Security Design: The Demand Side View

12

Consumer Liquidity Demand

12.1 Introduction

As studied in Chapter 5, corporations and financial intermediaries secure their liquidity on the asset side of their balance sheet through lines of credit and the hoarding of liquid assets. They also manage their liquidity on the liability side. Short-term debt drains liquidity much more than long-term debt or securities, such as preferred equity and common equity, that embody a valuable option of not being forced to pay (preferred or common equity) dividends if times get rough.

By assuming that investors' utility is represented by the present discounted value of their consumptions (with a discount rate normalized at 0), we have ignored their own liquidity demand. In practice, consumers face personal shocks and value the flexibility of being able to realize their assets when they need to. For example, ignoring differences in rates of return, they value demand deposits over and above savings that are locked in for a few months or years.[1] They hoard substantial amounts of liquid assets in order to insure against shocks. They are willing to sacrifice returns in order to make sure they will have enough money to buy a house or a car when the opportunity arises, to send their children to (more or less expensive) college, or to protect themselves against illness or unemployment. Thus, consumers compete with corporations for the available stock of liquidity.[2] Consumer liquidity demand has been the focus of a large and interesting literature, starting with the seminal papers of Bryant (1980) and Diamond and Dybvig (1983).

This chapter looks at three aspects of consumer liquidity demand:

(i) The role that financial institutions may play as (a) liquidity pools and (b) insurers. We will see that the first role, which prevents the waste associated with individual securing of liquidity (duplication of costly liquidity provision) is primordial, while the second, which aims at flattening the term structure of interest rates in order to reduce the cost of impatience, is more fragile as it is exposed to opportunistic arbitrage by financial markets.

(ii) The runs that may occur on financial intermediaries with (efficiently) limited liquid assets on their balance sheet.

(iii) The design of a menu of securities that fits the individual profiles—short term versus long term—of investors.

12.2 Consumer Liquidity Demand: The Diamond–Dybvig Model and the Term Structure of Interest Rates

12.2.1 Insuring against Liquidity Shocks

The Diamond–Dybvig model depicts the optimal contract between a financial intermediary and a consumer who faces uncertainty as to the timing of her consumption. The model, in its simplest and most common form, has three periods, $t = 0, 1, 2$.[3]

Consumer preferences. Consumers are *ex ante* identical. For notational simplicity, let us assume that they have no demand for consumption at date 0, and therefore invest their entire date-0 resources, 1 per consumer. More generally, their savings are

---

1. Unless they are worried about a time-inconsistency problem and do not want to be exposed to the temptation to consume (see, for example, Laibson 1997).

2. This competition has been little studied in the literature unfortunately.

3. Good alternative expositions of the Diamond–Dybvig model can be found in Bond and Townsend (1995), Freixas and Rochet (1997), and Gorton and Winton (2003).
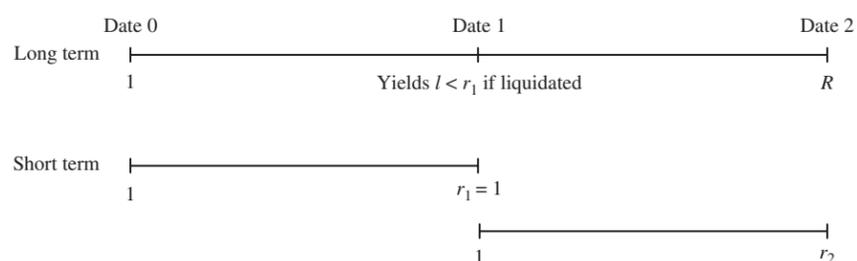
Figure 12.1 Liquid and illiquid investments.

equal to 1 per consumer. They have no further resources at dates 1 and 2, and have state-contingent preferences over date-1 and date-2 consumptions, c_1 and c_2, given by

$$\left.\begin{array}{ll} u(c_1) & \text{if impatient (probability } \lambda), \\ u(c_2) & \text{if patient (probability } 1 - \lambda), \end{array}\right\} \quad (12.1)$$

where the function u is increasing and *strictly* concave, and $u'(0) = \infty$. Consumers do not know at date 0 whether they will be impatient ("face a liquidity shock") or not. They learn at date 1 their "type" (patient or impatient).[4] In the simplest version of this model there is no aggregate uncertainty, and so exactly a fraction λ of consumers will want to consume at date 1.

The specification of consumer preferences embodied in (12.1) is a simple-minded way of formalizing the idea that the consumer does not know when she will need money.[5]

Technology. Date-0 resources are invested in short-term (liquid) and long-term projects. Short-term projects yield r_1 at date 1 per unit of date-0 investment. Similarly, 1 unit of investment (in a short-term project) at date 1 yields r_2 at date 2. Long-term projects yield $R > 1$ at date 2 per unit of date-0 investment, and nothing at date 1. Liquidity is costly as long-term projects have a higher yield:

$$r_1 r_2 < R.$$

In words, an investor with a long-term perspective, that is, one who would be unconcerned by the pos-

sibility of a liquidity shock ($\lambda = 0$), would invest in the long-term asset rather than in a short-term asset that she would roll over at date 1 (see Figure 12.1).

Without loss of generality, let us assume that

$$r_1 = 1.$$

This production function defines a *technological yield curve*. Let r_{LT} denote the per-period return on the long-term asset:

$$(1 + r_{LT})^2 = R \quad \text{or} \quad r_{LT} = \sqrt{R} - 1.$$

In comparison, a short-term investment at date 0 yields rate of interest

$$r_{ST} = r_1 - 1 = 0 < r_{LT}.$$

The technological yield curve is upward sloping.

Liquidating a long-term investment at date 1 yields a salvage value l per unit of date-0 investment. In this section, we will assume for simplicity that this salvage value is equal to 0, but more generally it will be assumed to be lower than $r_1 = 1$ (if $l \geqslant 1$, the long-term asset dominates and there is never any investment in short-term assets).

Lastly, the representative investor must decide how to allocate her savings between short-term investment i_1 and long-term investment i_2:

$$i_1 + i_2 = 1.$$

12.2.2 Self-Provision of Liquidity Is Inefficient

As is the case for corporate liquidity demand and for a related reason, self-provision of liquidity—a consumer's investing in liquid assets solely to cover her own liquidity shock—is wasteful. If the consumer happens not to face a liquidity shock, the costly liquidity that she has hoarded is wasted. Somehow,

---

4. These types are also called "late dier" and "early dier" in the literature.

5. For example, most of the insights derived below still hold if the consumer has utility $u(c_1 + c_2)$ when patient, provided that the rate of return between dates 1 and 2 (called r_2 below) exceeds 1: $r_2 \geqslant 1$.

the community of consumers should be able to use the law of large numbers to reduce their investment in liquid assets while enjoying the same amount of liquidity.

Autarky. To demonstrate the inefficiency of an "autarky situation" (as it is called in the literature), suppose that, as announced earlier,

$$l = 0,$$

and that the representative consumer invests i_1 and i_2 in short- and long-term assets. Then, because $r_1 = 1$,

$$c_1 = i_1 \quad \text{and} \quad c_2 = r_2 i_1 + R i_2 = R - c_1 (R - r_2).$$
(12.2)

As the consumption c_1 when the depositor is impatient grows from 0 to 1, the date-2 consumption enjoyed by the patient incarnation falls from R to r_2.

The representative consumer therefore maximizes her expected utility:

$$\max_{\{c_1\}} \{\lambda u(c_1) + (1 - \lambda) u(R - c_1 (R - r_2))\}. \quad (12.3)$$

Indeed, the optimization with respect to $\{i_1, i_2\}$ can be reduced to one over the date-1 consumption, since c_1 determines the investment in short-term assets needed and therefore the investment in long-term assets as well. This optimization yields either an interior or a corner solution:

$$\left. \begin{array}{ll} \text{either} & \dfrac{\lambda u'(c_1)}{(1 - \lambda) u'(c_2)} = R - r_2 \\[3mm] \text{or} & c_1 = 1 \quad \left(\text{if } \dfrac{\lambda u'(1)}{(1 - \lambda) u'(r_2)} > R - r_2\right). \end{array} \right\}$$
(12.4)

It is easy to check that the fraction invested in liquid assets $i_1 = c_1$ grows when the probability λ of facing a liquidity shock increases and when the "technological premium" associated with a long-term investment, $R - r_2$, decreases. Note, in particular, that everything is invested in the short-term asset when this premium is low or when the probability of a liquidity shock is high.

The benefits from pooling liquidity: mutual funds. The autarky outcome studied above precludes any resale. We have emphasized the excessive investment in liquid assets (possibly to the level of the consumer's entire savings). The flip side of the same

coin is that any investment in the long-term asset is thrown away (as $l = 0$) when the consumer turns out to be impatient. Somehow the opening of date-1 resale markets should generate gains from trade. Long-term assets held by impatient consumers are very attractive to patient consumers, who could use their own liquid assets to purchase the impatient consumers' long-term assets.

Along these lines let us show that a mutual fund enables consumers to enjoy the same date-1 consumption as under autarky, and a much larger date-2 return. Let $(\bar{c}_1, \bar{c}_2)$ denote the consumptions under autarky (they solve (12.2) and (12.4)). Let the consumers invest in a mutual fund with short- and long-term investments:

$$i_1 = \lambda \bar{c}_1 \quad \text{and} \quad i_2 = 1 - \lambda \bar{c}_1.$$

That is, they invest less in short-term assets and more into long-term ones than under autarky. This mutual fund distributes dividends equal to $i_1 r_1 = i_1$ at date 1, and $i_2 R$ at date 2. At date 1, the impatient consumers resell their share of the mutual fund to patient consumers, who compete with their scarce resources (the dividends they receive at date 1) for this valuable asset. The date-1 mutual fund price p is such that the patient consumers' resources, $(1-\lambda) i_1$, equal the value of the shares sold by the impatient ones:

$$(1 - \lambda) i_1 = \lambda p.$$

The impatient consumers then consume

$$c_1 = i_1 + p = \frac{i_1}{\lambda} = \bar{c}_1.$$

At date 2, the patient consumers each consume

$$c_2 = \frac{i_2 R}{1 - \lambda},$$

since they end up holding not only their initial shares, but also the shares of the impatient consumers and therefore own $1/(1 - \lambda)$ shares of the fund each. It is easily checked that

$$c_2 = \frac{[1 - \lambda \bar{c}_1] R}{1 - \lambda} > \bar{c}_2 = R - \bar{c}_1 (R - r_2).$$

These computations show that the resaleability of assets allows consumers to economize on liquidity provision and thereby increases their welfare.

Let us next compute the mutual fund's optimal portfolio: because

$$i_1 = \lambda c_1 \quad \text{and} \quad i_2 = \frac{(1-\lambda)c_2}{R}, \qquad (12.5)$$

the optimal portfolio solves

$$\max_{\{c_1\}} \left\{ \lambda u(c_1) + (1-\lambda)u\left(\frac{(1-\lambda c_1)R}{1-\lambda}\right) \right\}, \quad (12.6)$$

yielding

$$\frac{u'(c_1)}{u'(c_2)} = R. \qquad (12.7)$$

Note that the optimal mutual fund does not fully insure consumers against liquidity shocks ($c_1 < c_2$). It is optimal to take advantage of the upward-sloping yield curve and sacrifice some insurance. We will come back to this point shortly.

Comparison with corporate liquidity demand. To sum up, and as we have already noted, an important analogy between corporate liquidity demand (Chapter 5) and consumer liquidity demand (this chapter) is that corporations and consumers alike must obtain some insurance against liquidity shocks. Such insurance is costly when long-term investments have higher returns than short-term ones. Accordingly, liquidity ought to be hoarded sparingly and dispatched properly. When shocks are not perfectly correlated among economic actors (corporations, consumers), liquidity can be pooled and fewer low-yield investments are needed in comparison with the situation in which these actors self-provide liquidity. Or, put differently, autarky results in an overprovision of liquidity.

Consumer and corporate liquidity demands, however, differ in at least two respects:

• A key theme of corporate liquidity demand is that investments in short- and long-term assets, while competing for scarce resources at date 0, are later on *complements*, as liquidity enables long-term assets to bear their fruits. There is no such complementarity in the consumer liquidity demand model.

• *A consumer* consumes the cash that she receives, and *does not create any pledgeable income* (in the notation of Chapter 5, $\rho_0 = 0$). This observation has several consequences. First, the consumers' total investment is equal to their savings or "cash on hand" ($i_1 + i_2 = A = 1$ here), while firms can invest

more than their cash on hand ($i_1 + i_2 > A$).[6] In particular, the only way for consumers to satisfy their liquidity needs is to invest in real, low-yield, short-term assets. By contrast, Chapter 15 will show that, under some circumstances, the private sector may create enough "inside liquidity" and avoid having to invest in low-yield assets.

12.2.3 Optimal Liquidity Insurance

The mutual fund is only one of many ways available for pooling liquidity. Another familiar financial institution through which consumers pool their liquidity is the bank. Demand deposits allow consumers to choose the timing of withdrawals. A bank, of course, does not hold an amount of liquid assets equal to the level of demand deposits. Rather, it uses the law of large numbers to economize on liquid assets, as it knows that only a fraction of consumers will withdraw their deposits at any point in time.

More generally, one may wonder about the nature of the "optimal insurance scheme." The first point to note is that it is optimal to *match* the maturities of investments and consumptions. Given that there is no aggregate uncertainty and so one can predict exactly the levels of investment that are needed for date-contingent consumptions, investing $i_1 > \lambda c_1$ and rolling over the unneeded income ($i_1 - \lambda c_1$) is dominated by investing "just what is needed" for date-1 consumption ($i_1 = \lambda c_1$) and investing the rest in the higher-yield long-term asset. And so (12.5) holds.

---

6. For instance, the two-shock model of Section 5.3.1 can be rewritten by adapting the notation slightly to facilitate the comparison with the consumer liquidity demand. Recall that the entrepreneur chooses investment scale I that, if the liquidity shock is met at date 1, yields total income $\rho_1 I$ and pledgeable income $\rho_0 I$ (with $\rho_0 < \rho_1$). With probability λ, the firm must pay $x\rho I$ to salvage a fraction x of its assets. With probability $1 - \lambda$, it faces no shock at the intermediate stage. Letting $i_2 = I$ and $i_1 = \lambda x \rho i_2$, the breakeven and NPV conditions were given by

$$i_1 + i_2 - A = [\lambda x + (1-\lambda)]\rho_0 i_2$$

and

$$U_b = [\lambda x + (1-\lambda)]\rho_1 i_2 - (1 + \lambda x \rho)i_2.$$

Recall from Chapter 5 that, at the optimum, $x = 1$ if $\rho(1-\lambda) < 1$ and $x = 0$ if $\rho(1-\lambda) > 1$. In particular, i_1 and i_2 may both be positive, while in the optimal mutual-fund policy of the risk-neutral version of Diamond–Dybvig (the consumer's expected utility is $\lambda c_1 + (1-\lambda)c_2$), $i_2 = 1$ as long as $R > 1$ (concave versions of production technologies can also be studied so as to facilitate the comparison with the Diamond–Dybvig model with risk-averse consumers).

The optimal allocation must then solve (12.6), yielding, again, a solution characterized by (12.7). Let (c_1^*, c_2^*) solve (12.7) and $c_2^* = (1 - \lambda c_1^*)R/(1 - \lambda)$.

Implementation by a deposit contract. The optimal allocation can be implemented by a bank deposit contract provided that the rate of interest received by the consumer on this deposit depends on the date at which she withdraws. Namely, the consumer receives rates of interest r_{ST}^* and r_{LT}^* on deposits withdrawn at dates 1 and 2, such that

$$1 + r_{ST}^* = c_1^* \quad \text{and} \quad (1 + r_{LT}^*)^2 = c_2^*.$$

Let us follow Diamond and Dybvig (1983) and most of the subsequent literature in assuming that the consumers' coefficient of relative risk aversion exceeds 1:

$$\left| \frac{c u''(c)}{u'(c)} \right| > 1 \quad \text{for all } c.$$

This assumption is empirically reasonable (see, for example, Gollier 2001, Chapter 2). Equation (12.7),

$$\frac{u'(c_1^*)}{u'(c_2^*)} = R,$$

can then be shown to imply that[7]

$$1 < c_1^* < c_2^* < R. \tag{12.8}$$

We have

$$r_{ST}^* > r_{ST} \quad \text{and} \quad r_{LT}^* < r_{LT}.$$

In words, *the optimal insurance scheme flattens the yield curve relative to the technological yield curve.*

Note that, while the optimal insurance scheme flattens the yield curve relative to the technological one, no prediction can be made concerning its slope. If risk aversion is low (the coefficient of relative risk aversion is close to 1), the yield curve is close to the technological yield curve and is therefore upward sloping. In contrast, if risk aversion is very high (the coefficient of relative risk aversion goes to infinity), then consumptions at the two dates are almost equalized and so the yield curve is downward sloping (the interest on long-term deposits is

compounded and yet does not exceed the short-term deposit interest rate).

We have not yet wondered about whether this deposit contact is "incentive compatible." For example, would the patient consumers not want to withdraw at date 1 and reinvest the proceeds in the date-1 short-term technology yielding r_2? Indeed if $r_2 > 1$, and risk aversion is large, then $c_1^* r_2 > c_2^*$ from our previous analysis, and so it is indeed in the interest of patient consumers to feign impatience, cash out, and reinvest. Let us therefore assume at this stage that the bank is able to observe who is patient and who is not, or, equivalently, is able to prevent reinvestment elsewhere. This assumption is unrealistic, especially in a decentralized market economy, but it has the pedagogical merit of separating insurance concerns from incentive compatibility issues in a first step. Let us be "patient" and delay the discussion of incentive compatibility for a more general treatment in the next section.

More general preferences: suboptimality of mutual funds (advanced). The equivalence between mutual funds and demand deposits breaks down for more general specifications of preferences. Suppose with Jacklin (1987) that the representative consumer's preferences are more generally given by

$$u^{I}(c_1^I, c_2^I) \quad \text{with probability } \lambda \text{ (impatient)},$$
$$u^{P}(c_1^P, c_2^P) \quad \text{with probability } 1 - \lambda \text{ (patient)}.$$

To make sense of the terminology, one can imagine that the impatient type has a higher marginal rate of substitution between date-1 and date-2 consumptions $((\partial u / \partial c_1)/(\partial u / \partial c_2))$ than the patient type.

Ignoring again incentive compatibility questions, the optimal allocation then chooses investments and consumptions so as to solve

$$\max_{\{c_1^I, c_2^I, c_1^P, c_2^P\}} \{\lambda u^{I}(c_1^I, c_2^I) + (1 - \lambda) u^{P}(c_1^P, c_2^P)\}$$

s.t. $\qquad\qquad\qquad\qquad\qquad\qquad$ (12.9)

$$[\lambda c_1^I + (1 - \lambda)c_1^P] + \frac{[\lambda c_2^I + (1 - \lambda)c_2^P]}{R} = 1,$$

since $i_1 = \lambda c_1^I + (1 - \lambda)c_1^P$ is needed to deliver the total date-1 consumption and $i_2 = [\lambda c_2^I + (1 - \lambda)c_2^P]/R$ is what it takes to deliver the total date-2 consumption.

---

7. To show this, note that the assumption on the coefficient of relative risk aversion says that the function $c u'(c)$ is decreasing. Hence, $R u'(R) < 1 \cdot u'(1)$, and so at the feasible allocation $\{c_1 = 1, c_2 = R\}$, $u'(c_1)/u'(c_2) = u'(1)/u'(R) > R$. To obtain (12.8), one must increase c_1 above 1, and concomitantly reduce c_2 below R. To conclude, recall that $R > 1$ and (12.7) imply that $c_1^* < c_2^*$.

At the optimal allocation, marginal utilities are equalized across types:

$$\frac{\partial u^{\mathrm{I}}}{\partial c_1^{\mathrm{I}}} = \frac{\partial u^{\mathrm{P}}}{\partial c_1^{\mathrm{P}}} \quad \text{and} \quad \frac{\partial u^{\mathrm{I}}}{\partial c_2^{\mathrm{I}}} = \frac{\partial u^{\mathrm{P}}}{\partial c_2^{\mathrm{P}}}.$$

Furthermore,

$$\frac{\partial u^{\theta}}{\partial c_1^{\theta}} \bigg/ \frac{\partial u^{\theta}}{\partial c_2^{\theta}} = R \quad \text{for } \theta \in \{\mathrm{I, P}\}.$$

In contrast, a mutual fund mechanism equalizes only marginal rates of substitution: if p denotes the price (in terms of date-1 consumption) of shares in the date-2 dividend, then each type $\theta \in \{\mathrm{I, P}\}$ faces a date-1 budget constraint,

$$c_2^{\theta} - i_2 R = (i_1 - c_1^{\theta}) \left(\frac{i_2 R}{p} \right),$$

and maximizes $u^{\theta}(c_1^{\theta}, c_2^{\theta})$ subject to this constraint. Thus marginal rates of substitution are equalized:

$$\frac{\partial u^{\mathrm{I}}}{\partial c_1^{\mathrm{I}}} \bigg/ \frac{\partial u^{\mathrm{I}}}{\partial c_2^{\mathrm{I}}} = \frac{\partial u^{\mathrm{P}}}{\partial c_1^{\mathrm{P}}} \bigg/ \frac{\partial u^{\mathrm{P}}}{\partial c_2^{\mathrm{P}}}.$$

But, in general, the mutual fund scheme contains no mechanism to redistribute across types. The consumer enters date 1 with the same budget (dividend plus resale value) regardless of her type. This insurance shortage must be remedied through a different scheme, in which the consumer gets the solution to (12.9), $(c_1^{\mathrm{I}}, c_2^{\mathrm{I}})$ when impatient and $(c_1^{\mathrm{P}}, c_2^{\mathrm{P}})$ when patient. Assuming $c_1^{\mathrm{I}} > c_1^{\mathrm{P}}$ and $c_2^{\mathrm{I}} < c_2^{\mathrm{P}}$, this can be accomplished by a combination of long-term savings that are locked in until maturity and deliver c_2^{I} at date 2, together with a deposit contract that offers the option of withdrawing the total amount c_1^{I} at date 1 versus withdrawing the smaller amount c_1^{P} in exchange for return $[c_2^{\mathrm{P}} - c_2^{\mathrm{I}}]$ at date 2.

Even if we rule out reinvestments outside the bank offering such contracts, it is no longer clear that the optimal allocation is incentive compatible, that is, that type $\theta \in \{\mathrm{I, P}\}$ prefers $(c_1^{\theta}, c_2^{\theta})$ to $(c_1^{\theta'}, c_2^{\theta'})$ for $\theta' \neq \theta$ (while this created no difficulty with the more special preferences studied earlier). (Noninnocuous) conditions need to be imposed to guarantee that the optimal allocation is incentive compatible (see Jacklin 1987).

Interbank lending. As shown by Bhattacharya and Gale (1987), interbank lending performs a useful pooling function when banks suffer idiosyncratic shocks in their depositors' withdrawal rates. Thus, suppose that there are two *ex ante* identical banks. The fraction of impatient depositors will be high (λ_{H}) in one bank and low (λ_{L}) in the other. So there is no aggregate uncertainty. The average withdrawal rate is $\lambda = \frac{1}{2}(\lambda_{\mathrm{H}} + \lambda_{\mathrm{L}})$. But there is idiosyncratic risk: no one knows at date 0 which bank will face the high withdrawal rate.

The banks can reach the efficient outcome by granting each other credit lines. They invest $i_1 = \lambda c_1^*$ and $i_2 = (1 - \lambda) c_2^* / R$ per consumer each and redispatch the liquid asset between the two when the shocks accrue. The liquidity-poor bank (with withdrawal rate $\lambda_i = \lambda_{\mathrm{H}}$) can transfer some of the claim to the proceeds $i_2 R$ on its long-term investment to the liquidity-rich bank (with withdrawal rate $\lambda_i = \lambda_{\mathrm{L}}$) in exchange for $\frac{1}{2}(\lambda_{\mathrm{H}} - \lambda_{\mathrm{L}}) c_1^*$ at date 1.[8]

12.2.4 Financial Markets and the Jacklin Critique

A common theme in the economics of information and incentives is that markets conflict with the optimal provision of insurance (e.g., Pauly 1974; Helpman and Laffont 1975; Bernheim and Whinston 1986). Jacklin's (1987) critique of the Diamond–Dybvig model fits within this overall theme.

In a nutshell, Jacklin argues that financial markets' ability to arbitrage the implicit cross-subsidy in favor of the impatient relative to the technological yield curve undermines the overall insurance mechanism.

Suppose that a consumer initially bypasses the insurance system and invests her entire savings in the high-yield long-term asset ($i_2 = 1$). This strategy clearly delivers the highest possible payoff if the consumer turns out to be patient, since then

$$c_2 = R > c_2^*.$$

But what if the consumer ends up being impatient? The trick is then to sell the claim to the long-term payoff to the patient consumers, who use their ability to withdraw their deposits at the bank in order to finance the purchase. Normalize the number

---

8. The analysis by Bhattacharya and Gale (1987) is much broader than reported here. In particular, it also deals with situations in which banks are imperfectly informed about each other's solvency (investment in or return on the long-term assets, or the number of withdrawing depositors).

of shares issued by the consumer at one (divisible) share. A patient consumer can withdraw an amount c_1^* from the bank and is willing to pay price p per share for α shares such that

$$c_1^* = \alpha p,$$

as long as she gets at least as much consumption at date 2 as when she leaves her money at the bank:

$$\alpha R \geqslant c_2^*.$$

That is, the consumer who has invested in the long-term asset can obtain price p for the claim on this asset, such that

$$p = \frac{c_1^* R}{c_2^*} > c_1^*.$$

In effect, this opportunistic consumer free rides on the banks' costly provision of liquidity. She can have her cake and eat it too.

More generally, the same reasoning shows that any insurance scheme is undone by financial markets as long as $c_2 < R$. Hence, in the presence of financial markets, the best feasible allocation is

$$\hat{c}_1 = 1 \quad \text{and} \quad \hat{c}_2 = R. \tag{12.10}$$

Financial markets force the yield curve back to the technological yield curve. The reader will find in Allen and Gale (1997) useful complements on free riding and the underprovision of liquidity.

Remark (differential access to financial markets). Diamond (1997) studies the intermediate case in which some consumers have access to financial markets (as in Jacklin 1987) while others do not (as in Diamond and Dybvig 1983). Suppose, for instance, that everyone is *ex ante* identical. At date 1, the consumer learns her type. But there are now three types rather than two: an impatient type (receives c_1) and two patient types. In Diamond's terminology, those with access to financial markets are "type 2A," while those with no such access (who cannot reinvest the money they withdraw at date 1) are "type 2B." The bank is unable to tell the different types apart. The date-0 optimal contract offers return c_1 if the consumer withdraws at date 1 and c_2^B at date 2 if the consumer does not, where

$$1 < c_1 < c_2^B < R.$$

In equilibrium, patient consumers with no access to financial markets just consume c_2^B. Patient consumers with such access withdraw c_1 and reinvest in one of these long-term investment vehicles yielding R; they consume[9]

$$c_2^A = c_1 R > R.$$

The extent of flattening of the bank's yield curve relative to the technological yield curve then depends on the fraction of consumers with no access to financial markets. If this fraction is important, extensive cross-subsidies *à la* Diamond–Dybvig are doable; if not, then the bank must offer a steep yield curve, close to the technological yield curve.

12.2.5   Economizing on Liquidity by Rolling over Deposits

Let us ignore the Jacklin critique and address another potential enrichment of the Diamond–Dybvig model. By not describing the economy as an ongoing one, Diamond and Dybvig overestimate the need for low-yield liquid assets, at least in a relatively stationary context. The idea is that if investments by incoming generations of consumers (new investors) offset the disinvestments by earlier generations of investors facing liquidity needs, then no asset needs to be liquidated and everything can be invested in the high-yield long-term asset.

Following Qi (1994), consider an overlapping-generations (OLG) version of the Diamond–Dybvig model in which:

- a new generation ("generation t") invests its savings (1 per individual) at date t, and lives up to date $t + 2$;
- members of this generation learn at date $t + 1$ whether their utility function is $u(c_{t+1}^t)$ (probability λ) or $u(c_{t+2}^t)$ (probability $1 - \lambda$), where c_τ^t is generation t's consumption at date τ;
- the population is constant; and
- the technology is similar to that described above: 1 unit of "long-term investment" yields $R > 1$ two periods later; 1 unit of "short-term investment" yields 1 one period later.

---

9. One can envision that this arbitrage is enabled by financial entities that invest in the long-term asset and resell it at cost (1) at date 1 to these type-2A consumers.

Table 12.1 The OLG structure.

	t	$t+1$	$t+2$	$t+3$	$t+4$
Generation t	invests 1	$u(c_{t+1}^t)$ (prob. λ)	$u(c_{t+2}^t)$ (prob. $1-\lambda$)		
Generation $t+1$		invests 1	$u(c_{t+2}^{t+1})$ (prob. λ)	$u(c_{t+3}^{t+1})$ (prob. $1-\lambda$)	
Generation $t+2$			invests 1	$u(c_{t+3}^{t+2})$ (prob. λ)	$u(c_{t+4}^{t+2})$ (prob. $1-\lambda$)

Table 12.1 summarizes the timing.

Consider a bank that in *steady state* offers consumption profile {c_1 (for the impatient), c_2 (for the patient)} so as to maximize the depositors' expected utility:

$$\max\{\lambda u(c_1) + (1-\lambda)u(c_2)\}. \tag{12.11}$$

This bank needs not invest in low-yield short-term investments. At period $t+2$, say, it can employ the return R on the generation t's deposits invested in high-yield assets, to honor the deposit withdrawal by generation t's patient types and generation $t+1$'s impatient types. Thus, the budget constraint is

$$\lambda c_1 + (1-\lambda)c_2 \leqslant R. \tag{12.12}$$

Note that the maximization of (12.11) subject to (12.12) yields perfect insurance:

$$c_1 = c_2 = R.$$

This allocation, which exhibits a downward-sloping yield curve, however, is not incentive compatible if patient consumers can withdraw and reinvest in a similar bank (or the same bank under a different name). Such arbitrage indeed imposes that

$$(c_1)^2 \leqslant c_2. \tag{12.13}$$

That is, if the consumer can withdraw and reinvest, the yield curve must be either flat ($(c_1)^2 = c_2$) or upward sloping ($(c_1)^2 < c_2$). Given that the optimal yield curve in the absence of constraint (12.13) is downward sloping, *the constrained optimal yield curve is flat*:

$$(c_1)^2 = c_2,$$

which implies

$$c_2 > R > c_1 > 1.$$

This analysis requires that there be no aggregate uncertainty and that the economy be in a steady state. In particular, Qi (1994) looks at how a bank can get started. We refer to the paper for more detail.

While highly stylized, this OLG analysis captures an important aspect of reality. Banks make heavy use of the facts that demand deposits are rolled over, and that, to honor the promises made in previous deposit agreements, they can attract new deposits rather than liquidate their long-term assets. The same strategy plays an important role on the equity side as well. For example, the underlying assets in a closed-end mutual fund (whose shares are sold on the open market) are not liquidated when an investor wants to sell her share. Rather, this share is transferred to another investor.

Allen and Gale (1997, 2000, Chapter 6) analyze an OLG model with a safe and a risky asset. The safe asset can be accumulated over time. Financial markets allow cross-sectional risk-sharing opportunities to be exploited, but may provide insufficient intertemporal risk smoothing. An intermediated system fares better in the latter dimension. However, the intertemporal smoothing provided by a long-lived intermediary is fragile as arbitrage opportunities undermine the insurance it offers.

12.3 Runs

12.3.1 Depositor Panics

A substantial fraction of the literature on consumer liquidity demand, starting with Bryant (1980) and Diamond and Dybvig (1983), is preoccupied by the possibility of bank runs.[10] A basic hazard faced by financial institutions performing a maturity transformation function is the risk that depositors run for exit even when they do not actually experience liquidity needs. A run may occur when long-term assets are liquidated in order to honor the withdrawal demands. Thus, if other depositors withdraw, even a patient depositor has an incentive to withdraw since

---

10. Early analysis of bank runs can be found in Bagehot (1873) and Kindelberger (1978). Other useful references include Fulghieri and Rovelli (1998), Gale and Vives (2002), and Rochet and Vives (2004).

the financial institution then becomes an empty shell.

To understand the mechanics of bank runs, consider the technology described in the previous section, with

$$l = r_1 = r_2 = 1 \quad \text{and} \quad R > 1.$$

That is, a unit long-term investment yields R if carried to its maturity, but only 1 if it is liquidated at date 1. The short-term technology in each period is a storage technology that transforms 1 unit of good in a given period into 1 unit of good in the following period. The long-term investment here dominates the short-term investment, and we will therefore focus on investment policies in which the bank invests solely in the long-term asset:

$$i_1 = 0 \quad \text{and} \quad i_2 = 1.$$

The representative consumer, as before, saves 1 at date 0 and learns her type at date 1; with probability λ, the consumer is impatient and has utility $u(c_1)$, and with probability $1 - \lambda$, the consumer is patient and has utility $u(c_2)$. We assume that a patient consumer who withdraws at date 1 has access to the storage technology and can thus consume at date 2 what she withdrew at date 1.[11]

Consider the Diamond–Dybvig allocation (letting L denote the fraction of the long-term asset that is liquidated at date 1):

$$\max_{\{c_1, c_2, L\}} \{\lambda u(c_1) + (1 - \lambda) u(c_2)\}$$

s.t.

$$\lambda c_1 = L,$$

$$(1 - \lambda) c_2 = R(1 - L).$$

This program is equivalent to

$$\max_{\{c_1\}} \left\{ \lambda u(c_1) + (1 - \lambda) u \left(\left(\frac{1 - \lambda c_1}{1 - \lambda} \right) R \right) \right\},$$

yielding, as earlier,

$$\frac{u'(c_1)}{u'(c_2)} = R,$$

and so, provided that the consumers' coefficient of relative risk aversion exceeds 1,

$$1 < c_1^* < c_2^* < R.$$

---

11. Alternatively, we could assume that the patient consumer has utility $u(c_1 + c_2)$.
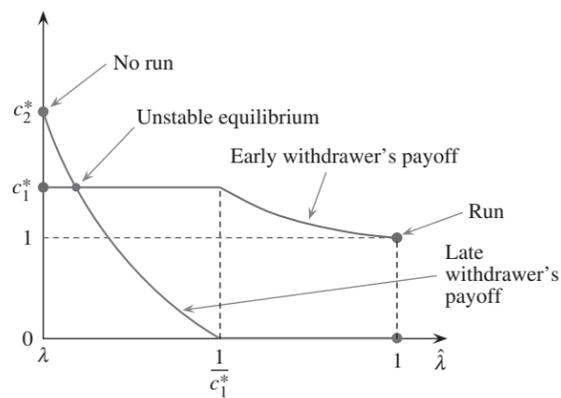
Figure 12.2 Incentive to run.

Let $\hat{\lambda} \geqslant \lambda$ denote the fraction of consumers who withdraw at date 1 (so $\hat{\lambda} = \lambda + (1 - \lambda)x$, where x is the fraction of patient consumers who run on the bank). Because $c_1^* < c_2^*$, the Diamond–Dybvig outcome $\hat{\lambda} = \lambda$ is an equilibrium. But this equilibrium is not unique.

A consumer receives

$$\min \left\{ c_1^*, \frac{1}{\hat{\lambda}} \right\} \qquad \text{if she withdraws at date 1,}$$

$$\max \left\{ \left(\frac{1 - \hat{\lambda} c_1^*}{1 - \hat{\lambda}} \right) R, 0 \right\} \quad \text{if she does not.}$$

To see this, note that the bank keeps liquidating long-term investments as long as it cannot honor the withdrawal requests. If $\hat{\lambda} c_1^* < 1$, then all such requests are satisfied, and the fraction $(1 - \hat{\lambda})$ of consumers who did not run receives the return R on the remaining long-term investment $(1 - \hat{\lambda} c_1^*)$, which is less than $c_2^* = [(1 - \lambda c_1^*)/(1 - \lambda)] R$.

The payoffs are represented as functions of $\hat{\lambda}$ for $\hat{\lambda} \geqslant \lambda$ in Figure 12.2.

An interesting property of the strategic interaction among depositors is that the incentive to run (the difference between the consumptions when withdrawing at date 1 and waiting) increases with the number of other consumers who withdraw (at least as long as $\hat{\lambda} < 1/c_1^*$, since beyond this value, a late withdrawer receives nothing anyway). This game exhibits "strategic complementarities" (my running increases your incentive to run). And indeed there is exactly one other stable equilibrium, in which all

consumers withdraw at date 1. This "bad or panic equilibrium" yields a low consumption for both the patient and the impatient types.[12]

Large depositor. Suppose now that a fraction $\mu > 0$ of deposits is held by a large depositor.[13] The fraction $1 - \mu$ is held by atomistic depositors (previously we had $\mu = 0$). Let us further assume that the large depositor has a single incarnation (patient or impatient), which, if we assume, as we will do, that the total fractions of "impatient and patient deposits" are fixed at levels λ and $1 - \lambda$, respectively, requires that $\mu \leqslant \min(\lambda, 1 - \lambda)$.[14] How is the analysis affected?[15]

Suppose first that the large depositor turns out to be impatient. Then the analysis is unaltered, since the only strategies of interest are those of the patient depositors, who face a real choice between withdrawing and leaving their deposits at the bank.

In contrast, the analysis is changed when the large depositor is patient. On the one hand, the no-run equilibrium still exists (since $c_1^* < c_2^*$). On the other hand, the panic equilibrium may disappear. A run can occur *only if* the large depositor does not find it in her interest to keep her money in the bank, or

$$1 \geqslant \frac{1 - \hat{\lambda} c_1^*}{1 - \hat{\lambda}} R \quad \text{with } \hat{\lambda} = 1 - \mu.$$

Put differently, the risk of a run disappears if

$$(1 - \mu)(Rc_1^* - 1) < R - 1.$$

In particular, for μ close to $(1 - \lambda)$ (most of the "patient deposits" are held by the large depositor), this latter condition is verified (from $c_1^* < c_2^*$), and so there is no panic equilibrium. More generally, *the panic equilibrium is less likely to exist, the larger the fraction of deposits held by a large player.* This is easily understood: panics are generated by a lack of coordination. This coordination problem is less likely to be an issue if deposits are concentrated in large

part in a single hand (it is no longer an issue with a single patient depositor).

12.3.2 Antirun Policies

As was recognized by Diamond and Dybvig and the subsequent literature, there are various ways to prevent bad equilibria from happening.

12.3.2.1 Suspension of Convertibility

One policy for preventing runs is a suspension of convertibility (Gorton 1985, 1988). Before the design of deposit insurance schemes, suspensions of convertibility occurred frequently. For example, the American banking system suspended convertibility eight times between 1814 and 1907.

The idea behind a suspension of convertibility is straightforward. Suppose that the bank announces that it will stop honoring demand deposit withdrawal once level λ is reached. Patient depositors then know that there will be enough long-term investment around at date 2 to honor their date-2 claim c_2^*. And so they have no incentive to run.[16]

Suspensions of convertibility are, of course, no panacea. They raise a moral-hazard problem on the bank's side. The run may actually be triggered by bad news about the bank's fundamentals (we will come back to this). In this case, the bank, if given the right to suspend convertibility may use this right to stop outflows even when its management, rather than a pure depositor panic, is the culprit for the run. This is why suspensions of convertibility are better entrusted to the central bank (or at the country level with the International Monetary Fund), even though these solutions are not without hazard either.

12.3.2.2 Credit Line and Lender of Last Resort

Second, the bank may have an explicit or implicit credit line with another financial institution or the central bank that protects it against a run. Again, if patient consumers know that long-term assets will not be forced to liquidation by a run, they have no reason to worry and therefore do not withdraw their deposits.

---

12. As indicated in the figure, there is a third equilibrium with $\lambda < \hat{\lambda} < 1$. This equilibrium is, however, unstable: suppose that a slightly higher fraction than $\hat{\lambda}$ withdraws. Then everyone else wants to withdraw.

13. To make things comparable, assume that the consumptions "c_t" of that depositor are consumptions per unit of deposit.

14. More generally, we could avoid this restrictive assumption, and assume that the large depositor suffers a liquidity shock corresponding to a (random) fraction of her deposits.

15. Large depositors are considered in Corsetti et al. (2002) and (in a version closer to that adopted here) Ventura (2001).

---

16. See Green and Lin (2003) and Peck and Shell (2003) for studies of more general contingent withdrawal contracts, in which the amount that can be withdrawn depends on the number of consumers who have already withdrawn.

Of course, in the case of a private sector arrangement, the credit line mechanism can protect only against a run on a single bank or a small number of banks. To avoid a run on a single bank, it suffices that each bank stand ready to liquidate a small amount of its long-term assets to come to the rescue of the endangered bank (or to hoard a little more liquidity than needed if $l < 1$).

However, such arrangements cannot protect the banking sector as a whole. If runs occur simultaneously on all banks, liquidity must be provided from elsewhere (the central bank or abroad).

12.3.2.3 Interbank and Other Liquidity Markets

Alternatively, banks can make up for temporary shortfalls in liquidity by borrowing liquidity in the interbank market. A solvent bank, with fully pledgeable income Ri_2 in the model, can credibly promise to repay any date-1 loan that is destined only to honor the deposit withdrawals.

While "bank runs" have a negative connotation and much thought has been given to how to avoid them, another strand of the literature, initiated by Calomiris and Kahn (1991), emphasizes the benefits of creating competition in monitoring. The possibility of a bank run keeps depositors (or, presumably, at least large ones) on their toes. They are then induced to collect information about the bank's performance. There is then a tradeoff between the inefficiency generated by liquidations and the disciplining benefit associated with the monitoring of banking moral hazard.[17]

12.4 Heterogenous Consumer Horizons and the Diversity of Securities

In the Diamond–Dybvig model, consumers are identical *ex ante* (although not *ex post*), and a single claim fits them all. In practice, consumers are heterogenous in several respects, including their savings horizon, or, to use the terminology of this chapter, the frequency of liquidity shocks. Gorton and Pennacchi (1990) provide an interesting extension of the Diamond–Dybvig model that allows for such heterogeneity.

Their study is motivated by the long-standing advice given by bankers to their clients: "If you save for the long term, invest in equities; if you are looking for liquidity, invest in debt instruments." The alleged "liquidity" benefits of debt in this advice does not quite refer to the possibility that equities cannot be resold quickly in well-functioning markets. Rather, it refers to the fear of trading against better-informed traders in such markets.

A useful innovation of the Gorton–Pennacchi model is to employ the consumer-liquidity-demand model to refine our understanding of market microstructure. In traditional models of markets microstructures (say, Kyle 1985), trade is driven by the presence of apparently irrational "liquidity traders" who trade assets without regard to their return. These liquidity traders generate value for the other traders and thereby give rise to trading volume.[18] The Diamond–Dybvig model allows the model to endogenize liquidity trading by explicitly modeling preference shocks that give rise to a demand for altering one's portfolio. The benefit of this "rationalization" of liquidity trading is not purely aesthetic. As we will see, it shows that liquidity trading in equities is highly responsive to the set of securities that are offered in the market.

The Gorton–Pennacchi model is similar to Diamond and Dybvig's, with two twists. First, the payoff of the long-term investment is uncertain and is not commonly observed at date 1. Second, the number of impatient consumers is also random and unobservable. In contrast, consumers are risk neutral, which eliminates the insurance focus that is so prominent in the Diamond–Dybvig literature.

There are three dates ($t = 0, 1, 2$).

Consumers. Consumers all have date-0 savings equal to 1, but are *ex ante* heterogenous with respect to their consumption horizon. More precisely, there are two categories of consumer.

---

17. See Chapters 8 and 9 for a discussion of the variety of ways in which incentives for monitoring can be designed.

18. Another approach is to assume that investors are risk-averse and learn over time news about their tastes or about the value of the components of their existing portfolios, and therefore want to rebalance these portfolios. This approach is much more complex (and depends on the set of futures and derivative markets allowed). Much of the microstructure literature therefore relies on the irrational-liquidity-traders approach.

Potential liquidity traders, in proportion α, have the following preferences:

$$u(c_1, c_2) = c_1 \qquad \text{with probability } \tilde{\lambda},$$
$$u(c_1, c_2) = c_1 + c_2 \quad \text{with probability } 1 - \tilde{\lambda}.$$

As in Diamond and Dybvig, these consumers learn their preferences at date 1; the realized fraction of liquidity traders, $\tilde{\lambda}$, takes two possible values, λ_L or λ_H, with $\lambda_H > \lambda_L$. The realization of $\tilde{\lambda}$ is unobservable.

Long-term investors, in proportion $1 - \alpha$, have the following preferences:

$$u(c_1, c_2) = c_1 + c_2 \quad \text{with probability } 1.$$

That is, long-term investors take a long-term perspective and never need money at date 1 (they are happy to get the return from their savings at date 2).

Technology. On the technology side, we will assume that the savings are invested in a long-term asset yielding a random $\tilde{R}$ at date 2, where $\tilde{R} = R_L$ or $R_H > R_L$. This long-term return is publicly observable only at date 2 (when realized).

States of nature. Let us now turn to the probability distribution over the state of nature $(\tilde{\lambda}, \tilde{R})$.[19] In principle, there are four possible states of nature as each of these variables can take on two values. To simplify the computations, we will make two innocuous assumptions. First, $\tilde{\lambda}$ and $\tilde{R}$ are perfectly correlated in the following way. There are only two states of nature:

$$(\lambda_L, R_L) \quad \text{with probability } q_L,$$
$$(\lambda_H, R_H) \quad \text{with probability } q_H,$$

with $q_L + q_H = 1$. Second, potential liquidity traders that are revealed patient do not have cash at date 1 to participate in the date-1 asset market. Only long-term investors (and possibly some newly arrived arbitrageurs, also with utility function $c_1 + c_2$) have date-1 resources to buy the shares sold by the impatient investors. (The second assumption is just meant to shorten the analysis by not having to consider the inferences drawn by the patient liquidity

traders about the state of nature from the observation that they individually are patient. The first assumption focuses the analysis on those two states of nature in which the asset price may not reveal publicly the state of nature. The reader can alternatively assume four states of nature and follow the lines of Section 8.3 to check that the analysis in no way hinges on these two assumptions.)

Speculator. To formalize the idea that small investors may "lose their shirt" when disposing of the asset at date 1, let us assume that an informed trader, called the speculator, appears at date 1, who learns the state of nature and may buy as many shares as he likes (he has a large enough date-1 endowment). The speculator cannot engage in short sales; neither can any other economic agent. The speculator also has preferences $c_1 + c_2$. He places at date 1 an order flow. The date-1 arbitrageurs (long-term investors or newly arrived arbitrageurs) observe only the *total* order flow, that is, the impatient investors' sales minus the speculator's purchase, but cannot decompose this order flow to figure out exactly how much is demanded by the speculator (otherwise they could infer the state of nature from the speculator's order flow, as we shall see shortly).

12.4.1 Trading Losses in the Stock Market

When informed that the state is L, the speculator knows that the long-term payoff is R_L and since the asset price P necessarily lies in the interval $[R_L, R_H]$, the speculator does not buy and so stays out of the market. The order flow is then equal to the impatient consumers' sales:

$$\alpha\lambda_L.$$

When learning that the state is H, the speculator buys a quantity $b > 0$ of shares. The dilemma facing the speculator is that a high demand reveals that the state is high, leading arbitrageurs to raise their own demand until the price is R_H and so there is no profit opportunity. More formally, the order flow is now

$$\alpha\lambda_H - b.$$

The only value of purchases by the speculator that does not reveal that the payoff is R_H is

$$b = \alpha(\lambda_H - \lambda_L).$$

---

19. Formally, the state of nature also includes the identity of those among potential liquidity traders who will face a liquidity shock. Because there is no aggregate uncertainty in this respect, we omit this description from that of the state of nature.

The equilibrium involves full pooling. Arbitrageurs learn nothing about the state of nature, and their posterior belief that the state is H is still q_H. And so the market price of shares at date 2 is always

$$P = q_H R_H + q_L R_L.$$

This pooling gives rise to adverse selection in the stock market. Arbitrageurs (who, after all, are not forced to trade) are not affected by this adverse selection, as they discount the price to reflect the asymmetry of information. The victims of adverse selection are the impatient consumers or liquidity traders, who sell at a price reflecting the *ex ante* expectation, even though the high state is more likely *per unit of sale* (the liquidity traders sell more in the high state).

The speculator makes (*ex ante*) expected profit,

$$\pi = q_H[\alpha(\lambda_H - \lambda_L)][R_H - P].$$

That is, the speculator trades only in the high state (probability q_H). He then trades as much as is consistent with not revealing his information ($\alpha(\lambda_H - \lambda_L)$), and makes profit $R_H - P$ per share purchased. The speculator's profit can be rewritten as

$$\pi = \alpha(\lambda_H - \lambda_L)q_H q_L(R_H - R_L).$$

Note, in particular, that this profit grows with the fraction of potential liquidity traders and with the uncertainty about the extent of their actual liquidity trading.

To confirm that the speculator feeds off the potential liquidity traders, let us compute the latter's expected loss:

$$
\begin{aligned}
q_H \lambda_H (R_H - P) &- q_L \lambda_L (P - R_L) \\
&= (\lambda_H - \lambda_L)q_H q_L(R_H - R_L) \\
&= \frac{\pi}{\alpha}.
\end{aligned}
$$

The speculator's profit is indeed equal to a potential liquidity trader's expected loss times the number (α) of such traders.

12.4.2 Debt as a Low-Information-Intensity Security and the Equity Premium

As in Chapter 8, the liquidity traders' loss can be interpreted as generating an equity premium. In order for potential liquidity traders to hold the stock,

they must be enticed by a date-0 price discount, or equivalently an equity premium (a higher return). There are (at least) two equivalent versions that can be offered for depicting this phenomenon in the context of this bare-bones model. First, potential liquidity traders demand to pay less than the expected return. Namely, the price discount per share is equal to π/α so that the issuer must price shares at $q_H R_H + q_L R_L - (\pi/\alpha)$ in order to arouse interest from liquidity traders. Second, were the stock sold solely to the long-term investors (which requires that they have enough savings to purchase all the shares), the price would jump by π/α to $q_H R_H + q_L R_L$.

This equity premium observation (which is not specific to the Gorton–Pennacchi model, and is rather a general implication of the logic of market microstructure) also fits well with the well-known fact that the return on equity grows with the holding length. As popular wisdom commands, the stock market is more appealing to long-term investors than to short-term ones.

Let us push the comparison with the analysis of Chapter 8 a bit further. Speculation (the acquisition of private information about returns in order to profit from trading securities) is here a purely *parasitical* activity. It is even socially wasteful if either the speculator incurs a cost (presumably smaller than π) in order to acquire the information, or if the potential liquidity traders are discouraged from buying the security because they will "lose their shirt" and do not find an alternative and substitutable security to invest in.

The perspective on speculation provided by Gorton and Pennacchi is therefore quite different from the Holmström and Tirole (1993) view exposited in Chapter 8. There, even though we stressed that there could be excessive speculation, we emphasized the benefits of market monitoring. We argued that speculators' greed creates a measure of the value of assets in place, and therefore allows firms to assess the performance of their management. In other words, market monitoring is an integral part of the firms' governance mechanism. We will later return to this discussion.

Returning to the Gorton–Pennacchi model, we observed that potential liquidity traders are willing to pay less than long-term investors for the shares. This

suggests that it is in the interest of the security designers to introduce a security that is better suited to their needs, and thereby offer a menu of securities. Indeed, suppose that

$$\alpha(q_H R_H + q_L R_L) \leqslant R_L. \qquad (12.14)$$

This condition is more likely to be satisfied when the projects have sufficient guaranteed income or collateral (R_L) and when there are few potential liquidity traders. The security designers[20] can then offer a fraction α (or more generally a fraction between α and $R_L/[q_H R_H + q_L R_L]$) of securities with *safe* payoff $q_H R_H + q_L R_L$ (or slightly less[21]) at date 2. The residual claim on the long-term projects is then sold to the public in equity shares. The safe debt security appeals to potential liquidity traders because it is not affected by adverse selection. Its final payoff is independent of the state of nature and is therefore common knowledge. Thus, as long as condition (12.14) holds, the equity premium, or equivalently the profit that can be enjoyed by an informed speculator, vanishes.

In contrast, if

$$\alpha(q_H R_H + q_L R_L) > R_L,$$

there are too many potential liquidity traders in the market to accommodate entirely with a safe claim. They must bear some of the risk and therefore the equity premium reappears.

12.4.3 A Broader Perspective

The issuance of debt illustrates a broader strategy already alluded to in Chapter 8: investors who may be forced to sell fear that they will be trading against better-informed players and try to avoid this likely loss by purchasing securities that are less exposed to this risk. This flight to low-information-intensity securities takes multiple forms, and debt is only one of these.

Another way of limiting costly trade with speculators is to buy bundles of indices on the grounds that they are less exposed to asymmetric information "thanks to the law of large numbers": stock index futures, closed-end mutual funds, real-estate investment trusts, etc. The general idea is that even though one may be poorly informed about the value of a particular firm, one is on average better informed about that of a bundle of firms as an overappreciation of a firm's value tends to be compensated by an underappreciation of another (see Subrahmanyam 1991; Gorton and Pennacchi 1993). This is easily illustrated in the context of "continuum of firms" with independent date-2 profit realizations. The per-firm *ex post* value of the index is then a deterministic $q_H R_H + q_L R_L$, and so potential liquidity traders can enjoy liquidity without any sacrifice in return.

There is some empirical support for this view. For example, the bid–ask spread (which in part measures the extent of the adverse-selection problem) for the index is about one-tenth of that in individual stocks. Furthermore, the spectacular development of index funds in the last two decades points to the benefits of such bundling.

This evolution toward debt and bundles of equity claim is privately rational for (at least short-term) investors. It is also socially desirable if one subscribes to the view of Gorton and Pennacchi. On the other hand, it also jeopardizes the role of financial markets as a monitoring device,[22] and therefore has potentially detrimental effects. The cost involved in turning companies public and in spinning off divisions to have them listed individually are evidence of a demand for market monitoring. An important research topic is therefore to combine the negative and positive aspects of market monitoring and to analyze whether the investors' private incentives will in the future affect the relevant tradeoffs.[23]

---

20. We can assume that these security designers correspond to the corporate entities that invest in the long-term projects. Alternatively, these corporate entities could issue just stocks, and financial markets could perform the repackaging of these stocks by stripping the debt component from the stocks and offering it as a safe debt derivative instrument. As long as financial markets are competitive and efficient, the initial stocks would not include an equity premium, due to the expectation of subsequent repackaging.

21. In order to make sure that the long-term investors are not attracted to buy the debt security.

22. At least of the speculative/passive type studied in Chapter 8; concerning *active* monitoring (see Chapter 9), index funds do have some influence as they are not swayed by business ties.

23. While this section has assumed that trading costs are governed by adverse selection in asset markets, another relevant consideration is the existence of transaction costs. Favero et al. (2005) analyze a Diamond-Dybvig model in which consumers can buy or sell at date 1 some assets with heterogeneous and exogenously determined transaction costs. Consumers can only trade the set of primary assets (and so, as in standard microstructure theory, consumers cannot economize on transaction costs by trading asset bundles or derivatives).

Supplementary Sections

12.5   Aggregate Uncertainty and Risk Sharing

The analysis of interest rates in Section 12.2 focused on the term structure and neglected the allocation of the interest rate risk in an economy by assuming that there was no aggregate risk. In practice, interest rate risk is a serious issue, and financial institutions have developed various instruments, such as interest rate swaps, to reallocate the risk among economic agents. Ultimately, someone—consumers, banks, corporations, or other agents—must bear the risk. A question confronting both the private sector and public policy (e.g., through the regulatory treatment of value at risk in banking institutions) is who should actually bear it.

To start analyzing interest rate risk, Hellwig (1994) extends the Diamond–Dybvig model to allow for an uncertain realization at date 1 of the date-2 return on short-term investment r_2. The randomness of r_2 is a metaphor for a more general uncertainty about the rate of return on new investments in the economy.[24]

Consumers' preferences are as described in Section 12.2: facing known probability λ of a liquidity shock, their expected utility is

$$E[\lambda u(c_1) + (1 - \lambda)u(c_2)],$$

where c_1 denotes the consumer's date-1 consumption in the state of nature in which she is impatient, c_2 the date-2 consumption when she is patient, and the expectation will refer to the impact of aggregate uncertainty on these consumptions. The utility function u's coefficient of relative risk aversion $(-cu''/u')$ exceeds 1.[25]

A consumer's date-0 savings, equal to 1, are allocated between the short- and long-term investments:

$$i_1 + i_2 = 1.$$

Technology is described as in Diamond and Dybvig except for the aggregate uncertainty about r_2. A unit of short-term (liquid) investment sunk at date 0 yields r_1 at date 1. A unit of short-term investment sunk at date 1 yields r_2 at date 2. The value of r_2 is publicly learned at date 1. A unit of long-term (illiquid) investment sunk at date 0 yields R at date 2, and $l < r_1$ if liquidated at date 1. To keep the model as closely related to Diamond and Dybvig as possible, let us assume that liquidating the long-term project never delivers a higher return that the long-term project itself:

$$lr_2 < R \quad \text{for all realizations of } r_2. \qquad (12.15)$$

The random variable r_2 is assumed to have a continuous distribution with support included in $[0, R/l)$.

12.5.1   Socially Optimal Insurance

The first-best outcome is a choice of investments i_1, i_2 and (r_2-contingent) consumptions c_1 and c_2 and liquidation level L solving

$$\max_{\{i_1, i_2, c_1(\cdot), c_2(\cdot)L(\cdot)\}} \{E[\lambda u(c_1) + (1 - \lambda)u(c_2)]\}$$

s.t.

$$\lambda c_1 \leqslant r_1 i_1 + lL \quad \text{for all } r_2,$$

$$(1 - \lambda)c_2 \leqslant R(i_2 - L) + r_2(r_1 i_1 + lL - \lambda c_1)$$
$$\text{for all } r_2,$$

$$i_1 + i_2 = 1,$$

$$0 \leqslant L \leqslant i_2 \quad \text{for all } r_2.$$

The first constraint expresses the fact that impatient consumers' consumption must be financed from the

---

---

return on the date-0 investment in liquid assets, plus, possibly, the liquidation of some of the long-term assets. As we will see, this constraint may or may not be binding. The second constraint says that the consumption of the patient consumers stems from the return on the unliquidated long-term asset, plus, possibly the leftover from period 1 reinvested at rate r_2; it is obviously binding since there is no point wasting resources.

We will say that there is *earmarking* (or matching of the term structures of investments and consumptions) if long-term investments serve to finance long-term consumption and short-term investments to finance short-term consumption:

$$c_1 = \frac{r_1 i_1}{\lambda} \quad \text{and} \quad c_2 = \frac{R i_2}{1 - \lambda}.$$

Note that, under an earmarking policy, returns on deposits are guaranteed; in other words, consumptions at dates 1 and 2 are *immunized* against interest rate shocks (they are not contingent on r_2).

Is it optimal to immunize depositors against interest rate risk? From the first-order conditions associated with the first-best program, it can be shown that it is never optimal to liquidate the long-term investment ($L = 0$). Intuitively, the liquid investment always yields more than the illiquid one when it comes to generating date-1 income.

If reinvestment takes place at date 1 (i.e., $r_1 i_1 > \lambda c_1$), the consumptions must solve the following *ex post* program (for a given r_2):

$$\max_{\{c_1, c_2\}} \{\lambda u(c_1) + (1 - \lambda) u(c_2)\}$$

s.t. $\qquad\qquad\qquad\qquad\qquad$ (12.16)

$$\lambda c_1 + \frac{(1 - \lambda)c_2}{r_2} = r_1 i_1 + \frac{R i_2}{r_2}.$$

We have written the constraint so as to highlight the role of $1/r_2$ as the relevant discount factor between dates 1 and 2 and the expression of the present discounted value of the endowment (on the right-hand side). Thus, if reinvestment occurs, then

$$u'(c_1) = r_2 u'(c_2). \qquad (12.17)$$

Insights. Let us now state the results and then give their interpretations and intuition:

(a) As depicted in Figure 12.3, earmarking is optimal for low-interest-rate realizations ($r_2 \leqslant r_2^*$) and
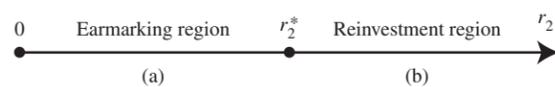
Figure 12.3 Incidence of interest rate risk: (a) $c_1 = r_1 i_1/\lambda$, $c_2 = R i_2/(1 - \lambda)$; (b) $c_1 < r_1 i_1/\lambda$, $c_2 > R i_2/(1 - \lambda)$, and $u'(c_1) = r_2 u'(c_2)$.

reinvestment for high-interest-rate realizations ($r_2 > r_2^*$).

(b) The level of investment i_1 in liquid assets at date 0 exceeds the level that would prevail if reinvestment opportunities did not exist or more generally always had a low return ("r_2 small").

To forge intuition about these results, let us begin with (a). One should think of the interest rate risk as creating an *option value* in this first-best world: if r_2 is large, then the date-1 consumption can be reduced in order to take advantage of the favorable reinvestment opportunities. This implies imposing some sacrifice on the impatient types to benefit the patient types. The impatient consumers are, of course, unhappy when the interest rate turns out to be high at date 1. But this is part of a deal a consumer is happy to accept at date 0. Conclusion (b) follows directly from the presence of an option value, which makes liquid investments more valuable.

12.5.2   Incentive Compatibility

Let us now assume more realistically that the patient consumers can feign impatience and reinvest their withdrawal elsewhere at rate r_2. The incentive to do so will, of course, depend heavily on the realization of r_2. A high-interest rate then becomes a double-edged sword. It offers investors an option value, but it also incentivizes them to behave opportunistically and to abuse the insurance deal.

The second-best solution is obtained by solving the first-best program to which is appended the incentive compatibility constraint:

$$r_2 c_1 \leqslant c_2 \quad \text{for all } r_2. \qquad (12.18)$$

This incentive compatibility condition creates a second rationale for a negative dependence of c_1 on r_2: generous terms on short-term deposits encourage opportunistic withdrawals. We refer to Hellwig's

paper for a full treatment of the second-best solution when r_2 exceeds 1 with probability 1. Two striking results (c) and (d) are as follows.

(c) In the reinvestment region, the impatient consumers bear the entire *valuation risk* of long-term investment while the patient consumers bear the entire *rollover risk* of short-term investment. In particular, no one is immunized against interest rate risk.

To see this, return to the first-best *ex post* program (12.16). We noted that the budget constraint corresponds to an income equal to the sum of the dividend on the liquid investment and the date-1 discounted dividend on the illiquid one. Condition (12.17) reflected a desire to provide some insurance to benefit the impatient type. This insurance, however, is undermined by the incentive compatibility condition: (12.17) together with $r_2 > 1$ implies that $c_1 < c_2$; the assumption that the coefficient of relative risk aversion exceeds 1 then implies that $c_1 u'(c_1) > c_2 u'(c_2)$ and so (12.17) yields $r_2 c_1 > c_2$. We thus conclude that (12.18) is binding:

$$r_2 c_1 = c_2.$$

This in turn implies that the two types can be given the same income $r_1 i_1 + (R i_2 / r_2)$ at date 1 (assuming there is no liquidation; the same is true if there is liquidation); and so

$$c_1 = r_1 i_1 + \frac{R i_2}{r_2} \quad \text{and} \quad c_2 = r_2(r_1 i_1) + R i_2. \quad (12.19)$$

These expressions make it clear that the impatient types fully bear the valuation (or execution) risk on the long-term asset and are hurt when the interest rate rises; conversely, the patient types fully bear the risk of rolling over the short-term return ($r_1 i_1$) and benefit from increases in the interest rate.

(d) Liquidation may become optimal.

To obtain a rough intuition as to why this may be the case, note that (12.18) suggests reducing c_1 and therefore first-period investment i_1. On the other hand, when r_2 is low, reinvestment does not pay off and incentive compatibility is not an issue. It may be optimal to increase c_1 beyond i_1 / λ by liquidating some of the long-term investment provided that l is not too low. Thus, liquidation, if it occurs

at all, is associated with low-interest-rate episodes. The second-best result according to which, in the reinvestment region, the impatient consumers bear the valuation risk and the patient ones the rollover risk extends to the case in which liquidation is optimal.

12.6 Private Signals and Uniqueness in Bank Run Models

As discussed in Section 12.3, a large literature in the last two decades has stressed the multiplicity of equilibria associated with deposit contracts. A recent and interesting strand of the literature, starting with Morris and Shin (1998), argues that the multiplicity tends to disappear provided that the economic agents receive private signals about the return to being patient and that their posterior beliefs have wide enough support.[26]

Morris and Shin's work is meant to address international financial crises.[27] As we will see, it captures some aspects of banking crises but not others. Like the bank run literature, it embodies a strategic complementarity: if other investors act in one way (say, run), that makes me more eager to act in that particular way (also run). But it also assumes that investors are better off when a run succeeds, while in banking models runs destroy the investors' value.

12.6.1 The Speculators' Game

Morris and Shin's stylized model of currency crises goes as follows: investors (also called speculators) can be thought of as being foreign investors. The central bank of a country has a level of foreign reserves θ unknown to investors. The central bank behaves mechanistically: it spends reserves to ward off speculation as long as there are some left. If S is the

---

26. Morris and Shin (1998) build their analysis on Carlsson and van Damme's (1993) work on global games. Earlier papers showing that private information together with a wide enough support allowing for the existence of dominant strategies, eliminates the multiplicity of equilibria in timing or bidding games include Fudenberg and Tirole (1986) on wars of attrition, Klemperer and Meyer (1989) on second-price auctions in supply-and-demand schedules, and Maskin and Riley (1986) on first-price auctions. See Frankel et al. (2003) for state-of-the-art results on unique strategy profiles surviving iterative dominance in games with strategic complementarities and slightly noisy payoff signals.

27. The version presented here is drawn from Corsetti et al. (2002).

Table 12.2 Payoffs in speculation game.

	Individual investor attacks	Individual investor does not attack
Attack succeeds ($S \geqslant \theta$)	$1 - c$	0
Attack fails ($S < \theta$)	$-c$	0

mass of financial resources mobilized by investors, then the currency collapses if and only if

$$S \geqslant \theta.$$

(This is meant to be a reduced form for a situation in which the country initially maintains a peg, and, if speculation is successful, the peg is abandoned and the currency devalued.) The level of investors' resources that can be mobilized to attack the currency is normalized to 1 (there is mass 1 of small investors[28]). And so $S \in [0, 1]$. In contrast, θ may exceed 1, in which case speculation against the currency is always unsuccessful, or be negative (say, because the country has contracted previous senior debts), in which case attacks always succeed regardless of their magnitude.

Assume that an investor individually pays a fixed cost $c \in (0, 1)$ when attacking the currency, and gains 1 when the attack is successful and he has been part of it.[29] The investors' decisions whether to attack the currency are simultaneous. An individual investor's contingent payoff is described in Table 12.2.

While the level of reserves θ is unknown, investors receive a signal as to its value. This signal y is equal to the true value plus noise:

$$y = \theta + \sigma \eta,$$

where η has mean 0 and σ measures (the inverse of) the precision of the signal. The variable η has cumulative distribution F with continuous density on, say, $(-\infty, +\infty)$.

In the public signal case, η is the same for all investors, who therefore have the same information.

28. See Corsetti et al. (2004) for the study of a similar game when there is a large investor.

29. In general, payoffs under successful and unsuccessful speculative attacks depends on the *ex post* exchange rate, which in turn depends on the size of the speculative attack and the government response to it. The speculation game may exhibit strategic complementarities or strategic substitutabilities (see Pathak and Tirole 2005).

In the private signal case, each investor receives his own signal; that is, η is i.i.d. across investors. (We could, of course, study the more general case in which investors receive both a public and a private signal. The results would be intermediate between those derived below.)

12.6.1.1 Public Signal

Under a public signal, the outcome resembles that in standard coordination games. There is a range $[\underline{y}, \bar{y}]$ of public signals for which there are multiple equilibria.

The *no-run* equilibrium exists provided that an individual investor does not find it profitable to attack the currency when others do not (and so the currency collapses only if $\theta \leqslant 0$):

$$(1 - c) \Pr(\theta \leqslant 0 \mid y) - c \Pr(\theta > 0 \mid y) \leqslant 0,$$

or

$$(1 - c)[1 - F(y/\sigma)] - cF(y/\sigma) \leqslant 0,$$

or else

$$F(y/\sigma) \geqslant 1 - c. \qquad (12.20)$$

Equation (12.20), taken as an equality, defines a unique $\underline{y}$. And so it is an equilibrium for no one to attack as long as $y \geqslant \underline{y}$.

Similarly, a *run equilibrium* exists provided that an individual investor prefers attacking when the others attack (and so the currency collapses whenever $\theta \leqslant 1$):

$$(1 - c) \Pr(\theta \leqslant 1 \mid y) - c \Pr(\theta > 1 \mid y) \geqslant 0$$

or

$$1 - c \geqslant F\left(\frac{y - 1}{\sigma}\right). \qquad (12.21)$$

Condition (12.21), taken as an equality, defines a threshold $\bar{y} > \underline{y}$, such that a run equilibrium exists if and only if $y \leqslant \bar{y}$.

Note that this "run equilibrium" cannot be called a "panic equilibrium." Indeed, when $y \in [\underline{y}, \bar{y}]$ investors are better off coordinating on an attack. In a sense, "panicking" corresponds to "staying put."

12.6.1.2 Private Signals

Let us now assume that investor i ($i \in [0, 1]$) receives signal

$$y_i = \theta + \sigma \eta_i,$$

and the noises are i.i.d. A simple "revealed preference" argument shows that in equilibrium investor i attacks the currency if and only if this signal lies below some threshold y_i^* (this is because the investor's net expected payoff to attacking the currency is decreasing in the signal). Let us look for a symmetric equilibrium (this is actually not restrictive): $y_i^* = y^*$.

The amount of resources involved in the attack is then

$$S(\theta) = F\left(\frac{y^* - \theta}{\sigma}\right),$$

and the currency collapses if and only if

$$S(\theta) \geqslant \theta.$$

Because S is decreasing in θ, the currency collapses if and only if $\theta \leqslant \theta^*$, where

$$F\left(\frac{y^* - \theta^*}{\sigma}\right) = \theta^*. \qquad (12.22)$$

Second, investor i attacks the currency if and only if

$$(1 - c)\Pr(\theta \leqslant \theta^* \mid y_i) - c\Pr(\theta > \theta^* \mid y_i) \geqslant 0.$$

And so y^* is defined by

$$1 - c = F\left(\frac{y^* - \theta^*}{\sigma}\right). \qquad (12.23)$$

Combining (12.22) and (12.23), we obtain

$$\theta^* = 1 - c. \qquad (12.24)$$

Thus, θ^* and y^* are uniquely determined. The uniqueness of equilibrium enhances predictive power.[30] When the investors' information is precise (σ close to 0), then y^* converges to θ^*.

12.6.2 The Depositors' Game

The bank run literature bears some resemblance to the analysis of the speculators' game in the previous subsection. But it differs from it in that runs are inefficient from the point of view of investors.[31] Another key difference is that, unlike the games considered in Carlsson and van Damme and by Morris and

Shin, the game does not quite exhibit strategic complementarities: as Figure 12.2 demonstrates, the net incentive to withdraw is not an increasing function of the number of other consumers who withdraw.

Let us return to the bank run model of Section 12.3, assuming that $l = 1$. Recall that if depositors are entitled to withdraw some arbitrary level $\bar{c}_1$ at date 1, and fraction $\hat{\lambda} \geqslant \lambda$ of depositors exercise this option, the consumptions of the early and late withdrawers are

$$c_1(\hat{\lambda}) = \min\left\{\bar{c}_1, \frac{1}{\hat{\lambda}}\right\}$$

and

$$c_2(\hat{\lambda}, R) = \max\left\{\frac{1 - \hat{\lambda}\bar{c}_1}{1 - \hat{\lambda}}R, 0\right\}.$$

Patient consumers have utility $c_1 + c_2$, and therefore choose the highest of the two. Let us extend the model of Section 12.3 in two respects:

- the date-2 return R is random and drawn from some cumulative distribution on $[0, \infty)$;
- this return is unobserved, but each depositor $i \in [0, 1]$ observes a private signal

$$y_i = R + \sigma\eta_i,$$

where the noises $\{\eta_i\}_{i \in [0,1]}$ have mean 0 and are i.i.d. across depositors; they are drawn from some cumulative distribution F with continuous density f.

We maintain the assumption that the bank offers a deposit contract, that is, the option to withdraw some fixed amount $\bar{c}_1$ at date 1.

A (symmetric) equilibrium is then defined by a threshold y^* such that depositor i, when patient, withdraws if and only if $y_i \leqslant y^*$, and a fraction $\lambda^*(y^*, R)$ of withdrawing depositors,[32] with

$$\lambda^*(y^*, R) = \lambda + (1 - \lambda)F\left(\frac{y^* - R}{\sigma}\right). \qquad (12.25)$$

It must also be the case that a depositor with signal y^* is indifferent between withdrawing and not withdrawing:

$$E[c_1(\lambda^*(y^*, R))] = E[c_2(\lambda^*(y^*, R), R)], \quad (12.26)$$

where expectations are taken with respect to the random variable R.

30. Angeletos et al. (2005) study a framework that is similar to that of Morris and Shin, but allow for a publicly observed policy choice by the policy maker before investors decide whether to attack. As in Morris and Shin, the equilibrium would be unique if the policy choice were exogenous. However, the endogeneity and observability of the policy reintroduce multiple equilibria in this model.

31. This distinction between the speculators' game and the depositors' game is drawn from Ventura (2001).

32. Thus $\lambda^*(y^*, R)$ is the counterpart of $S(\theta)$ in the speculators' game.

Goldstein and Pauzner (2005) analyze a related model (the technology succeeds or fails at date 2 and the probability of success is drawn from a continuous distribution). Their key insight is that while the depositor game does not exhibit strategic complementarities, it satisfies a weaker property (that they label "one-sided strategic complementarities"), namely, that the net incentive to withdraw increases with the number of withdrawing agents whenever this incentive is negative (see Figure 12.2).

They generalize the uniqueness result under this weaker property. They are then able to perform comparative statics exercises. For example, the probability of a bank run increases continuously with the degree of risk sharing offered by the intermediary.

12.7 Exercises

Exercise 12.1 (Diamond–Dybvig model in continuous time). Following von Thadden (1997), suppose that the representative consumer in the Diamond–Dybvig model has wealth 1 at date 0 and will need to consume at a time $t \in [0,1]$. Namely, the date of the liquidity shock, instead of taking two possible values (periods 1 and 2 in Diamond–Dybvig), belongs to an interval. It is distributed according to cumulative distribution function $F(t)$ ($F(0) = 0$, $F(1) = 1$) with continuous density $f(t)$. The representative consumer's expected utility is therefore

$$U = \int_0^1 u(c(t))f(t)\,\mathrm{d}t,$$

where $c(t)$ is her consumption if the liquidity shock occurs at time t.

On the technological side, suppose that one can at any point in time invest in "trees" that then grow until they are harvested. One unit of investment liquidated at maturity m yields $R(m)$. So an investment made at τ and "harvested" at $t \geqslant \tau$ yields $R(t - \tau)$ per unit. We assume that $R(0) = 1$, $\dot{R} > 0$ (where a dot indicates a time derivative), and $\dot{R}/R$, the instantaneous technological rate of return, is increasing in m. This implies in particular that a series of short-term investments yields less than a long-term investment with equivalent total length.

The choice is thus not about an allocation of investment at the initial date, and the exercise focuses entirely on the insurance aspects. Under autarky, the representative consumer receives expected utility

$$\int_0^1 u(R(t))f(t)\,\mathrm{d}t.$$

A bank offers a deposit contract in which a depositor chooses the date of withdrawal and obtains $c(t)$ if she withdraws at time $t \in [0,1]$. The depositors' liquidity shocks are i.i.d.

(i) Assume first that the realization of each depositor's liquidity shock is observable by the bank (so there is no incentive compatibility issue). Show that in the optimal insurance policy

$$u'(c(t))R(t)$$

is independent of t.

(ii) Assuming that the coefficient of relative risk aversion exceeds 1, conclude that there is "front loading,"

$$\frac{\dot{c}(t)}{c(t)} < \frac{\dot{R}(t)}{R(t)},$$

and so

$$c(t) > R(t) \quad \text{for } t < t^*$$

and

$$c(t) < R(t) \quad \text{for } t > t^* \text{ for some } t^* \in (0,1).$$

(iii) Show that the "first-best outcome" described above is not incentive compatible, in the sense that depositors may want to withdraw early and reinvest in the technology themselves.

Exercise 12.2 (Allen and Gale (1998) on fundamentals-based panics). Consider the Diamond–Dybvig model developed in Section 12.2 and add randomness in the payoff of the long-term asset. Consumers are Diamond–Dybvig consumers: they invest 1 at date 0, and learn at date 2 whether they are impatient (their utility is $u(c_1)$) or patient (their utility is $u(c_2)$). The probability of being impatient is λ.

The liquid or short-term technology yields one-for-one in each period: $r_1 = r_2 = 1$. The illiquid, long-term technology yields a random R (the same for all illiquid investments). The cumulative distribution is $F(R)$ and the density $f(R)$ on $[0, \infty)$. Liquidating the long-term asset yields nothing ($l = 0$).

One assumes
$$E(R) > 1.$$

The realization of R is publicly observed at date 1.

(i) Compute the socially optimal insurance contract $\{c_1(R), c_2(R)\}$, ignoring incentive compatibility (the ability of patient types to disguise as impatient ones). Note that this contract is incentive compatible.

(ii) Consider now a deposit contract. Consumers are promised, if they withdraw at date 1, a fixed payment $\bar{c}_1$, or a share of i_1 if total withdrawal demand exceeds i_1. The date-2 income is shared among depositors who did not withdraw at date 1. Long-term assets are never liquidated. One will denote by $x(R) \in [0, 1]$ the fraction of patient consumers who "join the run" (declare they are impatient, and store the money they have withdrawn from the bank).

Show that a judicious choice of $\bar{c}_1$ succeeds in implementing the social optimum described in (i).

Exercise 12.3 (depositors' game with a public signal). Consider the depositors' game of Section 12.6.2, except that the depositors receive the same signal:
$$y = R + \sigma\eta.$$

Determine the range of signals over which there exist multiple equilibria.

Exercise 12.4 (random withdrawal rate). Consider a three-date Diamond–Dybvig economy ($t = 0, 1, 2$). Consumers are *ex ante* identical; they save 1 at date 0. At date 1, consumers learn their preferences. A fraction λ has utility $u(c_1)$ and a fraction $(1 - \lambda)$ has utility $u(c_2)$.

At date 0, the consumers put their savings in a bank. They later cannot withdraw and invest in financial markets, so the Jacklin critique does not apply. That is, incentive compatibility issues are ignored in this exercise (a patient depositor cannot masquerade as an impatient one). The bank invests the per-depositor savings into short- and long-term projects: $i_1 + i_2 = 1$. The long-term technology yields (per unit of investment) $R > 1$ at date 2, but only $l < 1$ if liquidated at date 1. The short-term technology yields 1 (so $r_1 = r_2 = 1$).

(i) • Show that the optimal allocation (c_1, c_2) satisfies
$$u'(c_1) = Ru'(c_2).$$

• Suppose that $u(c) = c^{1-\gamma}/(1 - \gamma)$ with $\gamma > 1$. How do i_1 and i_2 vary with γ?

(ii) Suppose now that there is macroeconomic uncertainty, in that λ is unknown: $\lambda = \lambda_L$ with probability β and $\lambda = \lambda_H$ with probability $1 - \beta$, where $0 < \lambda_L < \lambda_H < 1$. Set up the optimal program (let y_ω and z_ω denote the fraction of short-term investment that is *not* rolled over, and the fraction of long-term investment that is liquidated, respectively, in state of nature $\omega \in \{L, H\}$). What does the solution look like for $l = 0$ and l close to 1? (Showoffs: characterize the solution for a general l!)

References

Acharya, V. V. and L. H. Pedersen. 2005. Asset pricing with liquidity risk. *Journal of Financial Economics* 77:375–410.

Allen, F. and D. Gale. 1997. Financial markets, intermediaries and intertemporal smoothing. *Journal of Political Economy* 105:523–546.

——. 1998. Optimal financial crises. *Journal of Finance* 53: 1245–1283.

——. 2000. *Comparing Financial Systems.* Cambridge, MA: MIT Press.

Angeletos, G. M., C. Hellwig, and A. Pavan. 2005. Coordination and policy traps. Mimeo, MIT, UCLA, and Northwestern University.

Bagehot, W. 1873. *Lombard Street: A Description of the Money Market.* London: H. S. King.

Bernheim, D. and M. Whinston. 1986. Common agency. *Econometrica* 54:923–942.

Bhattacharya, S. and D. Gale. 1987. Preference shocks, liquidity and central bank policy. In *New Approaches to Monetary Economics* (ed. W. Barnett and K. Singleton), pp. 69–88. Cambridge University Press.

Bond, P. and R. Townsend. 1995. Diamond–Dybvig models of banking. Teaching Notes for Money and Banking Course, University of Chicago.

Bryant, J. 1980. A model of reserves, bank runs, and deposit insurance. *Journal of Banking and Finance* 43:749–761.

Calomiris, C. and C. Kahn. 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81:497–513.

Carlsson, H. and E. van Damme. 1993. Global games and equilibrium selection. *Econometrica* 61:989–1018.

Corsetti, G., A. Dasgupta, S. Morris, and H. Shin. 2004. Does one Soros make a difference? The role of a large trader in currency crises. *Review of Economic Studies* 71:87–113.

Corsetti, G., P. Pesenti, and N. Roubini. 2002. The role of large players in currency crises. In *Preventing Currency Crises in Emerging Markets* (ed. S. Edwards and J. Frankel). NBER and Chicago University Press.

Diamond, D. 1997. Liquidity, banks, and markets. *Journal of Political Economy* 105:928–956.

Diamond, D. and P. Dybvig. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91:401–419.

Eisfeldt, A. 2003. Smoothing with liquid and illiquid assets. Mimeo, Northwestern University.

Favero, C., M. Pagano, and E. von Thadden. 2005. Valuation, liquidity and risk in government bond markets. Mimeo, Università Bocconi.

Frankel, D., S. Morris, and A. Pauzner. 2003. Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory* 108:1–44.

Freixas, X. and J. C. Rochet. 1997. *Microeconomics of Banking.* Cambridge, MA: MIT Press.

Fudenberg, D. and J. Tirole. 1986. A theory of exit in duopoly. *Econometrica* 54:943–960.

Fulghieri, P. and R. Rovelli. 1998. Capital markets, financial intermediaries, and liquidity supply. *Journal of Banking and Finance* 22:1157–1179.

Gale, D. and X. Vives. 2002. Dollarization, bailouts, and the stability of the banking system. *Quarterly Journal of Economics* 117:467–502.

Goldstein, I. and A. Pauzner. 2005. Demand deposit contracts and the probability of bank runs. *Journal of Finance* 60:1293–1327.

Gollier, C. 2001. *The Economics of Risk and Time.* Cambridge, MA: MIT Press.

Gorton, G. 1985. Bank suspension of convertibility. *Journal of Monetary Economics* 15:117–134.

——. 1988. Banking panics and business cycles. *Oxford Economic Papers* 40:751–782.

Gorton, G. and G. Pennacchi. 1990. Financial intermediaries and liquidity creation. *Journal of Finance* 45:49–71.

——. 1993. Security baskets and index-linked securities. *Journal of Business* 66:1–27.

Gorton, G. and A. Winton. 2003. Financial intermediation. In *Handbook of the Economics of Finance* (ed. G. Constantinides, M. Harris, and R. Stulz). Amsterdam: North-Holland.

Green, E. and P. Lin. 2003. Implementing efficient allocations in a model of financial intermediation. *Journal of Economic Theory* 109:1–23.

Hellwig, M. 1994. Liquidity provision, banking, and the allocation of interest rate risk. *European Economic Review* 38:1363–1390.

Helpman, E. and J. J. Laffont. 1975. On moral hazard in general equilibrium theory. *Journal of Economic Theory* 10:8–23.

Hirshleifer, J. 1971. The private and social value of information and the reward to inventive activity. *American Economic Review* 61:561–574.

Holmström, B. and J. Tirole. 1993. Market liquidity and performance monitoring. *Journal of Political Economy* 101:678–709.

Jacklin, C. 1987. Demand deposits, trading restrictions, and risk sharing. In *Contractual Arrangements for Intertemporal Trade* (ed. E. Prescott and N. Wallac), Chapter II, pp. 26–47. Minneapolis, MN: University of Minnesota Press.

Jacklin, C. and S. Bhattacharya. 1988. Distinguishing panics and information based bank runs: welfare and policy implications. *Journal of Political Economy* 96:569–592.

Kindelberger, C. 1978. *Manias, Panics, and Crashes.* New York: John Wiley.

Klemperer, P. and M. Meyer. 1989. Supply function equilibria in oligopoly under uncertainty. *Econometrica* 57:1243–1277.

Kyle, A. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–1335.

Laffont, J. J. 1985. On the welfare analysis of rational expectations equilibria with asymmetric information. *Econometrica* 53:1–29.

Laibson, D. 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112:443–478.

Maskin, E. and J. Riley. 1986. Uniqueness of equilibrium in sealed high bid auctions. Mimeo, MIT and UCLA. (Published in *Games and Economic Behavior* (2003) 45:395–409.)

Morris, S. and H. S. Shin. 1998. Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review* 88:587–597.

Pathak, P. and J. Tirole. 2005. Pegs, risk management, and financial crises. Mimeo, Harvard University and IDEI.

Pauly, M. V. 1974. Overinsurance and public provision of insurance: the roles of moral hazard and adverse selection. *Quarterly Journal of Economics* 88:44–62.

Peck, J. and K. Shell. 2003. Equilibrium bank runs. *Journal of Political Economy* 111:103–123.

Qi, J. 1994. Bank liquidity and stability in an overlapping generations model. *Review of Financial Studies* 7:389–417.

Rochet, J. C. and X. Vives. 2004. Coordination failures and the lender of last resort: was Bagehot right after all? *Journal of the European Economic Association* 2:1116–1147.

Subrahmanyam, A. 1991. A theory of trading in stock index futures. *Review of Financial Studies* 4:17–51.

Vayanos, D. 2004. Flight to quality, flight to liquidity, and the pricing of risk. National Bureau of Economic Research Working Paper 10327.

Ventura, J. 2001. Some thoughts on the role of large investors in currency crises: a comment to Corsetti et al. Mimeo, MIT.

Von Thadden, E. L. 1997. The term-structure of investment and the banks' insurance function. *European Economic Review* 41:1355–1374.