Recent Developments in Modeling Financial Intermediation

Stephen D. Williamson
Economist
Research Department
Federal Reserve Bank of Minneapolis

Financial intermediaries—such as banks, savings and loan institutions, and insurance companies—play a large and significant role in highly developed economies. In the United States, for example, the finance, insurance, and real estate sector accounted for about 15 percent of the real gross national product in 1986. The comparable figure for agriculture, forestry, and fisheries was less than 3 percent and for manufacturing, nearly 22 percent.[1] The economic significance of financial intermediaries results from their special role: making contractual arrangements that link borrowers and lenders more efficiently than if these agents had to trade directly.

Despite the size and significance of the intermediation sector, only recently have economists attempted to build models explicit enough to address such basic questions as

- Why do intermediaries exist?
- What makes them special?
- How do they interact with the rest of the economy?

In this paper I use recent advances in the theory of financial intermediation to construct a model that can address these questions. In this model, financial intermediaries arise endogenously in equilibrium: intermediaries exist because they economize on the costs of monitoring borrowers in a world where relevant information about specific borrowers is not freely available to lenders. The intermediaries of the model perform special functions similar to those performed by real-world intermediaries: they borrow from and lend to a large number of economic agents, they process information about investment outcomes, they issue liabilities with characteristics different from their assets, and they write debt contracts.

Having constructed the model, I then give an example of how it can be used to tell us more about business cycle phenomena and about the interplay between financial intermediation and the macroeconomy. The example shows how endogenous changes in the costs of intermediation can amplify business cycle fluctuations. The example also illustrates how this model can reproduce—as other business cycle models cannot—observed qualitative comovements between aggregate bankruptcies and other economic time series. In addition, the model predicts timing patterns between a nominal monetary aggregate and real output—patterns consistent with empirical observation.

Before presenting the model in detail, I briefly review developments in the theory of financial intermediation. This review allows us to place the model within the context of other approaches—old and new—to modeling intermediation.

Approaches to Modeling Intermediation

Newer approaches to modeling financial intermediation can be viewed as reactions to older micro- and

[1] These percentages are based on data from the Commerce Department's *Survey of Current Business*, July 1987, p. 57, Table 6.2.

macroeconomic approaches to the problem. Compared with older micro approaches, recent theories of intermediation are much more explicit about what makes intermediaries special. Compared with older macro approaches, recent theories tend to be less concerned with the role of money in intermediation.

Toward Greater Explicitness

In studying the microeconomic behavior of financial intermediaries, older approaches simply assumed the existence of intermediary firms. (See Baltensperger 1980 for a survey of the older literature.) For example, if we were to follow such an approach in modeling a bank, we would start by specifying a technology for producing outputs (different types of loans) from inputs (different types of deposits). As a result, a bank ends up looking much like any other firm; little in this approach distinguishes banking from, say, the production of toothbrushes. But what makes financial intermediation potentially worthy of study are its special functions (such as diversification, information processing, and asset transformation). We cannot expect to generate these special activities or derive many useful implications if our approach does not build on the economic features that cause financial intermediaries to arise in the first place.

More recent work on the theory of financial intermediation generally follows the route of first specifying an explicit economic environment (economic agents' preferences, their endowments, and the technology available to them), next showing that financial intermediation is a necessary element of an equilibrium in that environment, and then studying the features of the equilibrium. The environments in these models contain some sort of *private information;* that is, some economic agents know more about their own activities, preferences, endowments, or production technologies than do other agents.

One group of newer financial intermediation models, which I call *Type 1,* focuses specifically on banking, deposit contracts, and bank runs. Type 1 models attempt to rigorously justify the conventional wisdom that views banking as inherently unstable (and therefore as requiring special regulatory intervention). Of these models, the most well known is Diamond and Dybvig's (1983). (For related work, see Haubrich and King 1983, Smith 1984, Chari and Jagannathan 1987, and Bhattacharya and Gale 1987.)

In the Diamond–Dybvig model, banks provide insurance against an uncertain demand for liquidity. However, for any demand deposit contract supplying such insurance, a bad equilibrium exists where all agents run to the bank to withdraw their deposits; as a result, production is disrupted. In their model, this situation can be improved in some circumstances if the bank suspends withdrawals and in others if the government provides deposit insurance. Thus, Diamond and Dybvig construct an explicit environment where banking is in some sense essential, show what problems can develop, and evaluate the results of particular government interventions. Although their results about policy interventions are subject to debate (see, for example, Williamson, forthcoming *b*), the explicit nature of the approach makes it more likely that the debate will be conducted at a fruitful level.

In a second group of recent models, which I call *Type 2,* financial intermediation serves to economize on the costs of acquiring information. (It is this type of model that I present later in the paper.) Type 2 models focus on more general kinds of intermediary structures, rather than on banking in particular. An important example of this type appears in Boyd and Prescott 1986. In that paper, the intermediaries that arise are organizations which write sets of rules for compensating intermediary members (residual claimants) and nonmembers (depositors and borrowers) and for evaluating and funding borrowers' investment projects. In this way, the costly evaluation of investment projects and their funding occur more efficiently than if projects were evaluated and funded through decentralized markets.

In another Type 2 model (Diamond 1984), borrowers know more about the realized outcomes of their investment projects than do lenders. An arrangement with a diversified financial intermediary permits individual lenders to delegate the responsibility for monitoring borrowers to the intermediary and thereby to economize on the costs of monitoring investment outcomes. A similar delegated monitoring role for intermediaries is presented in Williamson 1986, which builds on the costly state verification approach in Townsend 1979. In Williamson's model, the contractual arrangements between intermediaries and their borrowers capture elements of those observed in the real world; that is, intermediaries write debt contracts and in some circumstances may ration credit.

Away From Money

In traditional macroeconomic approaches, financial intermediation has been regarded as important only because some intermediaries produce money. This view is reflected in the content of undergraduate money

and banking courses, where banking is introduced into the study of macroeconomic issues by examining the *money multiplier process* (by which an increase in base money gets *multiplied* by a fractional-reserve banking system). The idea that intermediation does not merit study by macroeconomists unless the intermediary in question issues liabilities having the characteristics of money is consistent with the monetarist views of Milton Friedman. (See, for example, Friedman 1960, 1969; Friedman and Schwartz 1963.)

Particularly illustrative of the monetarist approach is Friedman and Schwartz's (1963) interpretation of how bank failures contributed to the severity of the Great Depression. They argued that an increase in the incidence of bank failures caused the public to prefer holding currency rather than bank deposits. The money multiplier thus fell and, since the Federal Reserve did not conduct offsetting open market operations, the money supply dropped. According to Friedman and Schwartz, what mattered for the pace of economic activity was not the direct effects of the bank failures, but the fact that the money supply dropped. That is, in their analysis, the intermediation process itself is irrelevant for aggregate activity; what is relevant is how disturbances to the system affect monetary assets.

In the newer approaches to studying financial intermediation, the monetary nature of an intermediary's liabilities is of little consequence. For example, in the Diamond–Dybvig model cited earlier (Diamond and Dybvig 1983), bank runs arise in an environment in which there is no asset that could be identified as money. (Also, for the model presented later in this paper, it matters little that intermediary deposits are in some ways similar to currency.)

This element of the newer approaches was foreshadowed in some of James Tobin's writings (Tobin 1961, 1963; Tobin and Brainard 1963). However, his treatment of the intermediation process in macroeconomic analysis (for example, Tobin 1969) paralleled the earlier microeconomic studies of intermediary firm behavior, mentioned previously. That is, the existence of intermediaries is simply assumed, and their behavior is treated as identical to that of other economic agents. In Tobin's analysis, an intermediary is specified in terms of its assets and liabilities, ad hoc supply and demand functions for each of these assets and liabilities, and a balance sheet constraint. Since this analysis does not address the fundamentals of what makes financial intermediaries special, it cannot enlighten us much about the role intermediation plays in macroeconomic activity.

The approach to macroeconomics embodied in modern equilibrium business cycle theory virtually ignores financial intermediation. The monetary theory of the business cycle presented in Lucas 1972, and later popularized in Barro 1976, is consistent with Friedman's monetarist view that intermediation is unimportant. The real business cycle models of Kydland and Prescott (1982) and Long and Plosser (1983) go even further by abstracting from monetary factors as well as financial intermediation. (For an update on Kydland–Prescott, see also Prescott 1986.) This approach leaves observed money–output correlations unexplained, but King and Plosser (1984) have indicated how these correlations might be reproduced in a real business cycle model, due to the endogenous response of the banking sector to real disturbances. However, their model treats banking symmetrically with other types of economic activity, with intermediaries performing only one essential function, the production of transaction services.

A Dynamic General Equilibrium Model With Financial Intermediation

Having reviewed some theoretical trends, I now construct a Type 2 dynamic general equilibrium model of financial intermediation. Although the model resembles one in an earlier paper (Williamson, forthcoming *a*), it has been simplified to make my main points more accessible. The model illustrates how newer approaches capture the special functions of real-world intermediaries in explicitly specified economic environments. And by embedding the model in a dynamic framework, I show how such an approach can be used to confront macroeconomic issues.

The model's fundamental structure is that of an overlapping generations (OLG) model with borrowing and lending, and is similar to OLG models used by Wallace (1980) and Sargent and Wallace (1982). However, it differs from theirs in that, in equilibrium, one-period borrowing and lending is carried out through an intermediary structure, rather than by borrowers and lenders directly. The OLG apparatus is particularly convenient in this instance, since it easily accommodates a Type 2 financial intermediation model (resembling the one in Williamson 1986) within a dynamic framework. The model economy's inhabitants and their behavior are now described.

Lenders and Entrepreneurs

In each period t, where $t = 1, 2, 3, \ldots$, there are N agents born, each of whom lives for two periods. A fraction α

of these agents are *lenders* and the remainder $(1-\alpha)$ are *entrepreneurs.* Lenders born at time t are each endowed with e units of the time t consumption good in the first period of life. These lenders consume at t and at $t+1$ and maximize expected utility over their lifetime. That is, each lender maximizes

$$(1) \qquad E_t(\ln c_t + c_{t+1})$$

where E_t is the expectation operator conditioned on time t information, c_t is consumption in period t, and the lender's utility function is

$$(2) \qquad f(c_t, c_{t+1}) = \ln c_t + c_{t+1}.$$

To consume in the second period of life, lenders either acquire fiat money in the first period or exchange their endowment for some claim to the second period's consumption good. The nature of these claims is determined in what follows.

At $t=1$, a group of old lenders are collectively endowed with H units of fiat money. These lenders sell all this endowment of money so as to maximize consumption at $t=1$.

Entrepreneurs born at time t receive no endowment in either period of life and maximize $E_t c_{t+1}$. Each entrepreneur has access to an investment project that needs K units of the consumption good to operate. Because K is assumed to be much larger than e, it takes the endowments of a large number of lenders to finance the investment project of a single entrepreneur. If an investment project is funded at time t, it is either successful with probability π, yielding a return of w_t at $t+1$, or unsuccessful with probability $1-\pi$, giving a return of zero, where $0<\pi<1$. Entrepreneurs' returns are independently distributed, and entrepreneurs in a given generation have different π values. Thus, each entrepreneur has a *name,* or π value, and entrepreneurs' names in a given generation are uniformly distributed over the interval $[\pi_t^l, \pi_t^u]$. Mean project returns are therefore uniformly distributed over the interval $[\pi_t^l w_t, \pi_t^u w_t]$.

If a given project is funded in period t, only the entrepreneur who operates it can observe its return without incurring some cost. Any other agent can observe the return, but in doing so incurs a *monitoring cost* of β units of the consumption good. Thus, there is *costly state verification* (as in Townsend 1979). That is, it is costly for others to learn about the success or failure of a particular venture, but costless for the entrepreneur who actually operates the project (though all are equally informed before the project is funded about its potential payoffs and likelihood of success).

At the beginning of time t, agents know all details of the time t environment, but there is uncertainty about the returns to lending and to holding fiat money. Given that all agents are risk neutral [note that the utility of entrepreneurs is linear in c_{t+1}, as is that of lenders in equation (1)], all that is relevant for lenders' saving decisions is the highest available expected return on assets. In equilibrium, all assets held, including fiat money, bear the same expected return r_t. That is,

$$(3) \qquad r_t = (E_t p_{t+1})/p_t$$

where p_t is the price of fiat money in terms of the consumption good (and $1/p_t$ is the price level). If s_t is a lender's saving, then from (1) and assuming that $s_t > 0$, utility maximization implies

$$(4) \qquad s_t = e - (p_t/E_t p_{t+1}).$$

That is, lenders will save more if expected returns are high and less if expected returns are low.

If entrepreneurs in this environment were *pathologically honest* (they never lie, even if it were in their interest to do so), then there would be no need for monitoring investment outcomes, and the problem would be uninteresting. However, agents in the model are assumed to be *rational* (attentive to their own interests). Thus, without monitoring, an entrepreneur whose project was funded at time t and was successful would adopt the following utility-maximizing strategy: report at $t+1$ that the project was unsuccessful and then consume the entire return. In this situation we have what is known as a *moral hazard* problem. That is, the actions of an entrepreneur may affect the lenders' payoffs. Knowing this, lenders would not fund any projects if monitoring never occurred. Therefore, for projects to be funded, contracts must be written to provide for monitoring under some contingencies. They will be written in a way that gives entrepreneurs the incentive to truthfully report returns while allowing lenders to economize on monitoring costs.

Suppose that m lenders, indexed by $j = 1, 2, \ldots, m$, fund a particular entrepreneur's project in period t. Payments can be made to these lenders in $t+1$ only if the project is successful. Let x_j denote the payment to lender j if the project is successful, where total payments cannot exceed the entrepreneur's return ($\Sigma x_j \leq w_t$). Also assume that a given lender cannot observe what the entrepreneur pays the other lenders.

At $t+1$ the entrepreneur can make either of two declarations: the project succeeded or the project failed. At least one of these declarations must induce lender j to monitor, since otherwise the entrepreneur would always declare failure.[2] Clearly, a lender's optimal strategy is to monitor only if the entrepreneur declares the project failed.[3] Given this monitoring strategy, the entrepreneur always reports the true outcome to all lenders.

The total expected return to all m lenders, when each lends directly to the entrepreneur, is the expected value of total payments minus expected monitoring costs; that is,

$$(5) \qquad R_0^l = \pi \sum_{j=1}^{m} x_j - (1-\pi)\beta m.$$

The expected return to the entrepreneur is the expected project return minus expected payments to the lenders, or

$$(6) \qquad R_0^l = \pi \left(w_t - \sum_{j=1}^{m} x_j\right).$$

It seems clear that there must be a way to improve on this arrangement, since in the case of a bad outcome all lenders monitor and acquire the *same* information. Suppose, then, that the group of m lenders delegates the task of monitoring to one group member. If this monitoring agent always reports the true outcome to the group, the entrepreneur could receive an expected return of R_0^l, as in (6), while the group of lenders gets expected returns higher than in (5):

$$(7) \qquad R_1^l = \pi \sum_{j=1}^{m} x_j - (1-\pi)\beta > R_0^l.$$

However, the moral hazard problem crops up again because the monitoring agent does not have an incentive to report the truth. If the project succeeds, the monitoring agent could declare the outcome was bad and collude with the entrepreneur to split the payments that would have gone to the other lenders. Therefore, this arrangement does not work.

Intermediaries

An arrangement that optimally delegates the monitoring role does exist, however, and it shares most of the special functions associated with financial intermediation. In this arrangement, the financial intermediary is an individual lender (a monitoring agent) who lends to many entrepreneurs and borrows from many lenders. By lending to and borrowing from many agents, the intermediary exploits the *law of large numbers.* (As an example of this law, take the case of coin tosses. If we flipped a fair coin for a large enough number of times, then the fraction of flips that were heads would be arbitrarily close to one-half.) Thus a financial intermediary in this model can perfectly predict the fractions of entrepreneurs with good and bad outcomes. The intermediary can therefore commit to making fixed payments to its depositors (lenders), but without making the payments hinge on the outcomes of entrepreneurs' projects.[4] In equilibrium, each depositor receives a certain return at $t+1$ of r_t per unit deposited in period t. Because free entry into intermediation implies that no profit opportunities remain unexploited in equilibrium, intermediaries earn zero profits. Thus, an entrepreneur with probability π of a good outcome makes a payment $v_t(\pi)$ to the financial intermediary in the event of a good outcome. This payment is set to give the intermediary an expected return of $r_t K$ on the loan. That is,

$$(8) \qquad \pi v_t(\pi) - (1-\pi)\beta = r_t K.$$

Also, the entrepreneur's payment cannot exceed the project's return:

$$(9) \qquad v_t(\pi) \leq w_t.$$

Thus, under this optimal arrangement, all monitoring is delegated to the intermediary, and depositors need never incur any costs in verifying outcomes.

The financial intermediaries modeled here perform several of the special functions that characterize real-world intermediaries:

- *Intermediaries borrow from and lend to large numbers of economic agents.* Here, this diversification is critical to the functions performed by an intermediary, since it gives the intermediary the correct incentives as a monitoring agent and eliminates the moral hazard problem.

---

[2] I rule out stochastic monitoring by assumption. With stochastic monitoring, monitoring would occur with some prespecified probability, where this probability depends on the entrepreneur's particular declaration. This doesn't matter much for the analysis here, but it would make the equilibrium contract more complicated.

[3] For the purpose of this argument, I assume that each of the m lenders has access to enough of the consumption good (by holding some fiat money) in $t+1$ to cover the cost of monitoring. This assumption, however, is irrelevant for the analysis that follows.

[4] To be more precise, it is sufficient for this result to let the number of borrowers n grow large while the number of intermediaries k grows large, where $k = n^{\delta}$ and $0 < \delta < \frac{1}{2}$.

- *The characteristics of assets held by intermediaries differ from those of their liabilities.* Here, intermediary assets are loans, each of which pays off only in the event of a good outcome for the entrepreneur. Deposits are noncontingent liabilities that pay a fixed amount regardless of project outcomes.
- *Financial intermediaries process information.* Here, an intermediary obtains information in monitoring its borrowers, and its depositors need not concern themselves about monitoring.
- *Financial intermediaries, particularly banks, tend to rely on debt instruments in writing contracts with their borrowers.* Here, intermediary assets have several features associated with debt: a borrower makes a fixed payment for good outcomes and consumes zero (is *bankrupt*) in the event of a bad outcome. The monitoring cost β can be interpreted as a cost of bankruptcy. Debt contracts can also be obtained with more general probability distributions for the project return. The general result is that the entrepreneur consumes the project return minus a fixed payment when there is no monitoring, or consumes zero if monitoring occurs (see Williamson 1986).[5]

General Equilibrium

Given the expected return on assets r_t, it is generally the case, from equations (8) and (9), that some entrepreneurs will not receive loans. That is, even if such an entrepreneur had received a loan and paid w_t with a good outcome, the expected payment to the intermediary would be less than r_tK. Entrepreneurs not receiving loans are those with a low probability of success and a high expected cost of monitoring. Given (8) and (9), we can determine the cutoff point for lending by intermediaries. That is, there is some point π_t^* such that entrepreneurs with $\pi \geq \pi_t^*$ get loans and those with $\pi < \pi_t^*$ do not. From (8) and (9) we get

$$(10) \quad \pi_t^* = (r_tK+\beta)/(w_t+\beta).$$

Given equation (10), the quantity of loans made by intermediaries is

$$(11) \quad L_t = (1-\alpha)NK(\pi_t^u-\pi_t^*)/(\pi_t^u-\pi_t^l).$$

The stock of fiat money is fixed for all t, with H units issued to old lenders at $t=1$. In equilibrium, lenders save by holding fiat money or deposits with intermediaries. Therefore, since deposits are equal to loans,

$$(12) \quad \alpha NS_t = p_tH + L_t.$$

It remains to describe the nature of the stochastic disturbances that will drive the business cycle in the model. Attention will be confined to disturbances affecting investment technologies, that is, to shocks affecting the success probabilities and returns of projects.[6] Thus, given a specification of how π_t^l, π_t^u, and w_t evolve, equations (4), (10), (11), and (12) can be used to determine a rational expectations solution for the evolution of p_t, L_t, π_t^*, and s_t. A simple and tractable formulation is to assume that the state of the world, denoted by z_t, follows a two-state *Markov process.* That is, z_t equals either z_1 or z_2, and the future evolution of z_t depends only on its current value. We thus have

$$(13) \quad Pr[z_{t+1}=z_1|z_t=z_i] = q_i$$

where q_i is the probability of getting state 1 in the next period, given that the current state is i; $0 < q_i < 1$; $i = 1, 2$; and $q_1 > q_2$. Therefore, the probability of getting the same state in consecutive periods is higher than that of getting different states (that is, z_t is *positively serially correlated*). If $z_t = z_i$, then $\pi_t^l = \pi_i^l$, $\pi_t^u = \pi_i^u$, $w_t = w_i$, for $i = 1, 2$. A rational expectations equilibrium is then (p_i, L_i, π_i^*, s_i) for $i = 1, 2$, where $(p_t, L_t, \pi_t^*, s_t) = (p_i, L_i, \pi_i^*, s_i)$ if $z_t = z_i$. That is, variables depend only on the state of the world in equilibrium. Thus, from (13), we have

$$(14) \quad E_t[p_{t+1}|z_t=z_i] = q_ip_1 + (1-q_i)p_2.$$

Therefore, substituting using (4), (10), and (11) in (12) and using (14), the equilibrium prices for fiat money are determined by

$$(15) \quad \alpha N\{e-[q_1+(1-q_1)p_2/p_1]^{-1}\}$$
$$= [(1-\alpha)NK/(\pi_1^u-\pi_1^l)]$$
$$\times(\pi_1^u-\{[q_1+(1-q_1)p_2/p_1]K+\beta\}/(w_1+\beta))$$
$$+ p_1H$$

[5] Unfortunately, allowing for risk aversion or stochastic monitoring or both generally results in a contract that does not have these simple features (see Townsend 1987).

[6] Using this model, it would be straightforward to examine an equilibrium with shocks to the quantity of fiat money. But I choose to focus here on technology shocks because this helps relate the approach to the real business cycle literature and better illustrates what we can learn from the approach.

and

$$(16)\quad \alpha N\{e-[q_2p_1/p_2+(1-q_2)]^{-1}\}$$
$$=[(1-\alpha)NK/(\pi_2^u-\pi_2^l)]$$
$$\times(\pi_2^u-\{[q_2p_1/p_2+(1-q_2)]K+\beta\}/(w_2+\beta))$$
$$+p_2H.$$

In principle, (15) and (16) solve for p_1 and p_2, and (4), (10), and (11) then solve for s_t, π_t^*, and L_t. In addition, output at $t+1$, denoted y_{t+1}, is the sum of output produced from projects financed in period t and the endowments of lenders in period $t+1$, or

$$(17)\quad y_{t+1}=[(1-\alpha)N/(\pi_t^u-\pi_t^l)]$$
$$\times\int_{\pi_t^*}^{\pi_t^u}[\pi w_t-(1-\pi)\beta]d\pi+\alpha Ne.$$

Let b_{t+1} denote the number of bankruptcies at $t+1$ (that is, projects financed at t that get a bad outcome). Then

$$(18)\quad b_{t+1}=[(1-\alpha)N/(\pi_t^u-\pi_t^l)]\int_{\pi_t^*}^{\pi_t^u}(1-\pi)d\pi.$$

An Example of the Model at Work

I now present a numerical example that illustrates some of the model's properties and allows me to make my main points. For the example, parameter values (listed in Table 1) were chosen to produce fluctuations in output on the order of what might be observed in the real world and to ensure that lenders hold a relatively small part of their wealth as currency. Also, the shocks to investment project returns and success probabilities were set up so that $\pi_1^u w_1=\pi_2^u w_2$ and $\pi_1^l w_1=\pi_2^l w_2$. That is, whether $z_t=1$ or $z_t=2$, mean project returns are uniformly distributed over the same interval. The technology shocks thus do not affect production opportunities in terms of mean returns, but they do affect the riskiness of each entrepreneur's project. For example, if we use variance σ^2 as a measure of the riskiness of an investment project with an expected return R, then $\sigma^2=w_tR-R^2$, and risk increases as w_t increases.

Table 1

Parameter Values for the Example

Variable	Value	Variable	Value
N	1,000,000	π_2^u	.7
α	.98	π_2^l	.08
e	2	w_1	400
H	1	w_2	514
K	50	q_1	.6
π_1^u	.9	q_2	.4
π_1^l	.1		

Results

An equilibrium was computed for each of three values for monitoring costs: $\beta=0$, 40, and 80. The results are presented in Table 2. (Note that in the table, y_i and b_i are the level of output and the number of bankruptcies when $z_{t-1}=i$. Also, γ_i is the fraction of entrepreneurs who receive loans when $z_t=i$. Shading indicates positive monitoring costs.)

For the case where monitoring costs are zero ($\beta=0$), note that no fluctuations occur.[7] That is, since all agents are risk neutral and stochastic disturbances do not affect mean returns, production opportunities do not vary across states of the world when project returns can be verified without cost. In contrast, for the cases where monitoring is costly ($\beta=40$ and $\beta=80$), prices and quantities fluctuate. Note also that if $\beta=0$, then financial intermediation is inessential; borrowing and lending could be carried out through direct transactions between lenders and entrepreneurs. Thus, the ability of the model to produce fluctuations in response to shocks to the riskiness of investment project technologies is integrally related to the role for financial intermediation which arises here.

When monitoring costs are positive, fluctuations occur in the following manner. When the state of the world changes from 1 to 2, the primary effect of this disturbance is an increase in the average entrepreneur's probability of experiencing a bad outcome. Since monitoring occurs only if a project fails, the higher average probability of a bad outcome increases the expected cost of monitoring for the average borrower and thus increases the average cost of financial intermediation. Therefore, fewer loans are made in state 2 than in state 1 ($L_2<L_1$), and the demand for fiat money is higher in state 2, resulting in a higher price of fiat money ($p_2>p_1$) and thus a lower price level. Output is

[7] When $\beta=0$, the cutoff point for lending to entrepreneurs π_t^* fluctuates (as seen in Table 2). However, the size and characteristics of the group of entrepreneurs who receive loans are identical in states 1 and 2.

Table 2
Equilibria Computed for Different Monitoring Costs

Variable (subscripts denote state)	Monitoring Costs* $\beta = 0$	$\beta = 40$	$\beta = 80$
Price of Fiat Money			
p_1	11,250	113,020	198,867
p_2	11,250	113,610	201,240
Saving per Lender			
s_1	1.00000	1.00209	1.00479
s_2	1.00000	.99791	.99521
Lending Cutoff			
π_1^*	.12500	.20478	.27133
π_2^*	.09722	.16218	.21834
Fraction of Projects Financed			
γ_1	.96875	.86903	.78584
γ_2	.96875	.86434	.77408
Total Loans			
L_1	968,750	869,030	785,840
L_2	968,750	864,340	774,080
Output (when $z_{t-1} = i$)			
y_1	5,931,875	5,489,138	5,120,940
y_2	5,931,875	5,399,171	4,946,091
Bankruptcies (when $z_{t-1} = i$)			
b_1	n.a.	7,780	6,512
b_2	n.a.	9,835	8,373

n.a. = not applicable
*Shading indicates positive monitoring costs; the darker the shade, the higher the costs.

lower in period t if $z_{t-1} = 2$ than if $z_{t-1} = 1$, since fewer investment projects are funded at $t-1$ in the first case.

When monitoring costs are positive, the number of bankruptcies is affected in three different ways. First, given the nature of technological disturbances, the average probability of bankruptcy is higher if $z_t = 2$ than if $z_t = 1$; when $z_t = 2$, the number of bankruptcies tends to be higher at $t+1$. Second, fewer projects are funded if $z_t = 2$ ($\gamma_1 > \gamma_2$), and this tends to make the number of bankruptcies lower at $t+1$. Third, projects with a given mean return which go unfunded if $z_t = 2$ but which *would* be funded if $z_t = 1$ are those with the highest probability of going bankrupt among projects which get loans in state 1. That is, the marginal projects that get filtered out when $z_t = 2$ are the most likely to go bankrupt. This third effect tends to make the number of bankruptcies higher at $t+1$ if $z_t = 1$ than if $z_t = 2$. However, the first effect more than offsets the last two, so that $b_2 > b_1$.

Interpretations and Insights

A benefit of this explicit approach to modeling financial intermediation is the insight gained about the interaction between production technologies and the inter-

mediation process. In the example, the business cycle arises due to the effect of technology shocks on the costs of intermediation. That is, an increase in the probability of a bad return for all entrepreneurs, if mean returns are held constant, increases the magnitude of the incentive problem, since there is a higher probability that an intermediary must monitor a particular entrepreneur. Note that doubling the size of the monitoring cost from 40 to 80 magnifies fluctuations, as shown in Table 3. Therefore, the more important financial intermediation is in saving on resource costs (the higher β is), the greater the magnitude of fluctuations produced as the result of disturbances to project risk.[8]

Note that the monetary nature of an intermediary's liabilities is irrelevant to most of the analysis here. In contrast to the monetarist view of intermediation discussed earlier, it does not particularly matter for the role played by intermediaries that the deposits they issue might appear similar in some respects to fiat money.

For the most part, the business cycles generated by the model for the cases $\beta = 40$ and $\beta = 80$ fit the stylized facts of macroeconomic behavior. Note that since there is some persistence in the technology shocks (that is, $z_t = i$ tends to be followed by $z_{t+1} = i$, for $i = 1, 2$), there will also be persistence in output. Therefore, high output tends to be associated with a high quantity of loans and a low number of bankruptcies, as we observe in time series data. Also, prices tend to be high when output is high.[9]

The model also generates predictions about timing patterns among economic time series. For example, suppose we were to aggregate fiat money and intermediary deposits and call the sum (in nominal terms) *money,* denoted by M_t. We then have

$$(19) \quad M_t = \alpha N s_t / p_t.$$

If we compute the equilibrium values for money in this example as in Table 2, we get the following results:

Variable	$\beta = 40$	$\beta = 80$
M_1	8.68910	4.95152
M_2	8.60790	4.84648

Thus, periods with low real output are always preceded by periods with low nominal money. If we ran this model to generate long time series of money and output (as shown in Williamson, forthcoming *a*), we would find that in the statistical sense, money *causes* output. That is, upward movements in money tend to precede upward movements in output. However, this statistical relationship tells us nothing about the effects of changes in the stock of outside money on output. For example, a one-time unanticipated change in the quantity of fiat money, accomplished through transfers to old agents, would be *neutral;* it would increase the price level proportionally with no effect on the real allocation.

It should be noted that some of the time series correlations this model generates, such as the positive correlation between prices and output and the timing between money and output, could also be reproduced in models where financial intermediation is inessential. For example, the idea that observed money–output correlations can be explained by the endogenous response of inside money to exogenous disturbances is not new. This idea has been exposited in various forms by Tobin (1970), Sargent and Wallace (1982), King and Plosser (1984), and Freeman (1986)—all without the aid of an explicit model of financial intermediation. However, this model reproduces other phenomena that alternative models would have difficulty explaining, particularly the negative correlation between bankruptcies and output.

Table 3

Higher Monitoring Costs Magnify Business Cycle Fluctuations
(coefficients of variation*)

Variable	Monitoring Costs** $\beta = 40$	$\beta = 80$
Price of Fiat Money (p_t)	.0026	.0059
Total Loans (L_t)	.0027	.0075
Output (y_{t+1})	.0083	.0174
Bankruptcies (b_{t+1})	.1167	.1250

*Coefficients of variation have been computed unconditionally.
**Darker shading for higher monitoring costs.

[8]It can be shown more generally, by using calculus, that an increase in β increases the size of fluctuations.

[9]There is some debate as to how prices move over a typical business cycle; some, such as Lucas (1977), accept a positive correlation between the price level and output as a stylized fact, while Prescott (1983) finds a negative correlation in postwar detrended U.S. data.

Concluding Remarks

This paper has reviewed some recent advances in the theory of financial intermediation and then shown, by way of an example, what can be learned through the explicit modeling of financial intermediaries. The example shows how, in a business cycle driven by disturbances to the riskiness of investment projects, fluctuations are amplified as intermediaries' monitoring costs increase. In the example, an interplay exists between production technologies and the intermediation sector; fluctuations result from the effects of technology shocks on intermediation costs. The example shows how the model can generate qualitative comovements, similar to those observed empirically, among aggregate time series.

Due to its tractability, this model (or a more general one, such as presented in Williamson, forthcoming *a*), could be useful in examining other issues in macroeconomics. For instance, the model might lend insight into understanding particular historical events. Real business cycle models of the type studied by Prescott (1986) are capable of accurately reproducing the business cycle phenomena in postwar U.S. data, but such models have more difficulty explaining events like the Great Depression. Some authors, for example Bernanke (1983), argue that the severity of the Depression was in part due to a reduction in the pace of the credit allocation process. The analysis here is suggestive as to how this might have occurred, and Bernanke (1983) and Bernanke and Gertler (1986) have studied other possible mechanisms. Also, in principle, there seems to be no reason why the model couldn't be calibrated, as are the real business cycle models of Prescott (1986), to see how well it quantitatively replicates business cycle phenomena.

Much remains to be done in the theory of financial intermediation. Current research topics include the study of long-term financial relationships and the modeling of economies where some financial contractual arrangements involve intermediation while others do not. In addition, there are important unresolved questions concerning the role of government in the financial sector, and whether or not the special nature of financial intermediation implies market failures that special regulations could correct.

References

Baltensperger, Ernst. 1980. Alternative approaches to the theory of the banking firm. *Journal of Monetary Economics* 6 (January): 1–37.

Barro, Robert J. 1976. Rational expectations and the role of monetary policy. *Journal of Monetary Economics* 2 (January): 1–32.

Bernanke, Ben S. 1983. Nonmonetary effects of the financial crisis in the propagation of the Great Depression. *American Economic Review* 73 (June): 257–76.

Bernanke, Ben [S.], and Gertler, Mark. 1986. Agency costs, collateral, and business fluctuations. Working Paper 2015. National Bureau of Economic Research.

Bhattacharya, Sudipto, and Gale, Douglas. 1987. Preference shocks, liquidity, and central bank policy. In *New approaches to monetary economics*, ed. William A. Barnett and Kenneth J. Singleton. Proceedings of the Second International Symposium in Economic Theory and Econometrics, pp. 69–88. New York: Cambridge University Press.

Boyd, John H., and Prescott, Edward C. 1986. Financial intermediary-coalitions. *Journal of Economic Theory* 38 (April): 211–32.

Chari, V. V., and Jagannathan, Ravi. 1987. Banking panics, information, and rational expectations equilibrium. Unpublished manuscript, Federal Reserve Bank of Minneapolis and Northwestern University.

Diamond, Douglas W. 1984. Financial intermediation and delegated monitoring. *Review of Economic Studies* 51 (July): 393–414.

Diamond, Douglas W., and Dybvig, Philip H. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91 (June): 401–19.

Freeman, Scott. 1986. Inside money, monetary contractions, and welfare. *Canadian Journal of Economics* 19 (February): 87–98.

Friedman, Milton. 1960. *A program for monetary stability*. New York: Fordham University Press.

———. 1969. *The optimum quantity of money and other essays*. Chicago: Aldine Publishing.

Friedman, Milton, and Schwartz, Anna Jacobson. 1963. *A monetary history of the United States*, 1867–1960. Princeton, N.J.: Princeton University Press (for National Bureau of Economic Research).

Haubrich, Joseph G., and King, Robert G. 1983. Banking and insurance. Unpublished manuscript, University of Rochester and National Bureau of Economic Research.

King, Robert G., and Plosser, Charles I. 1984. Money, credit, and prices in a real business cycle. *American Economic Review* 74 (June): 363–80.

Kydland, Finn E., and Prescott, Edward C. 1982. Time to build and aggregate fluctuations. *Econometrica* 50 (November): 1345–70.

Long, John B., Jr., and Plosser, Charles I. 1983. Real business cycles. *Journal of Political Economy* 91 (February): 39–69.

Lucas, Robert E., Jr. 1972. Expectations and the neutrality of money. *Journal of Economic Theory* 4 (April): 103–24.

———. 1977. Understanding business cycles. In *Stabilization of the domestic and international economy*, ed. Karl Brunner and Allan H. Meltzer. Carnegie-Rochester Conference Series on Public Policy 5: 7–29. Amsterdam: North-Holland.

Prescott, Edward C. 1983. Can the cycle be reconciled with a consistent theory of expectations? Or a progress report on business cycle theory. Research Department Working Paper 239. Federal Reserve Bank of Minneapolis.

———. 1986. Theory ahead of business cycle measurement. *Federal Reserve Bank of Minneapolis Quarterly Review* 10 (Fall): 9–22. Also in *Real business cycles, real exchange rates and actual policies*, ed. Karl Brunner and Allan H. Meltzer. Carnegie-Rochester Conference Series on Public Policy 25 (Autumn): 1–44. Amsterdam: North-Holland.

Sargent, Thomas J., and Wallace, Neil. 1982. The real-bills doctrine versus the quantity theory: A reconsideration. *Journal of Political Economy* 90 (December): 1212–36.

Smith, Bruce D. 1984. Private information, deposit interest rates, and the "stability" of the banking system. *Journal of Monetary Economics* 14 (November): 293–318.

Tobin, James. 1961. Money, capital, and other stores of value. *American Economic Review* 51 (May): 26–37.

———. 1963. Commercial banks as creators of "money." Reprint 1971. In *Essays in economics.* Vol. 1, *Macroeconomics,* pp. 272–82. Chicago: Markham Publishing.

———. 1969. A general equilibrium approach to monetary theory. *Journal of Money, Credit, and Banking* 1 (February): 15–29.

———. 1970. Money and income: Post hoc ergo propter hoc? *Quarterly Journal of Economics* 84 (May): 301–17.

Tobin, James, and Brainard, William C. 1963. Financial intermediaries and the effectiveness of monetary controls. *American Economic Review* 53 (May): 383–400.

Townsend, Robert M. 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory* 21 (October): 265–93.

———. 1987. Private information and limited insurance: The revelation principle extended. Unpublished manuscript, University of Chicago.

Wallace, Neil. 1980. The overlapping generations model of fiat money. In *Models of Monetary Economies,* ed. John H. Kareken and Neil Wallace, pp. 49–82. Minneapolis: Federal Reserve Bank of Minneapolis.

Williamson, Stephen D. 1986. Costly monitoring, financial intermediation, and equilibrium credit rationing. *Journal of Monetary Economics* 18 (September): 159–79.

———. Forthcoming *a.* Financial intermediation, business failures, and real business cycles. *Journal of Political Economy.*

———. Forthcoming *b.* Liquidity, banking, and bank failures. *International Economic Review.*

The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.