Memorial Lecture for John Kuszczak: Notes on Optimal Capital Regulation

*Douglas Gale**

Introduction

Financial crises have a long and varied history. The development of central banking techniques preceded the advent of modern economic theory, so it is not surprising that it relied more on empirical experience than on a micro-economic theory. What is surprising is that contemporary policy making still relies more on experience than theory. The Basel Accords, which impose capital-adequacy requirements on the banking systems of signatory countries, are a case in point. Practitioners have become expert at mastering the details of a highly complex system for which there is no widely agreed rationale grounded in economic theory. What is the optimal capital structure? What market failure necessitates the imposition of capital-adequacy requirements? Why can't the market be left to determine the appropriate level of capital? We do not find convincing answers to these questions in the theoretical literature.

In the literature on capital adequacy, it is often argued that capital-adequacy requirements are necessary to control moral-hazard problems generated by deposit insurance. Deposit insurance was introduced in the 1930s to prevent bank runs or, more generally, financial instability. It is well known that

* I would like to thank the Bank of Canada for sponsoring the conference, the conference organizers for inviting me, and the conference participants for their helpful comments, especially my discussant, Vincenzo Quadrini, and the panellists Claudio Borio and Chuck Freedman. Onur Özgür provided valuable research assistance. The financial support of the C.V. Starr Center for Applied Economics at NYU and the National Science Foundation under Grant No. SES 0095109 is gratefully acknowledged.

deposit insurance encourages risk-shifting behaviour on the part of banks (see, e.g., Merton (1977)). Risk shifting or asset substitution can be controlled by requiring the shareholders to post a bond in the form of adequate levels of bank capital. Thus, capital-adequacy requirements are indirectly justified by the desire to prevent financial crises. A large literature investigates the effect of capital-adequacy requirements on risk taking. While the effect of capital-adequacy requirements is usually to decrease risk taking, the reverse is also possible (see, e.g., Kim and Santomero (1988), Furlong and Keeley (1989), Gennotte and Pyle (1991), Rochet (1992), and Besanko and Kanatas (1996)).

An exception in the literature on capital-adequacy requirements is the paper by Hellmann, Murdoch, and Stiglitz (1998). Rather than simply taking the existence of deposit insurance as given, the authors examine what happens in the absence of deposit insurance. Hellmann, Murdock, and Stiglitz (1998, 2000) develop a model that allows for the effect of both a higher charter value and capital-adequacy requirements on risk-taking incentives. Controls on deposit interest rates are necessary, in addition to capital-adequacy requirements, to achieve a Pareto-efficient allocation of resources. These interest rate controls increase charter value and provide an extra instrument for controlling risk taking. A Pareto improvement is possible even in the absence of deposit insurance.

It appears from this brief review of the literature that the justification for capital-adequacy requirements is often found in the existence of deposit insurance, but this begs an important question. One bad policy (deposit insurance) does not justify another (capital-adequacy requirements). In the absence of deposit insurance, one must find another reason why banks cannot be left to choose their own capital levels.

Bank capital has two functions. One is to provide a buffer or cushion against unexpected shocks. This is the **risk-sharing function** of bank capital. The other is to provide incentives for management to avoid taking excessive risks. This is the **incentive function** of bank capital. These functions provide a rationale for bank capital, but they do not necessarily provide a rationale for regulation. As Gale (2003) points out, if banks can internalize the costs and benefits of bank capital, the privately optimal capital structure will coincide with the socially optimal level of capital. In fact, Gale provides a model of the risk-sharing function of capital and shows that a laissez-faire equilibrium, in which banks are left to choose their capital structure as they see fit, is Pareto-efficient. To provide a welfare-improving role for regulation of capital structure, we have to show that a bank's choice of capital structure creates a **pecuniary externality** that imposes welfare costs on other banks.

In policy discussions, it is often assumed that financial fragility, the possibility of one distressed bank infecting the others, provides the relevant externality; but the story is more complicated than that. The failure of a bank creates a pecuniary externality directly through its effect on creditors and indirectly through its effect on asset prices. However, as is well known from the fundamental theorems of welfare economics, pecuniary externalities do not necessarily imply inefficiency. In a related series of papers, Allen and Gale (1998, 2000b, 2004) have argued that, under certain conditions, including complete markets for sharing risk, the incidence of financial crises is socially optimal in a laissez-faire system. The conditions required are not innocuous, but they at least provide a counter-example to the presumption that regulation of capital adequacy is required. Perhaps more importantly, they focus attention on the necessary conditions for welfare-improving regulation: if markets for aggregate risks are incomplete, there is scope for regulation to improve the allocation of risk bearing.

In a series of related papers, Allen and Gale (1998, 2000a–d, 2004) describe a model that integrates intermediation and capital markets in a way that proves useful for the analysis of asset-price volatility, liquidity provision, financial crises, and related issues. The model can be briefly described as follows. There are two types of assets in the economy: short-term assets that yield an immediate but low return, and long-term assets that yield a higher but delayed return. Risk-averse individuals want to invest to provide for future consumption. However, they are uncertain about their preferences regarding the timing of consumption. If they invest in the long-term asset, they earn a high return, but it may not be available when they want to consume it. If they invest in the short-term asset, they have the certainty that it will be available when they want it, but they have to forego the higher return of the long-term asset. In short, there is a trade-off between liquidity and rate of return.

Banks are modelled as institutions that provide an optimal combination of liquidity and return. In this respect, we are simply following Diamond and Dybvig (1983) and a host of other writers; see, e.g., Chari and Jagannathan (1988), Jacklin and Bhattacharya (1988), Postlewaite and Vives (1987), and Wallace (1988, 1990). Banks take deposits from consumers and invest them in a portfolio of long- and short-term assets. In exchange, the bank gives the individual a deposit contract, that is, an option to withdraw from the bank. The amount withdrawn depends on the date at which the option is exercised, but for a given date, liquidity is guaranteed. By pooling independent risks, the bank is able to provide a better combination of liquidity and return than an individual could achieve alone. The aggregate demand for liquidity is less volatile than individual risks, so the bank can guarantee the same degree of liquidity while investing a smaller fraction of the portfolio in short-term assets, thus giving the depositor the benefit of the higher returns from the long-term assets.

Bank behaviour can be represented as the solution of an optimal contracting problem. Banks compete for customers by offering combinations of a portfolio and a deposit contract. Free entry into the banking sector guarantees that banks will earn zero profit in equilibrium and will offer the combination of portfolio and contract that maximizes the depositors' expected utility. Otherwise, another bank could enter, offer a more attractive contract, and take away the first bank's customers.

Risk can take the form of shocks to asset returns or the demand for liquidity. In this paper, we focus on liquidity shocks. These shocks provide a role for financial markets. Specifically, we introduce markets for securities that allow banks to insure against aggregate shocks. We also introduce markets on which banks can buy and sell the long-term assets in order to obtain or provide liquidity.

The introduction of these two types of markets has important implications for the welfare properties of the model. First, the existence of markets on which assets can be liquidated ensures that bankruptcy involves no inefficiency ex post. Fire-sale prices transfer value to the buyer but do not constitute a deadweight loss. Second, ex ante risk sharing is optimal if there is a complete set of Arrow securities for insuring against aggregate shocks.

For a long time, policy-makers have considered it axiomatic that crises are best avoided. By contrast, in the present framework, with complete markets, a laissez-faire financial system achieves the constrained-efficient allocation of risk and resources. When banks are restricted to using non-contingent deposit contracts, default introduces a degree of contingency that may be desirable from the point of view of optimal risk sharing. Far from being best avoided, financial crises can actually be *necessary* to achieve constrained efficiency. By contrast, avoiding default is costly. It requires either holding a very safe and liquid portfolio and earning lower returns, or reducing the liquidity promised to the depositors. In any case, the bank optimally weighs the costs and benefits and chooses the efficient level of default in equilibrium.

The important point is that avoidance of crises should not be taken as axiomatic. If regulation is required to minimize or obviate the costs of financial crises, it needs to be justified by a microeconomic welfare analysis based on standard assumptions. Furthermore, the form of the intervention should be derived from microeconomic principles. After all, financial institutions and financial markets exist to facilitate the efficient allocation of risks and resources. A policy that aims to prevent financial crises has an impact on the normal functioning of the financial system. Any government intervention may impose deadweight costs by distorting the normal functioning of the system. One of the advantages of a microeconomic analysis of financial crises is that it clarifies the costs associated with these distortions.

The model described so far has no role for capital. Banks are like mutual companies, operated for the benefit of their depositors, with no investment provided and no return received by the entrepreneurs who set them up. We can add capital to the model by assuming the existence of a class of risk-neutral investors who are willing to invest in the bank in return for an equity share. These investors are assumed to have a fixed opportunity cost of capital, determined by the best investment returns available to them outside the banking sector. We assume this return is at least as great as the return on the long-term asset. These investors can also speculate on the short- and long-term assets, for example, holding the short-term asset in order to buy up the long-term asset at a fire-sale price in the event of a default. This kind of speculation provides liquidity. It is superfluous in the case of complete Arrow securities, but plays an essential role in equilibrium with incomplete markets.

In the sections that follow, we illustrate the general approach of Allen and Gale and apply it to the question of optimal capital structure. We begin by setting out a benchmark model in which markets for aggregate risk are assumed to be complete. Under standard assumptions, analogous to the conditions of the Fundamental Theorems of Welfare Economics, we can show that the laissez-faire equilibrium of this economy is Pareto-efficient.

To provide a rationale for intermediation, we asssume incomplete participation in markets for contingent commodities: more precisely, consumers are excluded from participating directly in the markets for contingent commodities and are forced to use intermediaries to access those markets. Intermediaries are modelled as risk-sharing co-operatives that offer incentive-compatible risk-sharing contracts to consumers and use markets to lay off some of the risk. We show that, just as in the previous model, the laissez-faire allocation is Pareto-efficient. We also show that the optimal capital structure of the intermediaries is indeterminate: complete markets provide a perfect substitute for capital. This is just a version of the Modigliani-Miller theorem: capital structure is irrelevant when markets are complete.

Turning to the case of incomplete markets, we show that capital structure is determinate but may still be optimal if the missing markets can be replaced by capital. In other words, incompleteness does not necessarily give rise to a welfare-relevant pecuniary externality. However, if a welfare-relevant pecuniary externality does exist, the privately optimal equilibrium capital structure is determinate and differs from the socially optimal capital structure. In this case, there is a clear role for regulation, subject to the usual caveat that the policy-maker may need a lot of information to implement the optimal policy. However, even if the capital structure chosen in equilibrium is inefficient, it does not follow that minimum capital requirements will improve matters: there may be too much or too little capital in equilibrium.

The rest of the paper is organized as follows. Section 1 presents the basic model. Section 2 defines an Arrow-Debreu economy with complete markets and the corresponding equilibrium. The efficiency of the Arrow-Debreu equilibrium is discussed in section 3. Section 4 introduces intermediaries and discusses the irrelevance of capital structure in the Arrow-Debreu economy. In section 5, we abandon the Arrow-Debreu assumption of complete markets and show that capital structure is now determinate. The privately optimal capital structure is still socially optimal, however, so again there is no rationale for policy interventions such as capital regulation. To provide a rationale for policy intervention, we introduce heterogeneity among financial institutions in section 6. Efficiency requires cross-sectional risk sharing, in which institutions insure each other against liquidity shocks; but this cannot be achieved without complete markets. The laissez-faire equilibrium is now inefficient, so, in principle, some policy intervention may be justified. However, the optimal policy need not take the form of imposing minimum capital requirements.

1 A Model of Risk

As a vehicle for our analysis, we use a variant of the model found in Gale (2003). Time is divided into three periods or dates denoted by $t = 0, 1, 2$. At each date, there is a single, all-purpose good that can be used for consumption or for investment. There are two assets, a short-term asset that matures after one period and a long-term asset that matures after two periods.

- The **short-term asset** is represented by a storage technology: one unit of the good invested at date 0 yields one unit at date 1.
- The **long-term asset** is represented by a constant-returns-to-scale investment technology: one unit invested at date 0 yields $R > 1$ units at date 2 (and nothing at date 1).

In choosing the optimal combination of the two assets, there is a trade-off between liquidity and the rate of return. The short-term asset provides greater liquidity (immediate access to returns), but the long-term asset provides a higher return per unit invested. These properties play a critical role when markets are incomplete.

The economic agents in this economy are divided into two groups: risk-averse consumers who provide a demand for liquidity and risk-neutral investors who supply the capital. There is a continuum of identical, risk-neutral investors with unit mass. Although investors are risk-neutral, we assume that their consumption must be non-negative at each date. Otherwise, it is impossible to make sense of limited liquidity. The investor's utility function is defined by

$$u(c_0, c_1, c_2) = \rho c_0 + c_1 + c_2,$$

where $c_t \geq 0$ denotes the investor's consumption at date $t = 0, 1, 2$. The constant $\rho > R$ represents the investor's opportunity cost of funds. An investor's endowment consists of a large (unbounded) amount of the good at date 0 and nothing at dates 1 and 2.

There is a continuum of identical, risk-averse consumers with unit mass. Each consumer has an endowment of one unit of the good at date 0 and nothing at dates 1 and 2. At date 1, each consumer receives a preference shock: with probability λ, the consumer becomes an **early consumer**, who only values consumption at date 1, and with probability $1 - \lambda$, the consumer becomes a **late consumer**, who only values consumption at date 2. A consumer's period utility function $U : \mathbf{R}_+ \to \mathbf{R}$ is twice continuously differentiable and satisfies the usual neo-classical properties,

$$U'(c) > 0, U''(c) < 0, \lim_{c \searrow 0} U'(c) = \infty .$$

The consumer's risk aversion, together with uncertainty about the preference shock (early or late), creates a demand for insurance. By pooling these risks, it is possible to provide liquidity to consumers at date 1 while holding a smaller amount of the short asset.

Since investors are risk-neutral, there is an opportunity to share risk with the consumers. However, since investors have no endowments at date $t = 1, 2$, and since consumption must be non-negative, they can only share risk by turning their endowment at date 0 into assets that yield returns at date $t = 1, 2$. Investor preference for immediate consumption means that risk sharing is costly: a consumer will have to pay a premium that covers the investor's opportunity cost ρ. The form that this risk sharing takes will depend critically on the markets available and the presence or absence of intermediaries.

We assume that the fraction of early consumers is equal to the probability λ that an individual agent turns out to be an early consumer. The probability λ is a random variable with a continuous density function $f(\lambda)$ with support $[\lambda_0, \lambda_1]$. All uncertainty is resolved at the beginning of date 1, when the state of nature λ is observed and each agent discovers whether he or she is an early or late consumer.

There are two types of risk sharing in this economy. The consumers are ex ante identical and receive preference shocks that are independent conditional on the state, so there is scope for **intertemporal smoothing** of consumption: consumers decide ex ante how the total consumption available will be divided between early and late consumers in order to maximize their ex ante expected utility. Similarly, consumers and investors can share risk **across states**: investors agree to decrease their consumption when the consumers' marginal utility of consumption is high and increase it when the consumers' marginal utility of consumption is low.

Since neither consumers nor investors have an initial endowment of goods at dates 1 and 2, future consumption is provided by holding quantities of the short- and long-term assets. The equilibrium allocation of risk- and asset-holding depends on the available markets. In the following sections, we introduce a sequence of different market structures and characterize the corresponding allocation of risk and asset holding.

2 Equilibrium with Complete Markets

We begin by describing an **Arrow-Debreu economy**, in which the optimal allocation of risk bearing can be achieved by trading a complete set of contingent commodities. Commodities are distinguished by their delivery date and the state of nature on which delivery is contingent. The state of nature is identified with the fraction λ of early consumers. Commodities are distinguished by the date and state in which they are delivered, so there is a single commodity, the good for immediate delivery, at the first date, and there is a contingent commodity corresponding to each state λ and each date $t = 1, 2$. Taking the good at date 0 as the numeraire, we denote by $q_t(\lambda)$ the price of the contingent commodity in state λ at date t, for $t = 1, 2$. A **price system** is denoted by the ordered pair $q = (q_1, q_2)$, where $q_t : [\lambda_0, \lambda_1] \to \mathbf{R}_+$ for $t = 1, 2$.

In equilibrium, the profits from holding assets will be zero, so it is immaterial who actually holds the assets. Without loss of generality, we can assume that assets are held by a notional profit-maximizing producer. The actual quantities of the two assets held in equilibrium will be determined by the market-clearing conditions.

The ability to invest in the short- and the long-term assets gives rise to a production technology that can be represented by a production set Y. A **production plan** is an ordered triple $y = (y_0, y_1, y_2)$, such that $y_0 \geq 0$ and $y_t : [\lambda_0, \lambda_1] \to \mathbf{R}_+$ for $t = 1, 2$. A production plan $y = (y_0, y_1, y_2)$ belongs to Y if it satisfies the following conditions, for some $0 \leq \theta \leq y_0$ and any λ:

$$y_1(\lambda) \leq \theta$$

and

$$y_1(\lambda) + y_2(\lambda) \leq \theta + (y_0 - \theta)R.$$

Here, y_0 is the input into the production process at date 0. The producer divides the input y_0 at date 0 into an investment of θ units in the short-term asset and an investment of $y_0 - \theta$ units in the long-term asset. This portfolio yields θ units of the good at date 1 (this is the return on the short-term asset) and $(y_0 - \theta)R$ units of the good at date 2 (this is the return on the long-term asset). The first inequality above says that the output at date 1 must be less than or equal to θ. The excess $\theta - y_1(\lambda)$ is stored until date 2, when the output must be less than or equal to $(y_0 - \theta)R + \theta - y_1(\lambda)$. This is the meaning of the second inequality.

The value of a production plan y is given by

$$E[-y_0 + q_1(\lambda)y_1(\lambda) + q_2(\lambda)y_2(\lambda)].$$

Because the investment technology exhibits constant returns to scale, the value of any production plan must be non-positive in equilibrium. For q to be an equilibrium price system, it must satisfy the following no-arbitrage conditions. First, there is no profit from investing one unit in the short-term asset at date 0 and holding it until date 1. This implies that

$$E[q_1(\lambda)] \leq 1, \tag{1}$$

with equality if there is positive investment in the short-term asset. Second, there is no profit from investing one unit in the long-term asset at date 0 and holding it until date 2. This implies that

$$E[q_2(\lambda)R] \leq 1, \tag{2}$$

with equality if there is positive investment in the long-term asset. Third, there is no profit from purchasing one unit of the good at date 1 in state λ and storing it until date 2. This implies that we must have

$$q_1(\lambda) \geq q_2(\lambda) \tag{3}$$

for every λ, with equality if there is storage between date 1 and date 2 in state λ. A price system is called **admissible** if it satisfies the conditions (1) – (3).

Let e_0 denote the investor's supply of capital at date 0 and let $e_t(\lambda)$ denote the investor's consumption at date $t = 1, 2$ in state λ. Then, a **consumption plan** for the representative investor is an ordered triple $e = (e_0, e_1, e_2)$, such that $e_0 \geq 0$ and $e_t : [\lambda_0, \lambda_1] \to \mathbf{R}_+$. Similarly, we let $c = (c_1, c_2)$ denote the representative consumer's **consumption plan**, where $c_t : [\lambda_0, \lambda_1] \to \mathbf{R}_+$ for $t = 1, 2$. The sets of consumption plans for investors and consumers are denoted by *E* and *C*, respectively.

An **allocation** consists of a production plan $y \in Y$, a consumption plan *e* for the investor, and a consumption plan *c* for the consumer. An allocation (c, e, y) is **attainable** if

$$1 + e_0 = y_0, \tag{4}$$

$$\lambda c_1(\lambda) + e_1(\lambda) = y_1(\lambda), \forall\lambda, \tag{5}$$

$$(1 - \lambda)c_2(\lambda) + e_2(\lambda) = y_2(\lambda), \forall\lambda. \tag{6}$$

The first market-clearing condition (4) says that, at date 0, the consumers' endowment plus the amount supplied by investors e_0 is equal to the amount invested in the two assets. The second market-clearing condition (5) says that, at date 1 in state λ, the consumption of the early consumers $\lambda c_1(\lambda)$ plus the investors' consumption $e_1(\lambda)$ is equal to the amount $y_1(\lambda)$ supplied by the producer. The third market-clearing condition (6) says that, at date 2 in state λ, the consumption of the late consumers $(1 - \lambda)c_2(\lambda)$ plus the investors' consumption $e_2(\lambda)$ is equal to the amount $y_2(\lambda)$ supplied by the producer.

Given an admissible price system, the decision problem of consumers is to choose $(c_1, c_2) \in C$ to

$$\begin{aligned} &\max \ E[\lambda U(c_1(\lambda)) + (1 - \lambda)U(c_2(\lambda))] \\ &\text{s.t.} \quad E[\lambda q_1(\lambda)c_1(\lambda) + (1 - \lambda)q_2(\lambda)c_2(\lambda)] = 1. \end{aligned} \tag{7}$$

Note that the budget constraint is written as an expected value. Multiplying the values in a particular state by the probability $f(\lambda)$ is just a (non-essential) normalization. The consumer only pays for the expected value of the goods consumed at each date: in the aggregate, there is no uncertainty about the number of early consumers in state λ, so the pricing is risk-neutral. Note that the choices of consumers do not need to satisfy an incentive constraint.

Consumer demand is automatically incentive-compatible in equilibrium, because the restriction $q_1(\lambda) \geq q_2(\lambda)$ implies that $c_1(\lambda) \leq c_2(\lambda)$, so a late consumer has no incentive to pretend to be an early consumer.

Similarly, the decision problem of investors is to choose a consumption plan $(e_0, e_1, e_2) \in E$ to

$$\begin{aligned} &\max \; E[-\rho e_0 + e_1(\lambda) + e_2(\lambda)] \\ &\text{s.t.} \quad E[-e_0 + q_1(\lambda)e_1(\lambda) + q_2(\lambda)e_2(\lambda)] = 0. \end{aligned} \tag{8}$$

Again, without loss of generality, the budget constraint can be written in terms of expected values.

A **competitive equilibrium** for this Arrow-Debreu economy consists of an admissible price system q and an attainable allocation (c, e, y), such that c and e solve the decision problems (7) and (8) for the consumer and investor, respectively, at the prevailing price system q.

3 Optimal Risk Sharing

The First Fundamental Theorem of Welfare Economics asserts that every competitive equilibrium is Pareto-efficient.

Theorem 1. *If (c, e, y, p) is a Walrasian equilibrium of the Arrow-Debreu economy, the allocation (c, e, y) is Pareto-efficient.*

From the definition of Pareto efficiency, it is clear that the allocation (c, e, y) is Pareto-efficient only if it maximizes the expected utility of the consumers, subject to a constraint on the investors' expected utility and the attainability conditions. In other words, we can characterize the optimal risk-sharing scheme as the solution to a planning problem. In equilibrium, investors earn zero profit, so their participation constraint is

$$E[-\rho e_0 + e_1(\lambda) + e_2(\lambda)] \geq 0.$$

Using the attainability conditions, we can rewrite the participation constraint equivalently as

$$\theta + (e_0 + 1 - \theta)R - E[\lambda c_1(\lambda) + (1-\lambda)c_2(\lambda)] \geq \rho e_0,$$

and the attainability conditions are given by equations (4) to (6). So, the planner's problem can be written as follows:

$$\begin{aligned} \max\ & E[\lambda U(c_1(\lambda)) + (1-\lambda)U(c_2(\lambda))] \\ \text{s.t.}\ & 0 \le \theta \le 1 + e_0 \\ & \lambda c_1(\lambda) + e_1(\lambda) \le \theta, \\ & \lambda c_1(\lambda) + (1-\lambda)c_2(\lambda) + e_1(\lambda) + e_2(\lambda) \le \theta + (e_0 + 1 - \theta)R, \\ & \theta + (e_0 + 1 - \theta)R - E[\lambda c_1(\lambda) + (1-\lambda)c_2(\lambda)] \ge \rho e_0. \end{aligned} \tag{9}$$

Theorem 2. *If* (c, e, y, q) *is an equilibrium of the Arrow-Debreu economy, then* (c, e, θ) *is a solution to the decision problem (9).*

The form of the optimal risk-sharing arrangement between consumers and investors is given by the following result.

Theorem 3. *If* (c, e, θ) *is a solution to the decision problem (9), the optimal consumption allocation is characterized by the parameters* (θ, e_0, d)*:*

$$c_2(\lambda) = \min\left\{d, \max\left\{\theta + (1 + e_0 - \theta)R, \frac{1 + e_0 - \theta}{1 - \lambda}\right\}\right\},$$

$$c_1(\lambda) = \min\left\{\frac{\theta}{\lambda}, \theta + (1 + e_0 - \theta)R\right\}.$$

Proof. The first-order conditions for the consumer's problem are

$$U'(c_1(\lambda)) = \mu q_1(\lambda) \tag{10}$$

and

$$U'(c_2(\lambda)) = \mu q_2(\lambda) \tag{11}$$

for all values of λ. For the investor's problem, the first-order conditions are

$$1 \le \varphi q_1(\lambda),$$

$$1 \le \varphi q_2(\lambda),$$

and

$$\rho \le \varphi,$$

with equality holding if the investor's consumption $e_t(\lambda)$ or supply of capital e_0 is positive. In what follows, we assume $e_0 > 0$. Otherwise, the theorem is trivial.

To characterize the efficient risk-sharing scheme, we consider two cases. In the first case, $q_1(\lambda) = q_2(\lambda)$ and, in the second, $q_1(\lambda) > q_2(\lambda)$.

Case 1: $q_1(\lambda) = q_2(\lambda)$. Then $U'(c_1(\lambda)) = U'(c_2(\lambda))$ implies that

$$c_1(\lambda) = c_2(\lambda) \le \theta + (1 + e_0 - \theta)R.$$

The inequality follows from the attainability condition. The first-order conditions imply that $c_1(\lambda)$ and $c_2(\lambda)$ are independent of λ as long as $q_1(\lambda) = q_2(\lambda)$.

Case 2: $q_1(\lambda) > q_2(\lambda)$. Then $1 \le \rho q_2(\lambda) < \rho q_1(\lambda)$ implies that $e_1(\lambda) = 0$, so

$$c_1(\lambda) = \frac{\theta}{\lambda}$$

and

$$(1 - \lambda)c_2(\lambda) \le (1 + e_0 - \theta)R.$$

There are two subcases to be considered. If $\rho q_2(\lambda) > 1$, then the inequality is strict, and if $\rho q_2(\lambda) = 1$, then $c_2(\lambda) = \phi(\mu/\rho)$, where $\phi(\cdot)$ is the inverse of $U'(\cdot)$. Call this *d*. Then $q_1(\lambda) > q_2(\lambda)$ implies that

$$c_2(\lambda) = \min\left\{d, \frac{(1 + e_0 - \theta)R}{1 - \lambda}\right\}.$$

To complete the proof, we have to show that $d \ge \theta + (1 + e_0 - \theta)R$. If this were not so, a reduction in e_0 and a corresponding reduction in θ would increase consumption at every point. This is because every unit of capital costs ρ in expected returns and can yield at most *R* when invested in assets. The difference comes out of consumption. The fact that $d \ge \theta + (1 + e_0 - \theta)R$ implies that $c_1(\lambda) = c_2(\lambda) = \theta + (1 + e_0 - \theta)R$ when $q_1(\lambda) = q_2(\lambda)$, i.e., $e_1(\lambda) = e_2(\lambda) = 0$. This completes the proof.

4 Intermediation and Capital Structure

In an economy with complete markets, individuals can achieve efficient risk sharing without the intervention of intermediaries. To provide a role for intermediaries, we assume **incomplete participation**; that is, individual consumers cannot participate in markets for contingent commodities, but intermediaries and investors can. To provide for future consumption, a consumer deposits his endowment of one unit of the good at date 0 with an intermediary in exchange for a promise to provide an early consumer with $c_1(\lambda)$ units of consumption at date 1 in state λ and a late consumer with $c_2(\lambda)$ units of consumption at date 2 in state λ. The intermediary can also raise e_0 units of capital at date 0 in return for the promise to pay dividends equal to $e_1(\lambda)$ units of consumption at date 1 in state λ and $e_2(\lambda)$ units of consumption at date 2 in state λ.

An **allocation** for the intermediated economy consists of a risk-sharing contract (c, e) proposed by the representative intermediary, a consumption plan f for the representative investor, and a production plan $y \in Y$. An allocation (c, e, f, y) is **attainable** if

$$1 + e_0 + f_0 = y_0,$$

$$\lambda c_1(\lambda) + e_1(\lambda) + f_1(\lambda) = y_1(\lambda), \forall\lambda,$$

$$(1 - \lambda)c_2(\lambda) + e_2(\lambda) + f_2(\lambda) = y_2(\lambda), \forall\lambda.$$

Free entry will ensure zero profits in intermediation, and competition among intermediaries forces them to maximize the expected utility of the depositor, subject to the participation constraint of investors. Given the equilibrium prices q, the intermediary will choose c and e to solve the following decision problem:

$$\begin{aligned} \max\ & E[\lambda U(c_1(\lambda)) + (1 - \lambda)U(c_2(\lambda))] \\ \text{s.t.}\ & E[q_1(\lambda)\{\lambda c_1(\lambda) + e_1(\lambda)\} + q_2(\lambda) \\ & \{(1 - \lambda)c_2(\lambda) + e_2(\lambda)\}] \le 1 + e_0 \\ & E[-\rho e_0 + e_1(\lambda) + e_2(\lambda)] \ge 0. \end{aligned} \tag{12}$$

If q is an equilibrium price system, then

$$E[-\rho e_0 + e_1(\lambda) + e_2(\lambda)] \ge 0$$

implies that

$$E[q_1(\lambda)e_1(\lambda) + q_2(\lambda)e_2(\lambda)] \ge e_0,$$

so a solution to the intermediary's decision problem must satisfy

$$E[-\rho e_0 + e_1(\lambda) + e_2(\lambda)] = 0$$

and

$$E[q_1(\lambda)e_1(\lambda) + q_2(\lambda)e_2(\lambda)] = e_0.$$

Thus, the intermediary's problem is equivalent to

$$\begin{aligned} &\max\ E[\lambda U(c_1(\lambda)) + (1-\lambda)U(c_2(\lambda))] \\ &\text{s.t.}\quad E[q_1(\lambda)\lambda c_1(\lambda) + q_2(\lambda)(1-\lambda)c_2(\lambda)] = 1, \end{aligned}$$

which is the individual's problem in the complete-markets equilibrium without intermediation.

The representative investor takes e as given and chooses f to solve

$$\begin{aligned} &\max\ E[-\rho f_0 + f_1(\lambda) + f_2(\lambda)] \\ &\text{s.t.}\quad E[-f_0 + q_1(\lambda)f_1(\lambda) + q_2(\lambda)f_2(\lambda)] \le 0. \end{aligned} \tag{13}$$

An **equilibrium** of the intermediated economy consists of an attainable allocation (c, e, f, y) and an admissible price system q, such that (c, e) solves problem (12) and f solves problem (13).

The following Modigliani-Miller-style result tells us that capital structure is irrelevant in an equilibrium with complete markets.

Theorem 4. *Let (c, e, f, y, q) be an equilibrium of the intermediated economy. For any $\hat{e}$ and $\hat{f}$ such that $\hat{e} + \hat{f} = e + f$ and $\hat{e}$ and $\hat{f}$ solve the problem (13), $(c, \hat{e}, \hat{f}, y, q)$ is an equilibrium, and the expected utilities of investors and consumers, respectively, are identical in the two equilibria.*

Proof. Clearly, $(c, \hat{e}, \hat{f}, y)$ is attainable, since the attainability conditions depend only on the sum $\hat{e} + \hat{f}$. By assumption, $\hat{f}$ satisfies the investor's decision problem (13), and $(c, \hat{e})$ solves the intermediary's decision problem (12), because $\hat{e} = e + (f - f)$ minimizes the cost of satisfying the investor's participation constraint.

Note that an equilibrium (c, e, y, q) of the Arrow-Debreu economy is isomorphic to an equilibrium of the intermediated economy of the form $(c, e, 0, y, q)$. This implies that any equilibrium is Pareto-efficient.

5 Equilibrium with Incomplete Markets

We have shown that bank capital is redundant when markets are complete: market trades are a perfect substitute for optimal capital structure. The optimal capital structure is indeterminate and includes the case where there is no external capital at all. In particular, since the equilibrium is Pareto-efficient, there is no rationale for capital regulation.

The assumption of complete markets is unrealistic, of course. It is the absence of complete markets that provides an essential role for bank capital. So, to understand the importance of capital, we need to study an economy with incomplete markets. For concreteness, suppose there are no markets for contingent commodities at date 0. The only markets that exist are spot markets for goods and assets. Specifically, at date 1, it is possible to sell the long-term asset. Formally, we model this by assuming there is a market at date 1 for goods delivered at date 2, but this is equivalent to a spot market for the long-term asset. This market structure is quite special, but it simplifies the analysis and allows us to make the essential points about the factors that determine the optimal capital structure and the efficiency of risk sharing in equilibrium.

We continue to assume that consumers have no access to capital markets and must use intermediaries to provide for future consumption. As before, a consumer deposits his or her endowment of one unit of the good at date 0 with an intermediary in exchange for a promise to provide an early consumer with $c_1(\lambda)$ units of consumption at date 1 in state λ and a late consumer with $c_2(\lambda)$ units of consumption at date 2 in state λ. The intermediary can also raise e_0 units of capital at date 0 in return for the promise to pay dividends equal to $e_1(\lambda)$ units of consumption at date 1 in state λ and $e_2(\lambda)$ units of consumption at date 2 in state λ.

When markets are complete, it does not matter who holds the assets, and we can assume that all assets are held by an anonymous firm and that investors, intermediaries, and consumers use complete markets to allocate consumption across states and dates. This is a corollary of the Modigliani-Miller theorem. By contrast, when markets are incomplete, holding assets is an essential way of hedging risk and redistributing consumption across states.

At the first date, the representative intermediary receives one unit of the good as a deposit from each consumer and e_0 units of capital from investors. In exchange, it promises consumption plans (c_1, c_2) and (e_1, e_2) to the consumers and investors, respectively. Because there are no markets for future contingent commodities, the only way for the intermediary to provide consumption in the future is to invest the funds it has received in a portfolio of the short- and long-term assets. The intermediary's investment strategy can be represented by a production plan $y \in Y$. Corresponding to any production plan y is a portfolio consisting of θ units of the long-term asset and $1 + e_0 - \theta$ units of the long-term asset at date 0 and a decision to store goods at date 1. There is no loss of generality in assuming that individual consumers hold no assets. Whatever the individual can do can be replicated by the intermediary.

We can assume without loss of generality that the role of investors is simply to provide capital to the intermediary through the contract $e = (e_0, e_1, e_2)$. While it is feasible for the investors to invest in assets at date 0 and trade them at date 1, it can never be profitable for them to do so in equilibrium. More precisely, the no-arbitrage conditions ensure that profits from trading assets are zero or negative at any admissible prices and the investors' preferences for consumption at date 0 imply that the investors will never want to invest in assets at date 0 and consume the returns at dates 1 and 2.

An **allocation** is an array (c, e, y), where c is the consumption plan for consumers, e is the capital structure (i.e., the investment and consumption plan for investors), and y is the intermediary's production plan. The allocation (c, e, y) is **attainable** if

$$1 + e_0 = y_0,$$

$$\lambda c_1(\lambda) + e_1(\lambda) = y_1(\lambda), \forall\lambda,$$

$$(1-\lambda)c_2(\lambda) + e_2(\lambda) = y_2(\lambda), \forall\lambda.$$

Let $p(\lambda)$ denote the price of the good at date 2 in terms of the good at date 1. The long-term asset will be held between date 1 and date 2 only if $p(\lambda) \le 1$ and the short-term asset will be held between date 1 and date 2 only if $p(\lambda) \ge 1$. A price system p is **admissible** for an allocation in which $0 < \theta < 1 + e_0$ if

$$p(\lambda) \le 1$$

for all λ and $p(\lambda) = 1$ if $\lambda c_1(\lambda) + e_1(\lambda) < \theta$. This condition is the analogue of the condition that $q_1(\lambda) \ge q_2(\lambda)$ for all λ in the Arrow-Debreu equilibrium.

Since the intermediary cannot hedge against preference shocks by trading contingent commodities, it must satisfy a budget constraint in each state λ at date 1:

$$(\lambda c_1(\lambda) + e_1(\lambda)) + p(\lambda)((1-\lambda)c_2(\lambda) + e_2(\lambda)) \le y_1(\lambda) + p(\lambda)y_2(\lambda).$$

The left-hand side is the present value of consumption promised to consumers and investors, and the right-hand side is the present value of the intermediary's portfolio.

An **intermediated equilibrium** for the incomplete markets economy consists of an attainable allocation (c, e, y) and an admissible price system p, such that (c, e, y) solves the intermediary's decision problem:

$$\begin{aligned} \max\ & E[\lambda U(c_1(\lambda)) + (1-\lambda)U(c_2(\lambda))] \\ \text{s.t.}\ & \lambda c_1(\lambda) + e_1(\lambda) + p(\lambda)\{(1-\lambda)c_2(\lambda) \\ & + e_2(\lambda)\} \le y_1(\lambda) + p(\lambda)y_2(\lambda),\ \forall\lambda, \\ & E[-\rho e_0 + e_1(\lambda) + e_2(\lambda)] \ge 0. \end{aligned}$$

Theorem 5. *Suppose that* (c^*, e^*, y^*, q^*) *is an equilibrium of the Arrow-Debreu economy and define* p^* *by putting*

$$p^*(\lambda) = \frac{q_2^*(\lambda)}{q_1^*(\lambda)}$$

for any λ. Then (c^*, e^*, y^*, q^*) is an intermediated equilibrium of the economy with incomplete markets.

Proof. It is sufficient to show that (c^*, e^*, y^*) solves the intermediary's problem at the defined prices p^*. It is easy to check that, if an allocation (c, e, y) satisfies the constraints of the intermediary's problem, then the consumption plan c satisfies the constraints for the consumer's problem in the Arrow-Debreu equilibrium. Thus, the intermediary cannot do better than choose (c^*, e^*, y^*).

In effect, the intermediary is replacing the missing markets. There is no need for trade among intermediaries, because they are assumed to be identical: each intermediary is a microcosm of the whole economy. The representative intermediary internalizes all the necessary trades between investors and consumers. It provides a risk-sharing contract for consumers, whereby early and late consumers optimally smooth consumption over time, and it provides a risk-sharing contract with investors, whereby the optimal capital structure (e_0, e_1, e_2) shares risk across states between investors and consumers. Implicit in the contract (e_0, e_1, e_2) is a premium paid to the investors for delaying their consumption. So, in spite of the missing markets, the laissez-faire equilibrium is still the first best.

Because the first best is achieved (i.e., all marginal rates of substitution are equalized), pecuniary externalities have no impact on efficiency. However, the absence of markets for sharing risk means that capital plays an essential role in achieving optimal risk sharing: *in this equilibrium, the capital structure is determinate, and the Modigliani-Miller theorem no longer* *holds*. The importance of incomplete markets is that they provide an essential role for bank capital in achieving optimal risk sharing.

6 Incomplete Markets and Heterogeneous Intermediaries

The equilibrium with incomplete markets (examined in the previous section) is efficient, because each intermediary is assumed to have a representative sample of the consumers in the economy. Since the intermediaries are identical, there are no gains from trade among them. An intermediary combining capital and deposits in the optimal proportions can behave like a central planner. To provide an essential role for markets, there must be gains from trade among intermediaries, and this can only come from **heterogeneity** among intermediaries.

As a practical matter, intermediaries are quite likely to be heterogeneous, and markets will provide an important channel for them to share risk. For example, suppose that intermediaries draw their depositors from distinct locations and that the consumers' types are correlated at any location. Then the fraction of early consumers will vary from location to location in a given state. If the state is λ, the proportion of early consumers in location i is denoted by $\lambda + \varepsilon_i$, where ε_i is a random variable with mean zero. If ε_i is i.i.d. across locations and there are a large number of locations, the ex ante probability of being an early consumer in any location and the average proportion of early consumers in the entire economy are both equal to the expected value of $\lambda + \varepsilon_i$, which is λ. Since intermediaries are heterogeneous ex post, there are gains from trade at date 1: intermediaries whose depositors are mainly early consumers will want to sell the long-term asset to get liquidity, while intermediaries whose depositors are mainly late consumers will use their liquidity to purchase the long-term asset.

Here we want to illustrate the importance of missing markets in the simplest possible way, and we can do this by assuming that heterogeneity among intermediaries takes a very special form. Specifically, we assume that, in any state, an intermediary's customers are either all early consumers or all late consumers. If the state is λ, this means that a fraction λ of the intermediaries consists entirely of early consumers and a fraction $1 - \lambda$ consists entirely of late consumers. Since intermediaries are identical ex ante, the probability that an intermediary consists entirely of early consumers is equal to λ when the state is λ.

When intermediaries are heterogeneous, the definition of equilibrium with incomplete markets has to be revised in two respects. First, the payment to investors will, in principle, be contingent on the realization of the intermediary's type (i.e., whether the consumers are early or late). Without loss of generality, we can assume that the investors are paid only at date 2 (this follows the fact that $p(\lambda) \le 1$ in equilibrium, so consumption at date 2 is cheaper than consumption at date 1) and let $e_t(\lambda)$ denote the payment to the investors when the intermediary's type is $t = 1, 2$. Second, the intermediary's budget constraint will depend on the realization of its type. If the intermediary is of type *t*, it offers consumers $c_t(\lambda)$ units of consumption at date *t* and investors $e_t(\lambda)$ units of consumption at date 2. The present value of consumption is $c_1(\lambda) + p(\lambda)e_1(\lambda)$ if $t = 1$ and $p(\lambda)(c_2(\lambda) + e_2(\lambda))$ if $t = 2$. Substituting these expressions on the left-hand side yields the appropriate budget constraint in each state (λ).

An **equilibrium** for this economy consists of an attainable allocation (c, e, y) and an admissible price function *p*, such that (c, e, y) solves the decision problem

$$\begin{aligned} \max\ & E[\lambda U(c_1(\lambda)) + (1-\lambda)U(c_2(\lambda))] \\ \text{s.t.}\ & c_1(\lambda) + p(\lambda)e_1(\lambda) \le y_1(\lambda) + p(\lambda)y_2(\lambda),\ \forall\lambda, \\ & p(\lambda)(c_2(\lambda) + e_2(\lambda)) \le y_1(\lambda) + p(\lambda)y_2(\lambda),\ \forall\lambda, \\ & E[\lambda e_1(\lambda) + (1-\lambda)e_2(\lambda)] \ge \rho e_0. \end{aligned}$$

6.1 The inefficiency of risk sharing

Suppose that the intermediary chooses a portfolio $(\theta, 1 + e_0 - \theta)$. The value of the intermediary's portfolio is

$$w_1(\lambda) \equiv \theta + p(\lambda)(1 + e_0 - \theta)R$$

at date 1, and

$$w_2(\lambda) \equiv \theta / p(\lambda) + (1 + e_0 - \theta)R$$

at date 2.

Theorem 6. *If (c, e, y, p) is an intermediated equilibrium of the economy with incomplete markets and heterogeneous intermediaries, then for some constant d, the optimal consumption allocation is*

$$c_2(\lambda) = \min\{w_2(\lambda), d\},\ \forall\lambda,$$

and

$$c_1(\lambda) = \min\left\{w_1(\lambda), \phi\left(\frac{\mu}{p(\lambda)}\right)\right\}, \forall\lambda,$$

where $U'(\phi(\mu/p(\lambda))) \equiv \mu/p(\lambda)$ and $\mu > 0$.

To simplify the comparison of the incomplete-markets equilibrium with the first best, we assume that

$$w_1(\lambda) \le \phi\left(\frac{\mu}{p(\lambda)}\right), \forall\lambda.$$

Then the decision problem faced by the intermediary is to maximize the consumers' expected utility

$$E[\lambda U(w_1(\lambda)) + (1-\lambda)U(\min\{w_2(\lambda), d\})],$$

subject to the constraints

$$0 \le \theta \le 1 + e_0,$$

and

$$E[(1-\lambda)(\max\{w_2(\lambda) - d, 0\})] \ge \rho e_0.$$

The equilibrium price function is given by

$$p(\lambda) = \min\left\{1, \max\left\{\frac{(1-\lambda)\theta}{\lambda(1+e_0-\theta)R}\right\}\right\}.$$

The value of the intermediary's portfolio is

$$w_1(\lambda) = \theta + p(\lambda)(1+e_0-\theta)R = \min\left\{\theta + (1+e_0-\theta)R, \frac{\theta}{\lambda}\right\}$$

at date 1, and

$$w_2(\lambda) = \theta/p(\lambda) + (e_0 + 1 - \theta)R = \max\left\{\theta + (1+e_0-\theta)R, \frac{(1+e_0-\theta)R}{(1-\lambda)}\right\}$$

at date 2. Substituting these expressions into the objective function, we get

$$\lambda U\left(\min\left\{\theta + (1 + e_0 - \theta)R, \frac{\theta}{\lambda}\right\}\right) +$$

$$(1-\lambda)U\left(\min\left\{d, \max\left\{\theta + (1 + e_0 - \theta)R, \frac{(1 + e_0 - \theta)R}{(1-\lambda)}\right\}\right\}\right).$$

A central planner, subject to the same constraints (incomplete markets) as the intermediaries but able to control the aggregate level of capital and investment in liquid assets would maximize the expected value of this objective subject to the constraints above. The difference between these two maximization problems captures the pecuniary externality that causes the constrained inefficiency of the equilibrium with incomplete markets.

6.2 Some policy experiments

We know from Geanakoplos and Polemarchakis (1986) that models with incomplete markets are generically constrained-inefficient, but the characterization of an optimal policy is difficult. Knowing that there exists a welfare-improving intervention is not the same thing as knowing what it is. To gain insight into the complexities and nuances of policy intervention, we introduce a simple example that can be solved numerically and consider the effect of various policy experiments.

We assume that the period utility function is logarithmic:

$$U(c) = \log c,$$

and the probability of being an early consumer is uniformly distributed

$$\lambda \sim \text{unif } [0, 1].$$

As usual, the return on the short-term asset is normalized to 1 and the investor's opportunity cost of funds is fixed at

$$\rho = 2.$$

The return on the long-term asset, R, is allowed to assume a number of values between 1 and 2. The cost of liquidity is measured by $R - 1$, the difference between the returns on the long-term and short-term assets. The cost of capital is measured by $\rho - R$, the difference between the opportunity cost of funds and the return on the long-term asset. So, as R increases, the cost of liquidity increases and the cost of capital decreases.

Intermediaries are assumed to be heterogeneous in the sense that their depositors are either all early consumers or all late consumers.

6.3 Complete markets

We begin by determining the efficient allocation of investment and risk. This will serve as a benchmark for the analysis of the incomplete-markets model. Rather than assume complete markets and calculate the Arrow-Debreu equilibrium, we consider an artificial economy with homogeneous intermediaries and solve for the incomplete-markets equilibrium. As we saw in section 5, when intermediaries are homogeneous, each intermediary is a microcosm of the economy, and an optimal capital structure allows a representative intermediary to achieve the first-best allocation, even if there are no markets for contingent commodities. Solving for an equilibrium of this artificial economy will yield an allocation that is equivalent to the complete-markets model with heterogeneous intermediaries and has the additional advantage that the capital structure is determinate, so we can easily compare the capital structures of the two economies.[1]

Table 1 shows the equilibrium values of capital e_0, investment in the short asset θ, and expected utility EU, for various values of R.

Since the total invested in short- and long-term assets at date 0 is $1 + e_0$, the ratio of capital to assets is $e_0/(1 + e_0)$, which is approximately the same as e_0 for small values of e_0. As R varies from 1.2 to 1.8, the optimal capital ratio varies from less than 2 per cent to around 13/113 = 11 per cent. The cost of capital is lower, the higher the return on the long-term asset, and intermediaries choose to hold more capital when its cost falls.

The demand for liquidity, measured by θ, the investment in the short-term asset, falls as R increases. When $R = 1.2$, the share of the short-term asset in the portfolio is nearly 74 per cent; when $R = 1.8$, the short-term assets share falls to 56/113 = 50 per cent. Again, this makes sense. There are two ways of providing liquidity at date 1: by holding the short-term asset or by holding the long-term asset and selling it if necessary. An increase in the return on the long-term asset causes intermediaries to hold more of the long-term asset and less of the short-term asset.

Obviously, an increase in the return on the long-term asset, other things remaining equal, must increase the expected utility of the typical consumer, and this is reflected in the right-hand column of Table 1.

---

1. In the complete-markets economy, capital structure is irrelevant, because markets provide a perfect substitute for capital, but there is a counterpart to capital, namely, the total amount invested in assets at date 0 by the investors.

Table 1
Equilibrium for the Arrow-Debreu economy

ρ	R	e_0	θ	EU
2	1.2	0.0135	0.7400	0.0293
2	1.3	0.0245	0.6845	0.0539
2	1.4	0.0375	0.6465	0.0810
2	1.5	0.0530	0.6185	0.1095
2	1.6	0.0725	0.5965	0.1388
2	1.7	0.0970	0.5785	0.1690
2	1.8	0.1300	0.5635	0.2000

6.3.1 Incomplete markets

Now let us assume that intermediaries are heterogeneous and solve for the incomplete-markets equilibrium. Since intermediaries are heterogeneous, there is a demand for hedging between intermediaries, but since there are no markets for hedging risks at date 0, intermediaries can only obtain liquidity ex post by selling the long-term asset to intermediaries that have an excess supply of liquidity (because their depositors are late consumers). The equilibrium values are given in Table 2 for the same parameter values.

Once again, we see that, as R increases, the share of capital rises, the holding of the short-term asset falls with R, and the expected utility of the typical depositor increases.

Comparing Tables 1 and 2, the first thing we notice is that the level of capital is much lower when markets are incomplete. For each value of R, the amount of capital held is a little more than half the amount held when markets are complete. This is the optimal response for the intermediary, but the equilibrium is not efficient. The second thing to notice is that the amount of the short-term asset held is greater. The share of the short-term asset is nearly 79 per cent when $R = 1.2$ and nearly 64 per cent when $R = 1.8$. Thus, the proportional increase in the share of the short-term asset varies from 6.7 per cent to 28 per cent.

Comparing the expected utilities in Tables 1 and 2, we see that the expected utility of the typical depositor is lower when markets are incomplete: the risk-sharing possibilities are reduced when markets are incomplete.

6.4 Capital requirements

It is clear that the first-best allocation is not achieved when markets are incomplete—this is why the expected utility is lower. However, the relevant notion of efficiency is constrained efficiency, not Pareto efficiency. Given the

Table 2
Equilibrium with incomplete markets

ρ	R	e_0	θ	EU
2	1.2	0.0085	0.7900	0.0279
2	1.3	0.0153	0.7460	0.0514
2	1.4	0.0228	0.7180	0.0770
2	1.5	0.0310	0.6990	0.1034
2	1.6	0.0399	0.6866	0.1302
2	1.7	0.0430	0.6680	0.1574
2	1.8	0.0602	0.6749	0.1835

incompleteness of markets, it is obvious that welfare will be lower. The interesting question is whether, taking the incompleteness of markets as a constraint, there is a simple policy intervention that will make everyone better off. If there are missing markets, it is presumably because there is some cost or technological constraint to which the policy-maker is also subject. So we constrain the policy-maker to using spot markets to share risk and to smooth consumption intertemporally. Equilibrium is said to be **constrained-efficient** if it is impossible to make everyone better off by changing the allocation of goods at date 0 and allowing markets to clear at dates 1 and 2. Otherwise, it is said to be constrained-inefficient.

One obvious policy experiment is to regulate capital. The policy-maker is assumed to dictate to the intermediaries the amount of capital e_0 they must raise at date 0, but the policy-maker allows them to choose their portfolios and consumption plans freely. The market-clearing prices, $p(\lambda)$, are also endogenously determined. Since the first-best capital ratios are much higher than the capital ratios under incomplete markets, it is natural to ask whether increasing the capital ratio will improve welfare. To answer this question, we compute equilibria in which intermediaries are required to hold different amounts of capital. Table 3 shows the equilibrium values corresponding to $\rho = 2$, $R = 1.8$, and different required levels of capital e_0 ranging from 0.04 to 0.08. This is approximately equivalent to requiring capital ratios from 4 per cent to 8 per cent of assets. The first two lines of the table give the equilibrium values with complete and incomplete markets for comparison purposes.

The striking feature of this exercise is that increasing capital requirements does not increase expected utility. In fact, to increase welfare, the required policy must *reduce* capital.

6.5 Liquidity requirements

We can try a similar thought experiment by regulating the amount of the short-term asset held in equilibrium, while allowing the intermediaries to choose the other variables freely. Tables 4a and 4b show the equilibrium values corresponding to two values of R in each of three different settings: complete markets, incomplete markets, and incomplete markets with the value of θ constrained to equal the first best. Table 4a shows equilibrium values when the return to the long-term asset is $R = 1.8$, and Table 4b shows equilibrium values when $R = 1.5$. In each table, we list first the equilibrium with complete markets, next the equilibrium with incomplete markets, and finally, the equilibrium with incomplete markets in which the value of θ is constrained to equal the value in the first-best equilibrium.

In both cases, we see that the amount of capital, which is chosen freely by the intermediary, increases, though it does not reach the first-best level, and the amount of the short-term asset is lowered, since the first best is less than the level in the incomplete-markets equilibrium. The expected utility increases and comes reasonably close to the first-best level. Certainly the impact of this intervention on welfare is much greater than the impact of capital regulation.

These are trivial examples in toy models, but they raise interesting questions about what is going on. For example, what is the general-equilibrium effect of capital-adequacy regulation and what do we know about the effect of capital structure and portfolio choices of intermediaries on asset pricing? Until we know a lot more, we will not have a handle on the microeconomics underlying the optimal capital-regulation policy.

Conclusions

Every intermediary chooses the optimal actions on behalf of its depositors in equilibrium, taking prices as given. But intermediaries do not take account of the effect of their collective choices on prices. When the regulator steps in and forces everyone to choose a different capital level, the result is to change the equilibrium prices. If the allocation is Pareto-efficient, these pecuniary externalities cancel out and have no effect on welfare. But if the markets are incomplete, the allocation will not be Pareto-efficient, and so pecuniary externalities will not cancel out. It is possible that for some changes in prices, everyone will be made better off. It is the change in prices that accounts for the increase in welfare.

Why should an increase in capital at date 0 lead to a change in prices that increases welfare? The problem with the equilibrium prices in incomplete

Table 3
Equilibrium with regulated capital ratios

	ρ	R	e_0	θ	EU
Complete	2	1.8	0.1300	0.5635	0.2000
Incomplete	2	1.8	0.0602	0.6749	0.1835
$e_0 = 0.04$	2	1.8	0.0400	0.6465	0.1824
$e_0 = 0.05$	2	1.8	0.0500	0.6602	0.1841
$e_0 = 0.06$	2	1.8	0.0600	0.6745	0.1835
$e_0 = 0.07$	2	1.8	0.0700	0.6889	0.1806
$e_0 = 0.08$	2	1.8	0.0800	0.7027	0.1756

Table 4a
Equilibrium with regulated asset holdings

	ρ	R	e_0	θ	EU
Complete	2	1.8	0.1300	0.5635	0.2000
Incomplete	2	1.8	0.0602	0.6749	0.1835
$\theta = 0.5635$	2	1.8	0.0952	0.5635	0.1979

Table 4b
Equilibrium with regulated asset holdings

	ρ	R	e_0	θ	EU
Complete	2	1.5	0.0530	0.6185	0.1095
Incomplete	2	1.5	0.0310	0.6990	0.1034
$\theta = 0.6185$	2	1.5	0.0461	0.6185	0.1092

markets is that they are too volatile: when there is a high demand for liquidity (high realization of λ), lots of intermediaries are selling assets, and this depresses $p(\lambda)$, increasing the cost of liquidity. To provide better insurance, one wants a policy that will stabilize prices. The way to do this is to increase the amount of the short-term asset being held, more precisely, the amount of the short-term asset relative to the amount of the long-term asset. This, at least, is the intuition behind the argument that capital may be too low in laissez-faire equilibrium. But there are other factors that must be taken into consideration.

First, reducing risk is not the only objective. If the degree of risk aversion is not too high, it may be more important to take advantage of the higher returns from investing in the long-term asset and to reduce the consumption at date 1. In Allen and Gale (2004), it is shown that, if relative risk aversion is above a critical value, there is too little liquidity, and, if it is below, there is too much. In fact, the critical value of risk aversion is one, the same as the logarithmic utility function used for the numerical examples above.

Second, capital is costly. Investors must be compensated for their opportunity costs ρ even though the return on bank assets is lower, and this requires the depositors to give up some of their share of the returns. Evidently, the benefit of higher capital, in terms of stabilizing prices, is not enough to offset the cost of capital. The odd thing is that intermediaries, who take prices as given, choose to set the level of capital too high.

Clearly, we cannot explain everything in terms of a simple unicausal story. It may be that if relative risk aversion were much higher, the intuitive explanation given above would be the dominant factor, and an increase in capital would increase welfare. More research into these questions is needed.

References

Allen, F. and D. Gale. 1998. "Optimal Financial Crises." *Journal of Finance* 53 (4): 1245–84.

———. 2000a. "Financial Contagion." *Journal of Political Economy* 108 (1): 1–33.

———. 2000b. "Optimal Currency Crises." *Journal of Monetary Economics*, Carnegie-Rochester Conference Series on Public Policy 53: 177–230.

———. 2000c. "Bubbles and Crises." *Economic Journal* 110 (460): 236–55.

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

———. 2004. "Financial Intermediaries and Markets." *Econometrica* 72 (4): 1023–61.

Besanko, D. and G. Kanatas. 1996. "The Regulation of Bank Capital: Do Capital Standards Promote Bank Safety?" *Journal of Financial Intermediation* 5 (2): 160–83.

Chari, V.V. and R. Jagannathan. 1988. "Banking Panics, Information, and Rational Expectations Equilibrium." *Journal of Finance* 43 (3): 749–61.

Diamond, D.W. and P.H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.

Furlong, F.T. and M.C. Keeley. 1989. "Capital Regulation and Bank Risk Taking: A Note." *Journal of Banking and Finance* 13 (6): 883–91.

Gale, D. 2003. "Financial Regulation in a Changing Environment." In *Framing Financial Structure in an Information Environment*, edited by T. Courchene and E. Neave, Kingston, Ontario: John Deutsch Institute for the Study of Economic Policy, Queen's University.

Geanakoplos, J.D. and H.M. Polemarchakis. 1986. "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete." In *Uncertainty, Information and Communication*: *Essays in Honor of Kenneth J. Arrow, Volume 3*. Cambridge and New York: Cambridge University Press, edited by W.P. Heller, R.M. Starr, and D.A. Starrett, 65–95.

Hellmann, T.F., K. Murdock, and J. Stiglitz. 1998. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" Stanford Graduate School of Business Working Paper No. 1466.

———. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–65.

Jacklin, C.J. and S. Bhattacharya. 1988. "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications." *Journal of Political Economy* 96 (3): 568–92.

Kim, D. and A.M. Santomero. 1988. "Risk in Banking and Capital Regulation." *Journal of Finance* 43 (5): 1219–33.

Merton, R.C. 1977. "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees." *Journal of Banking and Finance* 1 (1): 3–11.

Postlewaite, A. and X. Vives. 1987. "Bank Runs as an Equilibrium Phenomenon." *Journal of Political Economy* 95 (3): 485–91.

Rochet, J.-C. 1992. "Capital Requirements and the Behaviour of Commercial Banks." *European Economic Review* 36 (5): 1137–72.

Wallace, N. 1988. "Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig Model with Sequential Service Taken Seriously." *Federal Reserve Bank of Minneapolis Quarterly Review* 12 (4): 3–16.

———. 1990. "A Banking Model in Which Partial Suspension Is Best." *Federal Reserve Bank of Minneapolis Quarterly Review* 14 (4): 11–23.