JOURNAL OF FINANCIAL INTERMEDIATION **7**, 177–197 (1998)
ARTICLE NO. JF980239

Intermediated versus Direct Investment: Optimal Liquidity Provision and Dynamic Incentive Compatibility*

Ernst-Ludwig von Thadden

Département d'économétrie et d'économie politique, Université de Lausanne, Lausanne, Switzerland

Received April 21, 1997

The existing banking literature leaves largely unanswered the question: what is the viability of bank liquidity provision if investors can dynamically readjust their portfolios? To address this question, I analyze the problem of optimal liquidity provision through bank deposit contracts in a simple continuous-time equilibrium model under uncertainty. My model introduces the possibility of investors' directly investing in the market, which gives rise to a moral hazard problem in the use of deposit contracts. I argue that this can severely restrict liquidity provision and characterizes incentive-compatible deposit contracts as second-best mechanisms to provide liquidity. The analysis shows that at the optimum, liquidity provision is negatively correlated with the degree of irreversibility of the market investment opportunity. In particular, when the market investment opportunity is completely reversible, deposit contracts cannot provide any insurance against liquidity risks. *Journal of Economic Literature* Classification Numbers: D 51, D 92, G 20, G 21.

I. INTRODUCTION

One of the important functions usually attributed to the banking sector of an economy is the provision of liquidity. This function is valuable if the economy has long-term investment opportunities, but individuals face short-term liquidity needs, as has been argued most prominently by Diamond and Dybvig (1983), building on earlier work by Bryant (1980).[1]

* I am grateful to Abhijit Banerjee, Ed Green, Martin Hellwig, Ken Judd, Nobu Kiyotaki, Raghu Rajan, Hyun Song Shin, Scott Stern, Anjan Thakor, Neil Wallace, Andrew Winton, and, in particular, a referee for helpful comments or discussions. For part of the research reported here, I have benefited from the hospitality of the Federal Reserve Bank of Minneapolis and of the Department of Economics at Stanford University. Financial support by the Schweizerischer Nationalfonds is gratefully acknowledged.

[1] See Bhattacharya and Thakor (1993) for a comprehensive survey.

1042-9573/98 $25.00

The present paper does not question the potential value of bank-intermediated investment, but investigates its viability. More precisely, it asks: what is the scope and structure of liquidity provision if individuals can also invest directly in the market, instead of using the intermediary? To address this question, which has been largely left open in the literature, I develop a generalized, continuous-time version of the Diamond–Dybvig model. There are mainly two reasons for developing the model in continuous time. First, as Diamond and Dybvig themselves note, the interaction between the customers of a bank (and particular, their withdrawal decisions) is dynamic, which cannot be fully captured in their three-period model. They use a trick to describe this interaction within one time period (random rationing) which allows them "to capture the flavor of continuous time (in which depositors deposit and withdraw at different random times) in a discrete model" (Diamond and Dybvig (1983, p. 408)). The model presented here provides a step towards studying this interaction directly.

More importantly, the analysis of a model with many time points—of which the continuous-time model is a convenient limiting case—allows for a richer understanding of the instabilities inherent in the bank–depositor relationship and of the interplay between technological factors, consumer preferences, and uncertainty influencing them. These instabilities arise because the principal macroeconomic benefit of maturity transformation conflicts, by its very nature, with the microeconomics of prudent banking, as expressed in the age-old "golden rule of banking."[2] The strict matching of assets and liabilities implied by this rule would eliminate any scope for liquidity transformation.

The concern expressed in the golden rule of banking is that depositors may have incentives to withdraw their short-term deposits and thus endanger the operation of banks with illiquid portfolios. The present paper studies such incentives in a dynamic model and derives their implications for the design of deposit contracts.

The incentive problem considered here is a direct consequence of the liquidity service of deposit contracts if investors have the choice between direct and intermediated investment. In the classical model of Diamond and Dybvig (1983), investors not only choose intermediated investment in an ex ante sense, because it provides higher expected utility than direct investment, but they also adhere to this choice in an ex post sense, once uncertainty is resolved. However, in a world with repeated investment and withdrawal possibilities and ongoing uncertainty, such investor behavior

[2] I do not know how old this rule actually is. A 140-year-old German textbook (Hübner, 1854) states it in the following way: "The credit a bank can give without risking not to be able to honor its liabilities must correspond, not only in its quantity, but also in its quality, to the credit it enjoys."

cannot be taken for granted. Precisely because bank deposits provide greater liquidity than the underlying investment opportunities, a depositor has an incentive to withdraw her deposit, thus realizing the liquidity premium it provides, in order to reinvest it elsewhere. Such a deviation from the intended purpose of the deposit contract (provision for unexpected consumption needs), however, has the downside that the alternative direct investment is less liquid. Hence, there is a tradeoff between sticking to intermediated investment and switching to direct investment, and this tradeoff is present, continuously, at all dates at which the investor has not yet consumed her deposits: stick to a deposit arrangement which provides high liquidity at lower levels of overall returns, or withdraw and reinvest to obtain extra returns which are less liquid.[3]

Jacklin (1987) was the first to study incentive problems related to nonintermediated finance in the Diamond–Dybvig model. The key observation in his analysis of the role of financial markets in the Diamond–Dybvig model is the following. Suppose a bank offers the Diamond–Dybvig risk sharing deposit contract and investors can meet and trade after having deposited their funds with the bank. Then each investor has an incentive not to invest via the bank but to invest directly. The reason is that if the investor has no liquidity needs, she will be happy to enjoy the long-term returns of her investment. If, on the other hand, she needs liquidity, she can trade her long-term claim with a more patient investor against that investor's bank deposit. Both investors are better off and this breaks the bank.[4]

In response to Jacklin's (1987) critique, Wallace (1988, 1990) has argued that deposit arrangements should be viewed as alternatives rather than as complements to trading mechanisms. According to Wallace (1988, 1990), individuals use deposit contracts precisely because they anticipate not to have access to trading opportunities when liquidity needs arise. Diamond (1997), in turn, has proposed a middle ground between the critique of Jacklin (1987) and the approach of Wallace (1988, 1990) by suggesting that individual market participation ex post plays a role in creating liquidity, but is limited. In his model, some agents are assumed to be able to trade

[3] The multiperiod investment problem studied here is somewhat complementary to the recent literature on liquidity provision in overlapping-generations models (in particular, Qi (1994) and Bhattacharya and Padilla (1996)). The OLG approach allows for intergenerational exchange, and therefore has one interesting similarity to the approach of this paper: OLG deposit contracts must satisfy a "no-redepositing constraint," a specific incentive compatibility constraint which prevents middle-aged agents from withdrawing and later redepositing their funds.

[4] Jacklin (1987, 1993) and Haubrich and King (1990) analyze in more detail the role of instruments other than fixed rate demand deposits in a model with ex post trading opportunities.

when liquidity needs arise, and others are not. As in Bhattacharya and Gale (1987) and von Thadden (1997), Diamond's (1997) model has investment opportunities of different term structures and banks can provide liquidity by pooling consumption risks in a portfolio of long- and short-term assets. Now agents will simultaneously hold bank deposits and invest directly: holding deposits provides liquidity through the term-structure pooling effect identified in von Thadden (1997), whereas investing directly yields the trading benefits identified by Jacklin (1987) if the agent has access to the secondary market. In this sense, Diamond's (1997) model combines the liquidity benefits of markets and of deposit contracts and shows how they interact.

As outlined above, I take a different perspective on the incentive problem. In my model, strategic withdrawals are not made for explicit trading purposes, but rather to invest (passively) in the market portfolio. In particular, the incentive problem of individual withdrawals arises even in Wallace's (1988, 1990) interpretation of deposit contracts as a substitute for trading opportunities. Taking these incentives for strategic withdrawal into account, however, makes the problem of designing deposit contracts technically more complex.

Yet, while the problem has little general structure, a remarkably simple structure obtains if the investors' intertemporal risk aversion is sufficiently high compared to the degree of irreversibility of investment. In this case, which is, as I argue later, quite realistic and is the focus of this paper,[5] incentive compatibility is binding over the whole time-horizon.

The analysis then shows that the value of maturity transformation, as described by the first-best, and the scope for it, as determined by the second-best, may be significantly at variance, depending on the parameters of the economy under consideration. From the first-best analysis, maturity transformation is, ceteris paribus, more valuable (to the investors) the larger the investors' intertemporal risk aversion. However, as the second-best analysis shows, the scope for maturity transformation is, independent of investors' preferences, restricted by the dynamic properties of the underlying investment opportunity. The more easily the latter is reversible (in the sense of Wallace (1988)), the less maturity transformation is possible. In the extreme case in which the aggregate investment opportunity is given by a sequence of short-term investments (no irreversibility), maturity transformation is completely impossible, and despite a possibly high demand for liquidity, deposit arrangements can only replicate the autarky allocation.

This finding also shows that one must be careful to distinguish different notions of liquidity.[6] In particular, one might argue that an investment is

[5] The full analysis is given in von Thadden (1996).

[6] I am grateful to the referee for encouraging me to discuss these differences explicitly.

liquid if it is completely reversible (in the sense that its liquidation entails no deadweight loss). Under this definition, liquidity is a characteristic of the investment opportunity which is unrelated to investors' preferences. Banks, then, cannot provide liquidity, but only transform the term structure of returns. In the above mentioned case of fully reversible investment and high intertemporal risk aversion, for example, one would therefore have a high demand for maturity transformation in the presence of a fully liquid asset. Although in the end this question is only semantic, in this paper I stick to the more conventional notion that liquidity measures how easily investors can realize the long-term value of their investment. Therefore, liquidity is related to investors' preferences, and, as in the whole Diamond–Dybvig literature, maturity transformation is the same as liquidity provision.

The remainder of this paper is organized as follows. Section II sets out the model, a direct adaption of the model by Diamond and Dybvig (1983) to continuous time, and describes the first-best. Section III analyzes incentive-compatible deposit contracts. Section IV provides an interpretation of the results and concludes. Proofs are in the appendix.

II. THE MODEL

Consider an economy with one good and a continuum of identical households $a \in [0, 1]$ who each live from time $t = 0$ to $t = 1$. Time is measured continuously, with $t \in [0, 1]$.

Each household is endowed with one unit of the good at time $t = 0$ that it can invest or store without depreciation. Everybody in the economy has access to the same constant-returns-to-scale investment opportunity,[7] which, over any time-interval $[\tau, t]$, $0 \leq \tau < t \leq 1$, yields a gross return of $R(t - \tau)$. In other words, an investment of $a \geq 0$ units of funds invested at time τ yields $aR(t - \tau)$ units of funds when liquidated at time t. Investment and its liquidation are costless.[8] The return function R is assumed to be differentiable and increasing, with $R(0) = 1$. Furthermore, $g := R'/R$ is nondecreasing on $(0, 1)$.

The last condition implies that investing funds for a time-span of t yields at least as much as investing them for a time-span of $\tau < t$, liquidating

[7] For the case of multiple investment opportunities see von Thadden (1997) and Hellwig (1994).

[8] It is not difficult to incorporate transaction costs into the analysis. The main consequence of transaction costs would be to make the depositors' incentive compatibility constraint introduced in Section III less tight, which in turn would facilitate the provision of liquidity. More generally, one can consider market return functions $R(\tau, t)$ for returns from investment at date t accruing in τ, and much of the analysis would carry over (here we are concerned with $R(\tau, t) = R(\tau - t)$).

them, and investing the proceeds for another period of $t - \tau$.[9] Absent fixed costs for liquidation or investment, this must hold if funds are used efficiently. If $g(t)$ is strictly increasing, then "the real investment technology has an irreversibility, or goods-in-process, feature" (Wallace, 1988): a sequence of short-term investments is strictly inferior to a long-term investment.

Following Diamond and Dybvig (1983) and much of the literature, households' preferences exhibit the following strong intertemporal asymmetry. In "normal" times, households consume a constant, perfectly predictable flow of funds, not modelled here and normalized to zero. However, for each household a there is a time T_a at which its demand is singular and where it needs to consume all its wealth. This time-point is not predictable individually; it hits the household as an exogenous random shock. The random variables $\{T_a\}_{a\in[0,1]}$, $0 < T_a \leq 1$, are assumed to have c.d.f. F and to satisfy the Law of Large Numbers (see Judd, 1985). Hence, for each individual household the date T_a is a random event on the interval (0, 1] with c.d.f. F, whereas for the population as a whole the distribution of shocks is non-stochastic with empirical distribution function F. To simplify the exposition assume that F has a continuous density $F' =: f > 0$.[10]

To emphasize the importance of the unforeseen consumption shock, a household's utility over its lifetime is assumed to depend solely on what it can consume at time T_a.[11] Households' expected utility as of time $t = 0$ then is

$$U = \int_0^1 u(w(T_a))f(T_a)dT_a,$$

where $w(t)$ is the household's wealth at time t, and its instantaneous utility $u : IR_+ \to IR$ satisfies the usual Inada conditions $u' > 0$, $u'' < 0$, $u'(0) = \infty$, $u'(\infty) = 0$.

This completes the description of the model, a straightforward generalization of the model of Diamond and Dybvig (1983) to continuous time. The four-period production model of Postlewaite and Vives (1987) is also easily translated into the above framework if $R(0) < 1$ is admitted.

Absent any interaction in the economy, each household can invest its funds in $t = 0$ in order to liquidate the investment in T_a. This yields the autarkic utility level $\overline{U} := \int_0^1 u(R(t))f(t)\,dt$.

However, if the consumption shocks T_a are commonly observable and

[9] Formally, $R(t) \geq R(t - \tau)R(\tau)$.

[10] See von Thadden (1996) for the general case.

[11] For Diamond and Dybvig's (1983) three-period model, Jacklin (1987) discusses the case of more symmetric intertemporal preferences.

contractible, households can usually do better by acting collectively. Under autarky, an individual household is forced to consume a random amount $\tilde{c}$ with c.d.f. $F \circ R^{-1}$. Since uncertainty washes out in the aggregate and the fluctuations in consumption given by $F \circ R^{-1}$ will typically be neither in line with the households' risk preferences nor with their preferences for intertemporal substitution, there are gains from reallocating funds over time between households. Such a reallocation can be achieved by investing endowments collectively at $t = 0$ and liquidating them according to a collectively agreed upon rule, subject to the constraint that aggregate consumption be feasible.

Therefore, suppose the households pool their funds at $t = 0$ and invest collectively. In the collective, it is clearly a dominated strategy to liquidate assets in order to reinvest them. Hence, all funds are invested in $t = 0$ and liquidated only for consumption purposes. Let $S(t)$, $0 \leq t \leq 1$, denote the aggregate amount of funds available at time t. Let $x(t)$ be the (per capita) share of deposits an individual can withdraw if it needs liquidity on date t, yielding consumption of $x(t)R(t)$ (under autarky, for example, we have $x(t) = 1$ for all t). The evolution of S is determined by x. In the absence of liquidation, S would instantaneously evolve as R, i.e., with a growth rate of R'/R. Hence,

$$S'(t; x) = g(t)S(t; x) - x(t)R(t)f(t). \tag{1}$$

Straightforward integration of (1), together with the initial condition $S(0) = 1$, yields

$$S(t; x) = R(t)\left(1 - \int_0^t x(\tau)f(\tau)d\tau\right). \tag{2}$$

The problem of finding the first-best optimal aggregate consumption path $x^*(\cdot)$ is now simply that of maximizing ex ante expected utility subject to the constraint that each investor receives one unit on average, i.e., that $S(1; x) = 0$:

$$\max \int_0^1 u(x(t)R(t))f(t)\,dt \tag{3}$$

$$\text{subject to} \qquad x \geq 0 \text{ integrable on } [0, 1], \tag{4}$$

$$\int_0^1 x(t)f(t)\,dt = 1. \tag{5}$$

Let $q := (u')^{-1}$ denote the inverse of the marginal utility function. The following proposition provides a simple characterization of the first-best.

PROPOSITION 1. *The unique solution to problem* (3)–(5) *is*

$$x^*(t) = q\left(\frac{C^*}{R(t)}\right)\Big/ R(t), \tag{6}$$

where the constant C^* is given by

$$\int_0^1 x^*(t) f(t)\, dt = 1. \tag{5}$$

Proposition 1 has a straightforward economic interpretation, with a somewhat surprising aspect. Equation (6) states that at the optimum, marginal utility of consumption at time t, $u'(x^*(t)R(t))$, weighed with the marginal rate of transformation from time 0 to time t, $R(t)$, must be constant over time. In particular, the gross rate of return on collective investment, $x^*(t)R(t)$, is a smooth and strictly increasing function of time, although the distribution of shocks, given by f, may be very irregular. The distribution of T_a influences x^* only through the constant C^* determined by the "no-waste" condition (5).

Notice that the economy could choose a constant consumption flow $x(t)R(t) \equiv \bar{c}$. However, although this would completely eliminate consumption risk, this is not optimal. This simple observation suggests that it is not only the households' risk aversion which makes collective investment attractive, but also the fact that the timing of investment returns does not correspond to the households' intertemporal consumption preferences. To see the impact of risk aversion more clearly, differentiate (6) to get

$$\begin{aligned} \frac{x^{*\prime}(t)}{x^*(t)} &= -\frac{R'(t)}{R(t)} \frac{u'(x^*(t)R(t)) + x^*(t)R(t)u''(x^*(t)R(t))}{x^*(t)R(t)u''(x^*(t)R(t))} \\ &= g(t)\frac{1-\gamma(x^*(t)R(t))}{\gamma(x^*(t)R(t))}, \end{aligned} \tag{7}$$

where

$$\gamma(c) := -\frac{cu''(c)}{u'(c)}$$

is the absolute value of the elasticity of households' marginal utility with respect to consumption. In decision theory under uncertainty, γ is the households' (static) coefficient of relative risk aversion. In deterministic problems of intertemporal consumption choice with time separable utility, γ denotes the relative importance of the income effect as compared to the substitution effect, and also the inverse of the intertemporal elasticity of substitution. Hence, with the simple preferences employed here, several aspects, in particular risk sharing and intertemporal substitution effects, determine the shape of x^*.

An important insight of Bryant (1980) and Diamond and Dybvig (1983) has been to relate the demand for liquidity to γ. Following them and the literature they have inspired, we assume that $\gamma(c)$ is sufficiently large and, for simplicity, that it is constant:

Assumption. $\gamma(c) \equiv \gamma > 1$ for all $c > 0$.

Under this assumption, x^* is decreasing by (7). Therefore, the first-best consumption profile is always flatter than the return path available from the economy-wide investment opportunity. In other words, in an economy with relative risk aversion of more than unity, consumption along the first-best path is shifted from later to earlier times, in the sense that a household hit by an early consumption shock consumes more than its funds have physically produced up to that time, at the expense of households who will consume later. Conversely, in an economy with $\gamma(c) < 1$ for all $c > 0$, consumption would optimally be shifted towards later time points.

III. DEPOSIT CONTRACTS

If the consumption shocks T_a are not observable, the question is what intertemporal risk-sharing possibilities can be achieved through incentive-compatible contracting. The central contribution by Bryant (1980) and Diamond and Dybvig (1983) has been to show how to interpret demand deposits as mechanisms that provide such risk sharing possibilities for the economy.

Following them, suppose that the households in the economy set up an organization—henceforth called "the bank"—that collects the households' funds in $t = 0$ as deposits and invests them collectively. If a household wants to consume a fraction δ of its wealth at time t, it can withdraw it from the bank and obtains $\delta r(t)$, where $r(t)$ is the gross interest rate at time t, prespecified in the deposit contract in $t = 0$.

Here, a deposit contract between a bank and its customers should be properly viewed as a mechanism defining a "withdrawal game." In this game, all households deposit their funds in the bank, and the return to each customer from depositing depends on the deposit contract and the withdrawal decisions of all other customers. The question posed by Diamond and Dybvig (1983) is whether the first-best consumption path x^*R can be implemented as a unique Nash equilibrium of the withdrawal game between depositors.

In their model, the answer is in the affirmative. The only reason for an individual household to withdraw funds at a date before T_a would be the belief that other households will do so as well. To prevent an equilibrium in which such beliefs are correct (a "bank-run"), the mechanism only has to include a provision such as suspension of convertibility. If the bank stops paying out funds to depositors ("suspends convertibility") once withdrawal becomes excessive (compared to F), there is no danger that the bank will run out of funds prematurely, and everybody's demand for funds can be satisfied at T_a.[12] If everybody knows this, it is individually strictly optimal to withdraw only at T_a, and the first-best allocation of funds results.

This argument is one of individual opportunity costs: given that inefficient Nash equilibria are eliminated, in the only remaining equilibrium the individual cannot use her funds more profitably than leaving them in the bank. In particular, in the parameter constellation discussed by Diamond and Dybvig (1983), the option of reinvesting funds elsewhere is dominated. The following proposition shows that in general, this is not the case.

PROPOSITION 2. *The first-best path x^* cannot be implemented through deposit contracts with suspension of convertibility.*

Proposition 2 results from a basic tradeoff for the individual household. Since its intertemporal risk aversion is relatively high ($\gamma > 1$), the first-best flattens the aggregate return path by shifting returns forward in time. Suppose, then, that a deposit contract offers the first-best and that a household that is not hit by a consumption shock deviates from the contract by withdrawing its deposits at time $\varepsilon > 0$. If the household reinvests these funds, this generates a new individual return path of $x^*(\varepsilon)R(t)$ which had not been feasible before. Hence, the household must decide whether it prefers the relatively flat first-best consumption path $x^*(t)R(t)$ or the "unflattened" individual path. The premium of the first-best over autarky is highest at $\varepsilon = 0$ and, because x^* is strictly decreasing, high enough to even push the new individual path uniformly above the first-best path. Therefore, if withdrawal around date 0 is possible (which is the case if $f > 0$ around 0), everybody will demand it. In other words, any attempt to implement the first-best induces a bank run in strictly dominant strategies in $t = 0$.[13]

What is the reason for the apparent discrepancy between Proposition 2 and the result of Diamond and Dybvig (1983)? The Diamond–Dybvig model can be approximated as a special case of the present framework by concentrating the mass of F around $t = 1/2$ and $t = 1$ and by having

[12] This argument depends crucially on the absence of aggregate risk. For the case of aggregate risk, see Wallace (1990) and Green (1995).

[13] Strictly speaking, everybody attempts to withdraw "as soon after $t = 0$ as possible," since withdrawal at $t = 0$ is not possible. This trivial open-set problem causes non-existence of equilibrium rather than a bank-run equilibrium.

$R(1/2) \approx 1$. If $f \equiv 0$ on $(0, 1/2)$ this has two consequences: the bank can refuse to pay out at early t, and reinvestment, which therefore can only happen at later t, is unattractive for depositors. Thus, in this case depositors will leave their funds in the bank. However, if consumption needs arise over a longer time span and the bank is also forced to satisfy early withdrawal demands (because $f > 0$), Proposition 2 shows that early withdrawal and reinvestment is always more profitable for depositors.

Note that Proposition 2 holds regardless of risk preferences (γ) and production possibilities (R). The only assumption is that consumption needs arise early on in the life of the investment opportunity. As the following analysis will show, the structure of incentive-compatible liquidity provision will in general be determined by the interplay between all three factors—preferences, technology, and distribution of consumption shocks.

At each time point, a household not yet hit by a consumption shock must decide whether to leave its funds with the bank or to withdraw and reinvest them privately. Defining, for a given withdrawal schedule x, the function

$$H(t) := H(t; x) := \int_t^1 [u(x(\tau)R(\tau)) - u(x(t)R(t)R(\tau - t))]f(\tau)\, d\tau, \qquad (8)$$

$t \in [0, 1]$, a depositor's expected net gain from leaving funds with the bank at $t \in [0, 1]$ is $H(t)/(1 - F(t))$. Incentive-compatible liquidity provision therefore poses the problem

$$\max \quad \int_0^1 u(x(t)R(t))f(t)\, dt \qquad (3)$$

$$\text{subject to} \quad x \geq 0 \text{ integrable on } [0, 1], \qquad (4)$$

$$\int_0^1 x(t)f(t)\, dt = 1, \qquad (5)$$

$$H(t; x) \geq 0 \;\forall t \in (0, 1]. \qquad (9)$$

Problem (3)–(5), (9) differs from standard optimal control problems in several respects. First, it involves a continuum of constraints. More important, the constraints (9) involve averages over the future path of x together with point values. This feature, which, in particular, differentiates the above problem from the literature on optimal consumption under borrowing constraints (e.g., Scheinkman and Weiss (1986), He and Pagès (1993)), implies that the payout at one single point in time is restricted by the path of all future payouts. On the other hand, by (5), the path of past payouts obviously restricts future payouts. Finally note that the incentive constraint (9) renders the problem nonconvex.

Intuitively, the designer of an incentive-compatible deposit contract faces the following problem. At any given time $\tau \in (0, 1)$, the incentive-compatibility constraint (9) tends to push the value $x(\tau)$ down in order to make withdrawal and reinvestment unattractive. On the other hand, from the perspective of earlier $t < \tau$, future payouts $x(\tau)$ should be high in order to make sticking to the collective investment scheme as attractive as possible. The optimal contract has to balance these demands all along the path x.

The problem clearly has a recursive structure: given the maturity of the assets in place, at any point in time the objective function and constraints are forward looking. But the problem has even more structure. To see this, suppose that the incentive compatibility constraint (9) is slack at one point t_0. It is simple to see that, because of the strict concavity of u, any solution to (3)–(5), (9) cannot have jumps. Therefore, for a given solution x, $H(\cdot; x)$ is continuous, too. Because $H(t_0; x) > 0$ and $H(1; x) = 0$, there exists a $t_1 > t_0$ such that H is strictly positive on (t_0, t_1) and $H(t_1; x) = 0$. Intuitively, if the incentive compatibility at one point does not bind, it cannot bind locally around this point, but it must bind later, because at the very last moment everybody who has not yet done so will want to cash in.

This simple observation implies that on (t_0, t_1) an optimal incentive-compatible contract solves an optimization problem which is only constrained by the amounts of assets in place at time t_0 and t_1, not by incentive considerations. To make this precise, suppose that x is a solution to our problem and that $S(t_0; x) = S_0$ and $S(t_1; x) = S_1$. Then, on $[t_0, t_1]$ x must solve the unconstrained problem as in (3)–(5), modified such that the asset stock as given by (2) is depleted from S_0 to S_1 instead of from 1 to 0. Hence, on $[t_0, t_1]$ x solves

$$\max \int_{t_0}^{t_1} u(x(\tau)R(\tau))f(\tau)\, d\tau$$

$$\text{subject to} \qquad x \geq 0 \text{ integrable on } [t_0, t_1],$$

$$\int_{t_0}^{t_1} x(\tau)f(\tau)\, d\tau = \frac{S(t_0; x)}{R(t_0)} - \frac{S(t_1; x)}{R(t_1)}.$$

Using Proposition 1, we know therefore that x must have the form (using the assumption of constant relative risk aversion)

$$\begin{aligned} x(t) &= q\left(\frac{C}{R(t)}\right)\Big/ R(t) \\ &= C^{1/\gamma} R(t)^{(1-\gamma)/\gamma} \end{aligned}$$

on $[t_0, t_1]$ for some constant $C > 0$. Hence, the problem has a "local optimality property": solutions to the local unconstrained problems differ from the first-best only by a constant factor (which depends on the local data).

This result allows us to write out H explicitly on (t_0, t_1) (where it is strictly positive),

$$H(t; x) = \frac{1}{\gamma - 1} C^{(1-\gamma)/\gamma} G(t), \tag{10}$$

where

$$G(t) = \int_t^1 [R(t)^{(1-\gamma)/\gamma} R(\tau - t)^{1-\gamma} - R(\tau)^{(1-\gamma)/\gamma}] f(\tau)\, d\tau.$$

Note that G does not depend on the constant C and therefore not on t_0 and t_1. This property and the local optimality property of the overall problem make G a useful tool in characterizing the solution.

To simplify the exposition, the following proposition characterizes the incentive compatible deposit contract only for one important case, in which the solution takes a particularly simple form. This is the case where the households' coefficient of relative risk aversion is relatively large as compared to the degree of irreversibility of the investment opportunity. This case displays many features of the general solution and allows to draw some interesting conclusions.

PROPOSITION 3. *Suppose that* $\gamma g(0) \geq g(1)$. *The second-best problem* (3)–(5), (9) *has a unique solution* $\bar{x}$, *which is continuously differentiable. Along* $\bar{x}$, *the incentive constraints bind for all* $t \in [0,1]$, *and* $\bar{x}$ *is given by*

$$\frac{\bar{x}'(t)}{\bar{x}(t)} = \int_t^1 R(\tau - t)^{-\gamma+1} g(\tau - t) f(\tau)\, d\tau \Big/ \int_t^1 R(\tau - t)^{-\gamma+1} f(\tau)\, d\tau - \frac{R'(t)}{R(t)}. \tag{11}$$

Remember that $g(t)$ is the growth rate of funds invested at date 0. For $g(0) > 0$ we can interpret $g(1)/g(0)$ as an index of the irreversibility of the investment opportunity. At the smallest possible value of $g(1)/g(0)$, which is 1, $g(t)$ is constant and invested funds grow at a constant rate, hence $R(\tau) = R(\tau - t)R(t)$ for $\tau > t$. This means that the investment displays no irreversibility: a sequence of short-term investments is as good as one long-term investment.[14] The larger $g(1)/g(0)$, the greater the overall irreversibil-

[14] This case shows how one can make a conceptual distinction between maturity transformation and liquidity provision. A bank policy of holding only short-term liabilities and interrupting production whenever needed causes no deadweight loss in production—hence, the bank does not provide liquidity. Yet, if $\gamma > 1$ there is demand for maturity transformation.

ity of the investment opportunity, and the larger the relative disadvantage of short-term investments.

The proof and the idea of Proposition 3 are straightforward, given the above discussion of local optimality. If the incentive constraint were slack for one t_0, it would be slack for a whole interval, on which we could express H through G as given by (10). Now it can be easily shown that the condition $\gamma g(0) \geq g(1)$ implies that $G' \geq 0$ on $(0, 1)$, which means that the relative advantage of keeping deposits in the bank over withdrawing them prematurely grows over time. Then, of course, the incentive constraints must be slack for the whole remaining time, $(t_0, 1]$. However, we know that incentive compatibility binds trivially for $t = 1$, because everybody who has not done so yet will withdraw at the end.

Despite the nonconvexity of the problem, Proposition 3 shows that its solution is unique. The solution is characterized by a differential equation which is obtained by differentiating the identity $H(t; \bar{x}) = 0$ on $[0, 1]$ and by using the resource constraint (5). In general, the interplay between technical productivity (g), risk aversion (γ), and the distribution of consumption shocks (f) in determining the optimal deposit contract is quite complex. In particular, the incentive constraints may bind over some intervals and not over others.[15] Yet, in the case considered in Proposition 3 ($\gamma g(0) \geq g(1)$) such complications are absent.

In this case, there is no scope to relax the incentive-compatibility constraint anywhere, regardless of the distribution of consumption shocks. On the one hand, the higher the intertemporal risk aversion γ, the stronger the first-best tendency to front-load consumption, and therefore, the lower the implied returns from continuing to hold money in the bank. On the other hand, the lower the degree of irreversibility of investment, $g(1)/g(0)$, the lower the deadweight loss from switching away from the bank to outside investment. Taken together, then, if $\gamma \geq g(1)/g(0)$, any attempt to equate the weighted marginal utilities, $u'(x(t)R(t))R(t)$, over some time period would induce withdrawal and private reinvestment. Therefore, the optimal payout path is exclusively determined by incentive-compatibility considerations.

Although the construction of $\bar{x}$ has eliminated the incentives to withdraw and reinvest individually, in general, a deposit contract offering the payment path $\bar{x}R$ still provides some liquidity to households. Therefore, it is still vulnerable to expectation based bank-runs as identified by Diamond and Dybvig (1983). Hence, also in the case considered here, a second-best optimal mechanism must include a suspension-of-convertibility clause in order to eliminate inefficient equilibria. Such an arrangement—a deposit contract offering $\bar{x}R$ on demand, together with suspension of convertibility

[15] See, again, von Thadden (1996) for the solution in the general case.

if withdrawal demands exceed the rate of $f(t)$—achieves second-best liquidity provision as the unique Nash equilibrium of the withdrawal game between bank and depositors.

Equation (11) has an immediate consequence which allows us to characterize the extent of incentive compatible maturity transformation.

COROLLARY 1. *If* $\gamma g(0) \geq g(1)$, *the second-best deposit contract satisfies*

$$g(0) \leq \frac{\bar{x}'(t)}{\bar{x}(t)} + g(t) \leq g(1 - t). \tag{12}$$

Inequality (12) puts a bound on the incentive compatible growth rate of consumption, $\bar{x}'(t)/\bar{x}(t) + g(t)$, and, therefore, on how much maturity transformation can be achieved, regardless of its social value. To illustrate the point consider the extreme case of a fully reversible investment opportunity, $g(t) = a$ for all $t \in [0, 1]$. The (first-best) demand for liquidity transformation as expressed by (7), $x^{*\prime}(t)/x^*(t) = a(1 - \gamma)/\gamma$, requires the optimal consumption path to be the flatter the greater the households' intertemporal risk aversion, with constant consumption in the limit for $\gamma \to \infty$. However, by (12), $\bar{x}'(t)/\bar{x}(t) = 0$ on [0, 1]. Hence, $\bar{x} = 1$, and only autarky is incentive compatible, regardless of γ.

This example exhibits the complete breakdown of the liquidity role of deposit contracts in the case of full reversibility. But the point of Corollary 1 is more general. As seen in Proposition 1, the social value of liquidity transformation depends on intertemporal risk preferences (γ), technology (R), and uncertainty (F). The extent of liquidity provision, however, which deposit contracts can achieve is bounded by technological parameters alone. This implies that incentive compatibility necessarily restricts the provision of liquidity by deposit contracts, and it binds most when liquidity is most valuable, i.e., when intertemporal risk aversion is high.

IV. CONCLUSION

The analysis of this paper has identified a theoretical incentive problem in the provision of liquidity through banks which arises if investments in the economy are possible in several periods. The choice of continuous instead of discrete time for modeling the problem of this paper has largely been for esthetic reasons: the formulae become simpler and the reasoning becomes more compact. Most of the results could be derived in discrete time models with four or five periods, depending on the aspect under consideration. Continuous time provides a simple, unified framework for the analysis.[16]

The modeling of banks and reinvestment in this paper has two main interpretations, which both relate to the discussion, reviewed in the Introduction, of how to interpret the original model of Diamond and Dybvig (1983).

The first interpretation is the one given by Wallace (1988, 1990) of banks as institutions that provide intertemporal risk sharing opportunities to otherwise isolated households. The present model is consistent with this interpretation, but throws some doubt on the value of banking arrangements in this setting. In particular, if economywide investment opportunities display little irreversibility, the incentive problem of private reinvestment reduces the banks' ability to provide liquidity significantly.[17]

The main reason for the incentive problem identified here is the assumption that households and banks have the same investment opportunities available at every time—the banks' only advantage is the ability to exploit the Law of Large Numbers. If one were to endow the banking sector exogenously with a richer set of investment opportunities than the general public, the incentive problem associated with maturity transformation would be correspondingly reduced. This, however, would mean that the truly special feature of banks would be their lending capabilities, with liquidity provision being a derivative function. In other words, maturity transformation would be a rather fragile function to justify banks on its own.

Banks as providers of liquidity become even more fragile if one makes the model more realistic with respect to banking competition. In the model presented here, as in all of the maturity transformation literature, there is only one bank. While in the standard three-period model, banking competition can be introduced in a relative straightforward manner (see, e.g., Hellwig, 1994), bank competition renders the incentive problem worse in this model. The reason is that an incentive-compatible deposit contract of one bank must be robust not only with respect to deviations to private reinvestments, but even with respect to deviations to other banks' optimal contracts. Since these contracts provide potentially better risk sharing than R, the incentive constraints become stricter, and there is even less scope for liquidity provision.

The same remark applies to the possibility of redepositing funds with the same bank. In practice, of course, there are usually no limits on redepositing funds previously withdrawn from a deposit account (or a savings

[16] To be sure, the nonconvexities which create the mathematical difficulty of the analysis do not go away when using discrete time. See Engineer (1989) for an example of a four-period discrete-time model of liquidity provision in which the formulae and the different subcases become quite burdensome.

[17] Note that this criticism also applies to Diamond's (1997) synthesis of Wallace and Jacklin.

account) into the same account. In the present model, because agents obtain income only at $t = 0$, deposit contracts simply forbid redepositing. In a more general model, in which agents receive income more than once, redepositing will typically be allowed (albeit perhaps at a cost), which restricts the scope for liquidity provision by the initial contract even further. All these considerations throw doubt on the illiquidity of banks even in Wallace's interpretation of the Diamond–Dybvig model.

The other interpretation of the Diamond–Dybvig model is that the model's bank coexists with financial markets and that banking activity must be robust against trading possibilities available in the market (Jacklin, 1987). My model has a natural interpretation along these lines, if one adopts a partial equilibrium perspective.[18] In this perspective, R represents the returns generated by the market portfolio which is available to everybody, and banks transform the maturity of this asset. For maturity transformation to be compatible with market activity, banks then only have to respect the incentive compatibility constraint introduced in Section III.

This constraint is less severe than the one considered by Jacklin (1987), who assumes that investors who do not deposit their funds with the bank can freely transact with depositors. The reason is the following. Because of irreversibility in the present model, the original return stream $R(\cdot)$ when obtained through trading at date t is more valuable than the return stream $R(\cdot - t)$ obtained through private reinvestment. Therefore, a depositor who trades at date t with an investor who holds the original asset $R(\cdot)$ obtains weakly more than what he would get through directly reinvesting privately. How much of the surplus he obtains depends on the organization of markets, but certainly his incentive constraint will be (weakly) stricter than (9).[19]

The advantage of the formulation of this paper in a partial equilibrium perspective is that one can ignore the question of where the market return R comes from. It suffices to know that investors can obtain this return without specifying what trading mechanisms are available to individuals. The analysis then shows how much liquidity provision through banks is possible in the presence of such investment opportunities, depending on the features of R.

[18] Like most of the work following them, the models by Diamond and Dybvig (1983) and Jacklin (1987) are general equilibrium models. As Jacklin's analysis shows, this necessitates a lot of work to model all relevant institutions and markets. A partial equilibrium perspective is, of course, less ambitious, but usually more tractable.

[19] At one extreme of the distribution of market power, the original return stream would be priced such that the depositor would just get the value of her private reinvestment opportunity. This would yield incentive constraints identical to (8) and (9). At the other extreme, he would get the full value of the original return stream, resulting in ("strictly") stricter incentive constraints.

The conclusions to be drawn from the analysis for the actual scope of maturity transformation depend on the interpretation of the function R. If one believes that aggregate investment opportunities exhibit significant irreversibilities, then the present analysis confirms, in a dynamic framework, the interpretation of demand deposits given by Diamond and Dybvig (1983): demand deposits facilitate intertemporal risk sharing.

If one believes, on the other hand, that aggregate investment is relatively liquid, be it because of substitution effects between inflowing and outflowing capital or of intermediation services outside the banking system, then the present analysis has identified a gap between the demand for liquidity and the scope for liquidity provision through demand deposits. This gap may be quite large given the relatively high empirical estimates of individuals' relative risk aversion. Furthermore, if one interpretes $1/\gamma$ as an indicator of intertemporal substitution elasticities (which cannot be distinguished from risk aversion in the present simple framework), the work by Hall (1988) and others also suggests relatively high values of γ.

If this constellation is the relevant one—little irreversibility in aggregate investment and high intertemporal risk aversion—then one might have to consider other arrangements than demand deposits in order to achieve good intertemporal risk sharing. One such arrangement, which is very much in the spirit of liquidity provision advocated by Diamond and Dybvig (1983), is life insurance. The lifetime payout structure in a life insurance policy is almost exactly what the present analysis predicts as the first-best for high intertemporal risk aversion and back-loaded densities f: a constant high payout whenever the contractual contingency arises. But, of course, here we have changed the underlying market structure. With life insurance, there are no missing markets, because one can condition on an event which is very difficult to fake and extremely costly to bring about fraudulently: the death of the investor.

APPENDIX

Proof of Proposition 1. Denote by $\mathscr{F} \subset L^1$ ([0, 1]) the constraint set defined by (4)–(5). $\mathscr{F}$ is convex. By the Inada conditions on u, x^* is well defined. Clearly, $x^* \in \mathscr{F}$.

Denote the objective in (3) by $\Phi: \mathscr{F} \to IR$, $\Phi(x) := \int_0^1 u(x(t)R(t))f(t)\, dt$. For any $x \neq x^* \in \mathscr{F}$, the Gateau derivative of Φ at x^* in the direction of x is defined as

$$\partial\Phi(x^*; x - x^*) := \lim_{h \searrow 0} \frac{1}{h} [\Phi(x^* + h(x - x^*)) - \Phi(x^*)].$$

We have

$$\partial\Phi(x^*; x - x^*) = \int_0^1 C^*(x(t) - x^*(t))f(t)\,dt = 0, \tag{A1}$$

where the second equality in (A1) follows from (5).

Because Φ is concave, for every $\alpha \in (0,1)$

$$\Phi((1-\alpha)x^* + \alpha x) \geq (1-\alpha)\Phi(x^*) + \alpha\Phi(x)$$
$$\Leftrightarrow \Phi(x) - \Phi(x^*) \leq \frac{1}{\alpha}[\Phi(x^* + \alpha(x - x^*)) - \Phi(x^*)].$$

Hence (A1) implies that $\Phi(x) \leq \Phi(x^*)$. By the strict concavity of Φ this inequality is strict. ■

Proof of Proposition 2. In any equilibrium that implements the first-best, the bank has to pay out $x^*(t)R(t)$ at $t > 0$ according to the distribution F. Since $f > 0$ and the T_a are noncontractible, the bank must repay every depositor with probability 1 whenever she demands her funds (in particular, if convertibility were suspended in equilibrium with positive probability, the outcome would not be first best).

Consider a household that has not been hit by a consumption shock up to time $\varepsilon > 0$. Funds left in the bank yield future returns of $x^*(t)R(t)$, $t > \varepsilon$, per unit. Funds withdrawn can be reinvested privately where they will yield $x^*(\varepsilon)R(\varepsilon)R(t - \varepsilon)$ per unit. The latter return stream pointwise dominates the former iff

$$x^*(\varepsilon)R(\varepsilon)R(t-\varepsilon) > x^*(t)R(t), t > \varepsilon. \tag{A2}$$

Because x^* is strictly decreasing, (A2) holds for $\varepsilon = 0$. By continuity, it also holds for small positive ε. At each such time-point it is strictly optimal for each household to withdraw all funds from the bank, independent of what all other households do. ■

Proof of Proposition 3. We first verify that problem (3)–(5), (9) has a solution. By the incentive constraint (9), $x(t)$ is bounded globally from above if it is bounded locally at $t = 1$, which is straightforward to show. Hence, we can assume that $x(t) \in [0, K]$, with K large. Therefore, the problem is compact. The integrand in (3) and the right-hand side of (1) are continuous. Hence, standard compactness existence theorems for variational problems apply.

Suppose now that $H(t_0; \bar{x}) > 0$ for a $t_0 \in [0, 1]$. As argued before, then, there exists a $t_1 > t_0$ such that $H(t_1; \bar{x}) = 0$ and

$$\bar{x}(t) = q\left(\frac{C}{R(t)}\right)\Big/ R(t),$$

and, therefore,

$$H(t; \bar{x}) = \frac{1}{\gamma - 1} C^{(1-\gamma)/\gamma} G(t)$$

on $[t_0, t_1]$ for some constant $C > 0$.

By straightforward calculation,

$$G'(t) = \frac{\gamma - 1}{\gamma} R(t)^{(1-\gamma)/\gamma} \int_t^1 R(\tau - t)^{1-\gamma}[\gamma g(\tau - t) - g(t)]\, dF(\tau).$$

Therefore, if $\gamma g(0) \geq g(1)$, then

$$\gamma g(\tau - t) - g(t) \geq \gamma g(0) - g(1) \geq 0;$$

hence, $G'(t) \geq 0$ $\forall t \in [0, 1]$. It follows that $H'(t; \bar{x}) \geq 0$ on $[t_0, \mathrm{t}_1]$, which contradicts $H(t_1; \bar{x}) = 0$.

Hence, $H(t; \bar{x}) \equiv 0$ on $[0, 1]$. Differentiating this identity with respect to t yields (11). The right hand side of (11) is continuous in (t, x) (continuity at $t = 1$ follows by l'Hôpital's rule) and Lipschitz in x. Hence, for each initial value $\bar{x}(0)$ there is exactly one solution of (11). The unique solution is then determined by the resource constraint (5). ■

REFERENCES

Bhattacharya, S., and Gale, D. (1987). Preference shocks, liquidity, and central bank policy, *in* "New Approaches to Monetary Economics" (W. A. Barnett and K. J. Singleton, Eds.), pp. 69–88. Cambridge Univ. Press, Cambridge, UK.

Bhattacharya, S., and Padilla, A. J. (1996). Dynamic banking: A reconsideration, *Rev. Finan. Stud.* **9,** 1003–1032.

Bhattacharya, S., and Thakor, A. V. (1993). Contemporary banking theory, *J. Finan. Intermed.* **3,** 2–50.

Bryant, J. (1980). A model of reserves, bank runs, and deposit insurance, *J. Bank. Finan.* **4,** 335–344.

Diamond, D. W. (1997). Liquidity, banks, and markets, *J. Polit. Econ.* **105,** 928–956.

Diamond, D. W., and Dybvig, Ph. H. (1983). Bank runs, deposit insurance, and liquidity, *J. Polit. Econ.* **91,** 401–419.

Engineer, M. (1989). Bank runs and the suspension of deposit convertibility, *J. Monet. Econ.* **24,** 443–454.

Green, E. J. (1995). Illiquidity from a mechanism-design perspective, unpublished working paper, Federal Reserve of Minneapolis.

Hall, R. E. (1988). Intertemporal substitution in consumption, *J. Pol. Econ.* **96,** 339–357.

Haubrich, J. G., and King, R. G. (1990). Banking and insurance, *J. Monet. Econ.* **26,** 361–386.

He, H., and Pagès, H. F. (1993). Labor income, borrowing constraints, and equilibrium asset prices, *Econ. Theory* **3,** 663–696.

Hellwig, M. (1994). Liquidity provision, banking, and the allocation of interest rate risk, *Eur. Econ. Rev.* **38,** 1363–1390.

Jacklin, Ch. J. (1987). Demand deposits, trading restrictions, and risk sharing, in "Contractual Arrangements for Intertemporal Trade" (E. C. Prescott, and N. Wallace, Eds.), pp. 26–47. University of Minnesota Press, Minneapolis.

Jacklin, Ch. J. (1993). Market rate versus fixed rate demand deposits, *J. Monet. Econ.* **32,** 237–258.

Judd, K. L. (1985). The law of large numbers with a continuum of IID random variables, *J. Econ. Theory* **35,** 19–25.

Postlewaite, A., and Vives, X. (1987). Bank runs as an equilibrium phenomenon, *J. Polit. Econ.* **95,** 485–491.

Qi, J. (1994). Bank liquidity and stability in an overlapping generations model, *Rev. Finan. Stud.* **7,** 389–417.

von Thadden, E. L. (1996). Intermediated versus direct investment: Optimal liquidity provision and dynamic incentive compatibility, unpublished working paper, Université de Lausanne.

von Thadden, E. L. (1997). The term structure of investment and the banks' insurance function, *Eur. Econ. Rev.* **41,** 1355–1374.

Wallace, N. (1988). Another attempt to explain an illiquid banking system: The diamond and dybvig model with sequential service taken seriously, *Quart. Rev. Federal Reserve Bank of Minneapolis*, Fall 1988, 3–16.

Wallace, N. (1990). A banking model in which partial suspension is best, *Quart. Rev. Federal Reserve Bank of Minneapolis*, Fall 1990, 11–23.