Economics* 37, 397–420.

Cavalcanti, R. and N. Wallace (1999a) A model of private bank-note issue. *Review of Economic Dynamics* 2, 104–136.

Cavalcanti, R. and N. Wallace (1999b) Inside and outside money as alternative media of exchange. *Journal of Money, Credit and Banking* 31 Part 2, 443–57.

Friedman, M. (1959) *A Program for Monetary Stability.* New York: Fordham University Press.

Kehoe, T. and D. Levine (1993) Debt-constrained asset markets. *The Review of Economic Studies* 60, 865–888.

Kocherlakota, N. (1998) Money is memory. *Journal of Economic Theory* 81, 232–251.

Kocherlakota, N. and N. Wallace (1998) Incomplete record-keeping and optimal payment arrangements. *Journal of Economic Theory* 81, 272–289.

Ostroy, J. (1973) The informational efficiency of monetary exchange. *The American Economic Review* 63, 597–610.

Shi, S. (1995) Money and prices: A model of search and bargaining. *Journal of Economic Theory* 67, 467–498.

Townsend, R. (1989) Currency and credit in a private information economy. *Journal of Political Economy* 97, 1323–1344.

Trejos, A. and R. Wright (1995) Search, bargaining, money and prices. *Journal of Political Economy* 103, 118–141

APPENDIX: FORMAL CHARACTER-IZATION OF ALTERNATIVE MON-ETARY ARRANGEMENTS

A.1. STEADY-STATE AND FEASIBILITY CON-STRAINTS

In describing the steady-state and feasibility requirements imposed on state transitions, I anticipate the satisfaction of one of the constraints described later: no free disposal of assets.

Because each person must be in one of the states, the fractions of each specialization type in each state must satisfy

$$\sum_i x_i^b = B \text{ and } \sum_i x_i^n = 1 - B. \tag{A.1}$$

Additionally, because each person can be in only one state at a particular point in time, state transitions must satisfy

$$p_{ij}^{kl}(h) = 1 \text{ if and only if } p_{ij}^{kl}(g) = 0 \text{ for all } g \neq h,$$

$$q_{ij}^{kl}(h) = 1 \text{ if and only if } q_{ij}^{kl}(g) = 0 \text{ for all } g \neq h,$$

$$r_{ij}^{kl}(h) = 1 \text{ if and only if } r_{ij}^{kl}(g) = 0 \text{ for all } g \neq h. \tag{A.2}$$

A steady-state distribution of agents over states requires that the fraction of bankers in each state and the fraction of nonbankers in

31

each state be constant. This can be expressed by equating the inflow and outflow of each state for nonbankers and bankers. These are, for each $i \in A$,

$$\sum_{h \neq i} x_h^n \sum_j x_j^b [p_{hj}^{nb}(i) + q_{jh}^{bn}(i) + (S-2)r_{hj}^{nb}(i)] = x_i^n \sum_j x_j^b [\sum_{h \neq i} p_{ij}^{nb}(h) + q_{ji}^{bn}(h) + (S-2)r_{ij}^{nb}(h)]$$

(A.3)

for nonbankers and

$$\sum_{h \neq i} x_h^b \sum_j x_j^n [p_{hj}^{bn}(i) + q_{jh}^{nb}(i) + (S-2)r_{hj}^{bn}(i)] = x_i^b \sum_j x_j^n [\sum_{h \neq i} p_{ij}^{bn}(h) + q_{ji}^{nb}(h) + (S-2)r_{ij}^{bn}(h)]$$

(A.4)

for bankers.

There are also feasibility constraints implied by the preservation of asset holdings in meetings. The general conditions can be written as

$$p_{ii}^{kl}(i) = q_{ii}^{kl}(i) = r_{ii}^{kl}(i) = 1, \qquad (A.5)$$

$$p_{ij}^{kl}(j) = 1 \text{ if and only if } q_{ij}^{kl}(i) = 1, \qquad (A.6)$$

$$r_{ij}^{kl}(j) = 1 \text{ if and only if } r_{ji}^{lk}(i) = 1. \qquad (A.7)$$

Condition (A.5) says that if both agents have the same asset holdings, then they leave with the same asset holdings. Conditions (A.6) and (A.7) say that if agent A's next-period state is agent B's current state, then B's next-period state is agent A's current state (they swap asset holdings). Somewhat weaker conditions than (A.6) and (A.7) that

will become relevant for the mixed mechanism are

$$
\begin{aligned}
p_{ij}^{kl}(j) &= 1 \text{ if and only if } q_{ij}^{kl}(j) = 0, \\
q_{ji}^{kl}(j) &= 1 \text{ if and only if } p_{ji}^{kl}(j) = 0, \\
r_{ij}^{kl}(j) &= 1 \text{ if and only if } r_{ji}^{lk}(j) = 0. \tag{A.8}
\end{aligned}
$$

These are weaker than (A.6) and (A.7) because for agent A to enter agent B's state, all that is required is that agent B leave his state (they do not have to swap states).

A.2. VALUE FUNCTIONS AND INCENTIVE CONSTRAINTS

I now describe a general set of participation, truth-telling, and free-disposal constraints. I begin by describing the expected discounted utilities of agents. These are all expressed given that no one else defects.

Let v_i^k denote the no-defection expected discounted utility of an agent of information type k who is in state i at the start of a period. For a single-coincidence meeting in which the producer is of information type k in state i and the consumer is of information type l in state j, let P_{ij}^{kl} and Q_{ij}^{kl} be producer and consumer payoffs, respectively, from following the suggested outcome in the second stage of a

meeting. Then

$$P_{ij}^{kl} = -y_{ij}^{kl} + \beta \sum_h p_{ij}^{kl}(h)v_h^k \tag{A.9}$$

and

$$Q_{ij}^{kl} = u(y_{ij}^{kl}) + \beta \sum_h q_{ij}^{kl}(h)v_h^k. \tag{A.10}$$

For no-coincidence meetings, let R_{ij}^{kl} be the payoff to the agent of information type k in state i when his partner is of information type l in state j from following the suggested outcome in the second stage of a meeting. Then

$$R_{ij}^{kl} = \beta \sum_h r_{ij}^{kl}(h)v_h^k. \tag{A.11}$$

Given these definitions, v_i^k is

$$v_i^k = \sum_{l \in \{b,n\}} \sum_j \frac{x_j^l}{S}[P_{ij}^{kl} + Q_{ji}^{lk} + (S-2)R_{ij}^{kl}]. \tag{A.12}$$

I calculate recursively the initial-defector expected discounted utility. Let $\widetilde{P}_{mj\tau}^{bl}$ and $\widetilde{Q}_{jm\tau}^{lb}$ be producer and consumer payoffs, respectively, to an undiscovered defecting banker who first defected τ periods ago and announced state m in the first stage of a single-coincidence meeting with a trading partner of information type l in state j from following the suggested outcome in the second stage. Then

$$\widetilde{P}_{mj\tau}^{bl} = -y_{mj}^{bl} + \beta \sum_h p_{mj}^{bl}(h)\widetilde{v}_{h,\tau+1}^b \tag{A.13}$$

34

and

$$\widetilde{Q}_{jm\tau}^{lb} = u(y_{jm}^{lb}) + \beta \sum_h q_{jm}^{lb}(h)\widetilde{v}_{h,\tau+1}^b. \qquad (A.14)$$

Similarly, define $\widetilde{R}_{mj\tau}^{bl}$ for no-coincidence meetings as

$$\widetilde{R}_{mj\tau}^{bl} = \beta \sum_h r_{mj}^{bl}(h)\widetilde{v}_{h,\tau+1}^b. \qquad (A.15)$$

Now consider a banker in the first stage of a meeting whose true state is i. The announced state of an undiscovered defecting banker depends on the ability of that banker to misrepresent her state, what is known about the state of her trading partner, whether she is in a single-coincidence or no-coincidence meeting, and in the case of a single-coincidence meeting, whether she is a producer or consumer.

Let $\widetilde{A} \subset A$ be the set of states over which a banker can misrepresent. For the outside-money mechanism $\widetilde{A} = \emptyset$, for the inside-money mechanism, $\widetilde{A} = A$, and for the mixed mechanism, $\widetilde{A} = \{0,1\}$. Let I_i be an indicator variable that equals 1 if $i \in \widetilde{A}$ and is 0 otherwise. Thus, if $I_i = 1$, then a banker can announce a state in $\widetilde{A}$, whereas if $I_i = 0$, then she is constrained to report truthfully. Similarly, let J_j^l be an indicator variable that equals 1 if $j \in \widetilde{A}$ and is 0 otherwise, where j is the true state of the banker's trading partner and l is her trading partner's information type.

Let $\mu \in \{p,q,r\}$ denote the type of meeting a defecting banker is in where p indicates that she is a producer in a single-coincidence meeting, q indicates that she is a consumer in a single-coincidence

35

meeting and r indicates that she is in a no-coincidence meeting. Then define $m_{ij}^{\mu l}(I_i, J_j^l)$ to be the optimal message of a defecting banker who is in meeting type μ, whose true state is i, and whose trading partner is of information type l in state j.

There are four types of optimal messages for a banker in a meeting of type μ: $m_{ij}^{\mu l}(0,0)$, $m_{ij}^{\mu l}(0,1)$, $m_{ij}^{\mu l}(1,0)$, and $m_{ij}^{\mu l}(1,1)$. Let $\widetilde{M} = \widetilde{P}_{mj\tau}^{bl}$ if $\mu = p$, $\widetilde{M} = \widetilde{Q}_{jm\tau}^{lb}$ if $\mu = q$, and $\widetilde{M} = \widetilde{R}_{mj\tau}^{bl}$ if $\mu = r$. Then the optimal messages are

$$
\begin{aligned}
m_{ij}^{\mu l}(0,0) &= m_{ij}^{\mu l}(0,1) = i, \\
m_{ij}^{\mu l}(1,0) &= \arg\max_{m} \widetilde{M}, \\
m_{ij}^{\mu l}(1,1) &= \arg\max_{m} \sum_{i \in \widetilde{A}} \frac{x_j^l}{\sum_{j \in \widetilde{A}} x_j^l} \widetilde{M}.
\end{aligned}
\tag{A.16}
$$

The first message reflects the fact that if a banker cannot misrepresent her type ($I_i = 0$) then she reports truthfully. The final two messages say that when given the freedom to misrepresent, a defecting banker chooses the state that gives the highest expected discounted utility; $m_{ij}^{\mu l}(1,0)$ indicates that the state of the trading partner is known with certainty, whereas $m_{ij}^{\mu l}(1,1)$ indicates that what is known is that the state of the trading partner is an element of $\widetilde{A}$.

Finally, consider a defecting banker in the second stage. The announcements concerning states have been revealed and now a defecting banker chooses whether to agree to the suggested outcome. Let $\widehat{P}_\tau^{bl}[m_{ij}^{pl}(I_i, J_j^l), j] = \widetilde{P}_{mj\tau}^{bl}$ such that $m = m_{ij}^{pl}(I_i, J_j^l)$. Similarly, define

$\widehat{Q}_\tau^{lb}[j, m_{ij}^{ql}(I_i, J_j^l)]$ and $\widehat{R}_\tau^{bl}[m_{ij}^{rl}(I_i, J_j^l), j]$.

Then for $i \in A$ and $\tau \in \{1, 2, \ldots, T-1\}$, the expected discounted utility of an undiscovered defecting banker is

$$
\begin{aligned}
\widetilde{v}_{i\tau}^b &= \sum_{l \in \{b,n\}} \sum_j \frac{x_j^l}{S} \{ \max[\widehat{P}_\tau^{bl}(m_{ij}^{pl}(I_i, J_j^l), j), \beta \widetilde{v}_{i,\tau+1}^b] \\
&\quad + \max[\widehat{Q}_\tau^{lb}(j, m_{ij}^{ql}(I_i, J_j^l)), \beta \widetilde{v}_{i,\tau+1}^b] \\
&\quad + (S-2) \max[\widehat{R}_\tau^{bl}(m_{ij}^{rl}(I_i, J_j^l), j), \beta \widetilde{v}_{i,\tau+1}^b] \} \qquad \text{(A.17)}
\end{aligned}
$$

with the terminal condition that

$$
\widetilde{v}_{iT}^b = 0 \qquad \text{(A.18)}
$$

for all $i \in A$.

Now consider the constraints that are relevant for implementation. Participation constraints require that agents are ex post sequentially rational. This is equivalent to the requirement that they receive non-negative gains from trade. For nonbankers, the participation constraints are

$$
\min\{P_{ij}^{nl}, Q_{ji}^{ln}, R_{ij}^{nl}\} \geq \beta v_i^n \qquad \text{(A.19)}
$$

for all $i, j \in A$ and $l \in \{b, n\}$. The right-hand side of (A.19) is due to the fact that defecting nonbankers can only be punished with no trade at that date because they will never be discovered.

For bankers, the participation constraints are

$$\min\{P_{ij}^{bl}, Q_{ji}^{lb}, R_{ij}^{bl}\} \geq \beta \tilde{v}_{i1}^b \qquad \text{(A.20)}$$

for all $i, j \in A$ and $l \in \{b, n\}$. The right-hand side of (A.20) reflects the fact that if a banker does not agree to the suggested outcome, she does not trade at that date and becomes an initial defector.

Now consider truth-telling constraints on bankers. Bankers must report their true state in the first stage of a meeting for all possible meetings:

$$m_{ij}^{\mu l}(I_i, J_j^l) = i \qquad \text{(A.21)}$$

for all $i, j \in A$, $l \in \{b, n\}$ and $\mu \in \{p, q, r\}$.

Finally, the free-disposal constraints for nonbankers and nondefecting bankers are

$$v_i^k \geq v_0^k \qquad \text{(A.22)}$$

For undiscovered defecting bankers, the free-disposal constraints are

$$\tilde{v}_{i\tau}^b \geq \tilde{v}_{0\tau}^b \qquad \text{(A.23)}$$

for all $\tau \in \{1, 2, \ldots, T - 1\}$.

A.3. DEFINITIONS

The outside-money mechanism has the largest number of feasibility restrictions because outside-money holdings must be preserved in all

meetings, whereas the inside-money mechanism has the smallest number of feasibility constraints because inside-money holdings must be preserved only in meetings between nonbankers.

The incentive constraints are most restrictive for the inside-money mechanism, and least restrictive for the outside-money mechanism. There are two reasons for the difference. The first is that the outside-money mechanism has, in effect, no truth-telling constraints, the mixed mechanism has truth-telling constraints only for some bankers, and the inside-money mechanism has truth-telling constraints for every banker. The second is that, because the different mechanisms have different numbers of truth-telling constraints, the value to a banker from making an initial defection, $\widetilde{v}_{i1}^{b}$, may also vary with the mechanisms. This leads to stricter participation constraints for bankers in mechanisms where $\widetilde{v}_{i1}^{b}$ is higher.

The following definitions formalize the conditions necessary for an allocation to be implementable via each mechanism.

DEFINITION 3. *An allocation is implementable via outside money if it satisfies (A.1)–(A.7) for $i,j \in A$, $k,l \in \{b,n\}$ and (A.9)–(A.23) for $i,j \in A$, $k,l \in \{b,n\}$, $\mu \in \{p,q,r\}$, $\tau \in \{1,2,...,T\}$, where $I_i = J_j^l = 0$ for all $i,j \in A$.*

DEFINITION 4. *An allocation is implementable via inside money if it satisfies (A.1)–(A.4) for $i,j \in A$, $k,l \in \{b,n\}$, (A.5)–(A.7) for $i,j \in A$, $k = l = n$, and (A.9)–(A.21) for $i,j \in A$, $k,l \in \{b,n\}$, $\mu \in \{p,q,r\}$, $\tau \in \{1,2,...,T\}$, where $J_j^n = 0$ and $I_i = J_j^b = 1$ for all*

39

$i, j \in A$, and (A.22) for $i \in A, k = n$.

DEFINITION 5. *An allocation is implementable via outside and inside money if it satisfies (A.1)–(A.4) for $i, j \in A$, $k, l \in \{b, n\}$, (A.5) for $i \in A, k = l = n$ and k and/or $l = b$ with $i = 2$, (A.6) and (A.7) for $i, j \in A, k = l = n$, (A.8) for $j = 2, i \neq 2$, and k and/or $l = b$, (A.9)–(A.21) for $i, j \in A$, $k, l \in \{b, n\}$, $\mu \in \{p, q, r\}$, $\tau \in \{1, 2, ..., T\}$, where $I_2 = J_j^n = J_2^b = 0$ and $I_i = J_j^b = 1$ for $i, j \in \{0, 1\}$, and (A.22) and (A.23) for $i \in A, k = n$ and $k = b$ with $i = 2$.*

TABLE 1. Trades in the example allocation

Consumer			Producer					
			Banker State			Nonbanker State		
			0	1	2	0	1	2
Banker State		0	$(y^I,0,0)$	$(y^I,1,0)$	$(y^I,2,1)$	$(y^I,1,0)$	nt	nt
		1	$(y^I,0,1)$	$(y^I,1,1)$	$(y^I,2,1)$	$(y^I,1,1)$	nt	nt
		2	$(y^O,2,0)$	$(y^O,2,0)$	$(y^I,2,2)$	$(y^O,2,0)$	nt	nt
Nonbanker State		0	nt	nt	nt	nt	nt	nt
		1	$(y^I,1,0)$	$(y^I,1,0)$	$(y^I,2,0)$	$(y^I,1,0)$	nt	nt
		2	$(y^O,2,0)$	$(y^O,2,0)$	nt	$(y^O,2,0)$	$(y^S,2,1)$	nt

TABLE 2. Value functions for the example allocation

State	v_i^n	v_i^b	$\widetilde{v}_{i1}^b(I)$	$\widetilde{v}_{i1}^b(M)$
0	2.432	3.742	3.595	3.253
1	2.712	3.742	3.595	3.253
2	2.821	3.877	3.595	3.303