Chapter 17

Financial Intermediation

In this chapter we consider the problem of how to transport capital from agents who do not wish to use it directly in production to those who do. Some agents are relatively wealthy and already have all of the productive capital they need. Others accumulate capital for retirement, not production. In each case, the agents would want to lend their surplus capital to other agents, who would then use it in production. In the real world this lending takes the form of loans to individuals and businesses for the purpose of undertaking risky ventures.

Capital transportation of this form is known as *financial intermediation*. The institutions that stand between *savers* (those with surplus capital) and *borrowers* (those with less capital than they would like to use in their productive technology) are known as financial intermediaries. The most common financial intermediary is the bank, so the study of intermediation is sometimes also known as banking.

In this chapter we will examine how banks operate, starting from the bank's balance sheet, its role in matching lenders and borrowers and continuing on to its ability (or desire) to make loans. We will study competitive equilibria in banking. And we will consider the inherent instability built into many banking systems.

The balance sheet of a bank is a little unusual at first sight. The key fact to remember is that, for the bank, loans are assets, in exactly the same way that vault cash and government bonds are, while deposits are liabilities. Every financial instrument that is an asset on one balance sheet must be a corresponding liability on another balance sheet. For example, loans on a bank's balance sheet are assets, while those same loans are liabilities on the borrower's balance sheet.

We will consider a completely worked-out model with a competitive equilibrium in banking. This model, originally due to Williamson, provides several important insights. First, banks will act as aggregators of deposits, bundling together several small deposits to make one large loan, as we actually observe in practice. Second, some agents will be completely unable to get loans in equilibrium because the bank finds it too expensive to make loans to them. These agents are *credit rationed*, which acts something like a "credit crunch" in reality. Third, banks will use a *pure debt* contract with *default* in dealing with borrowers. That is, the bank will loan a borrower an amount, say $100 at an interest rate r, say 5%, and expect to be repaid $105 at the end of the period no matter how the borrower's finances have changed in that period. Even if the borrower's house burns down and employer goes bankrupt, the bank wants its $105. The borrower's only recourse is to declare bankruptcy, hand over all assets and consume nothing (or some very low amount). This is known as *defaulting.* The bank could have written a very different loan contract, something like, "Pay me $110 if your house doesn't burn down, but $10 if your house does burn down". Assuming that there is only a 5% chance that the borrower's house will burn down, the bank will in expected value get $105, and the borrower would vastly prefer such an insurance contract. The absence of such contracts must be explained. The explanation we use here is one of *moral hazard.* The bank has no way of knowing whether or not the borrower's house has really burned down without paying an *audit cost.* Thus the borrower always has an incentive to lie (hence the term "moral" hazard) and claim misfortune.

Next we turn our attention another model with moral hazard. In this model there are no audit costs, but agents will supply a secret amount of labor effort. Because labor effort is secret, lenders will not be able to directly contract on it, and borrowers will supply lower-than-optimal amounts of effort. Agents will also be of different wealth levels, which will allow us to think about how credit is provided to rich people as opposed to poor people. We will see that poorer agents will work less hard, pay a higher interest rate and default more frequently than richer agents. This effect will be so strong that certain very poor agents will be credit rationed. These very poor agents will save their meager assets.

Finally we will consider the celebrated model of *bank runs* by Diamond and Dybvig. Bank runs refer to financial panics in which depositors rush to their bank to liquidate their assets, usually because they doubt their bank's ability to make payments. The most famous bank runs happened during the Great Depression and indeed, according to Diamond and Dybvig, they might have greatly contributed to the economic collapse of that period. The model has continuing appeal because, although the American banking system has been somewhat insulated from panics, the banking systems of other countries continue to succumb to panics. It is an unfortunate fact of life that banking panics still plague us. In Diamond and Dybvig's model, agents will look out of their windows, see other agents running to the bank and be immediately compelled to also run. The first (luckier or fleeter) agents to the door of the bank withdraw all of their deposits, leaving nothing for the remaining (slower) agents.

17.1 Banking Basics

In this section we briefly review some important concepts in modern American banking. We begin with a discussion of accounting for banks. We then turn to the fractional reserve banking system, and examine how the government manipulates the money supply. We conclude this section with a discussion of an important banking reform proposal, that of narrow banking.

Assets and Liabilities of a Bank

Banks are businesses, and like all enterprises they incur liabilities and accumulate assets. The confusing thing about banks is that by accepting a deposit, which is after all an inflow of money to the institution, the bank has incurred a liability. Since a liability is an obligation to pay, the bank has, by accepting the deposit, promised to pay the depositor the amount of his deposit plus accumulated interest either on demand or at a particular time. In the same way, by making a loan, which is an outflow of money from the bank, the bank has accumulated an asset. An asset is a claim to payment, and by making the loan, the bank has a claim on some repayment schedule of principal plus interest. Not all loans are repaid, so the bank must estimate the expected value of loans that will not be repaid and count this against its assets. Thus bank balance sheets have an item marked “Outstanding Loans net of Loss Reserve”. This loss reserve is a polite term for the expected value of loan defaults.

A bank also holds, by law, a certain proportion of its deposits in zero-interest accounts with the Federal reserve system. These are also assets, although pretty low-yield ones. Banks hold a very small amount of “vault cash” which is currency (notes and coins) held at the bank (usually in impressive safes). This is used to meet the cash needs of depositors day-to-day. Banks also directly hold securities like U.S. Government debt (bonds). The nature of these securities is limited by law, so in the U.S. banks are not big stock market players. Banks also often directly own property, such as the bank building itself.

Like other business, banks make operating profits or losses as the value of assets and liabilities fluctuate. If a bank makes a profit, so that assets exceed liabilities, a residual liability is added to balance assets and adjusted liabilities. Thus profit is a liability. The opposite is true for losses.[1]

In this chapter we are going to assume that banks are zero cost enterprises with no assets other than loans and no liabilities other than deposits. We will assume that banks make zero economic profit in expectation, so the expected return on loans must cover the amount owed to depositors.

[1] For more information on this and other oddities in banking, see Mathis Dewatripont and Jean Tirole *The Prudential Regulation of Banks*, 1994, MIT PRess: Cambridge, MA.

Fractional Reserve Banking and the Money Supply

In the U.S. today, banks are required to hold a certain fraction of their deposits in reserve, that is, not lend them out (these reserves are held on deposit by the retail banks at Federal Reserve member banks). Since the reserve requirement is not 100%, banks may lend out the portion of their deposits not held on reserve. Reserve requirements are indexed by the nature of deposits and loans, so banks have to hold extra reserves against riskier loans, but let us suppose for a moment that they are constant at 10%. Imagine that the government prints \$100 and gives it to household 0, so $H_0 = 100$. This household immediately places it in the bank or spends it. Any amount spent must go as a payment to some other household, which then faces the same choice: deposit or spending. And so on, until the banking system has absorbed the entire \$100 of cash.

In this way, the fractional reserve banking system can multiply an infusion of cash, augmenting the M1 money supply by more than the infusion. Reserves held at the Fed will also affect the money supply in the same way. For this reason, cash and reserves held at the Fed are often called *base money* or *high-powered money.* The sum of all cash and deposits held at the Fed is called the *monetary base.*

Assume for simplicity that each household simply deposits the cash. Assume also that there is only one common bank. Now since the reserve requirement is 10%, the bank places \$10 of its new deposits on reserve at the local Federal Reserve system member bank. The remaining \$90 it lends out again to some other household, household 1, so $H_1 = 90$. This household spends or deposits the money, as before, so a further \$90 of deposits appear in the bank. Now the bank sends \$9 to the Federal Reserve, and lends out \$81 to household 2, so $H_2 = 81$. This process continues until the bank is lending out, to household i an amount H_i:

$$H_i = \$100(1 - 0.10)^i.$$

The amount of *new money* created is just the sum of all loans made to households as a result of the original \$100 transfer, plus that \$100. That is:

$$M' - M = \sum_{i=0}^{\infty} H_i = \sum_{i=0}^{\infty} 100(1 - 0.10)^i = 100 \sum_{i=0}^{\infty} 0.90^i = 100 \frac{1}{1 - 0.90} = 1000.$$

Here M' is the new stock of money and M is old stock of money. Thus via the fractional reserve banking system, the \$100 initial cash transfer of the government grows to be \$1000 worth of new deposits in the banking system, which can then be used for payments purposes.

Targeting “The” Interest Rate

The U.S. government does not, as a rule, print money and then hand it out to agents (as we shall see in Chapter 18, however, other governments do exactly that). Instead, monetary policy is controlled by the Federal Reserve system (known as the "Fed"). The specific instrument of monetary policy most commonly used are *open-market operations* which the Fed uses to target "the" interest rate. The level of "the" interest rate is closely watched by industry and the media.

The interest rate that is at the center of all this fuss is called the "Fed funds rate". It is associated with a very specific and abstract credit contract that is executed only between banks. No private individual has ever executed a Fed funds contract. Here is how the Fed funds market works: Banks are constantly gaining and losing deposits as individuals move, experience good or bad fortune or die. At the close of business each day, banks must meet their reserve requirement at the Fed. Imagine that a bank, say Hyde Park Bank, sees a sudden surge in deposits one day (say the first day of classes at the University of Chicago). Hyde Park Bank must place more funds on reserve at the Fed to meet its reserve requirement. Since deposits at the Fed earn no interest, it is likely that Hyde Park Bank does not have any excess reserves. Moreover, it takes time and money to make a deposit at the Fed. To meet the sudden shortfall in required reserves, Hyde Park Bank turns to other banks, who might have surplus reserves (caused by a sudden outflow of deposits). Thus Hyde Park Bank borrows reserves held at the Fed from other banks. Reserves held at the Fed are called Fed funds. The market for these reserves is called the Fed funds market. The interest rate on Fed funds (the interest rate that Hyde Park Bank pays) is the Fed funds rate. It is "the" interest rate.

Recall that deposits at a bank are liabilities. Thus Fed funds, an asset of banks, are a liability of the government. However, they are not an onerous liability since the government does not pay interest on these deposits. The Federal reserve system is the banker's bank, and its books must also balance.

The Fed does indeed announce a target for the Fed funds rate. However, it affects the rate with more than just the moral suasion of an announced target: the Fed affects the supply of Fed funds directly through open market operations (OMOs). In an OMO *purchase* the Fed trades Fed funds for assets held at banks. Thus the Fed approaches a bank, say Hyde Park Bank, which holds millions of dollars of U.S. Government debt (bonds) and offers to buy a thousand dollars of these bonds. How will the Fed pay for these bonds? With "store credit", that is, by creating $1000 of new reserves held at the Fed in the name of Hyde Park Bank. Hyde Park Bank is neither a net loser or a net gainer from this operation, since it has traded $1000 of one safe asset (government bonds) for $1000 of another (Fed funds). What will Hyde Park Bank do with the new Fed funds? It could in principal trade them for cash. For this reason, we say that cash is a *liability of the government* in exactly the same way that Fed funds are. However, Hyde Park Bank will probably just float them on the Fed funds market, driving down the Fed funds rate. Thus OMO purchases are associated with decreased Fed funds rates and an increased money supply (since Hyde Park Bank can trade its new Fed funds for cash). This is an *expansionary* monetary policy.

In the opposite way, an OMO *sale* tends to increase the Fed funds rate and decrease the money supply. In an OMO sale, the Fed approaches Hyde Park Bank and offers to sell it some government bonds which the Fed holds. The Fed accepts as payment Fed funds.

Thus the stock of Fed funds decreases and the Fed funds rate increases. Since Hyde Park Bank had to give up an asset that could be traded directly for cash (Fed funds), OMO sales cause a decrease in the money supply. This is a *contractionary* monetary policy.

Narrow Banking

In modern America, the banking system has two jobs. Intermediation is one job (and the primary focus of this chapter). The other job is acting as a *payments system.* That is, deposits at banks can be accepted as legal payment for goods and services, often in the form of checks (drafts on deposits). Although we are accustomed to thinking of these two jobs as linked, there is absolutely no inherent reason why they should be. In fact, true banking reform would institute so-called *narrow banking*, in which banks are not allowed to lend. They would have to back deposits completely with a riskless asset (usually government bonds). This is equivalent to a 100% reserve requirement. In this scenario, banks would concentrate only on providing low-cost payments services like checks, smart cards, direct deposit and direct withdrawal.

In a world of narrow banking, households seeking loans, for example a home mortgage, would not approach banks. Neither would households seeking a greater return on their investment than that provided by the narrow banks. Instead they would approach non-bank financial intermediaries. These intermediaries would be like mutual funds: savers would own an equity share of the mutual fund, which would have some constantly quoted price. Borrowers would negotiate with the mutual fund in much the same way that they would with a bank. The key difference is that the value of the mutual fund would not be fixed, like the value of deposits at a bank. In good economic times the mutual fund will perform better than in bad economic times. As we shall see in the section on bank runs below, this kind of arrangement effectively prevents financial panics.

Payments Systems

Before turning to models of intermediation, it is worth briefly discussing the payments system[2]. The payments system is the infrastructure, law and custom governing our ability to pay for goods and services. The most common methods of payment by households are cash, check and credit card. Of these three methods, only cash requires no up-front investment by the household. To have a checking account costs (on average) $80 a year, and since credit cards are a form of unsecured debt, households must frequently pay some minimum amount each year to use one. Households too poor to afford either must use so-called "Currency Exchanges", which are ubiquitous features of working-class neighborhoods. These institutions will cash paychecks and welfare checks (for a fee), allow households to pay

[2]The numbers in this subsection are drawn from the paper "Retail Payments Instruments: Costs, Barriers, and Future Use" by David Humphrey and Lawrence Pulley, presented at the Conference on Bank Structure and Competition in Chicago, May 7, 1998.

their utility bills in cash (for a fee) and write “money orders” (a form of check), again for a fee. A year’s worth of transactions at one of these exchanges can easily cost much more than $80, but the costs are directly contingent upon use, unlike checking accounts or credit cards, and do not require anything in the way of legal documentation.

Although there is about $400 billion in U.S. currency (cash) outstanding, only 10%-15% of this total cash (based on surveys of households) is used by households. Over 60% of total cash is thought to be outside the U.S. In the U.S., roughly 70% of all transactions are settled in cash, while the same figure is 78% in Holland, 83% in Finland, 86% in Germany and 90% in the United Kingdom. The average value of these transactions is small, below $10.

In the U.S., of all non-cash transactions, 75% are settled by check, and the average check is for $1,158. About 20% are settled by credit card, with an average value of $61. It is interesting to note that although less than 0.5% of transactions are settled by wire transfer, the average size of a wire transfer is over $4.2 million. The relatively large values of these figures stems from the fact that they are used by the government and other institutions. Households’ non-cash transactions average below $50.

The total cost of the U.S. payments system to payors, banks and payees is $204 billion a year (about 3% of GDP) or roughly $1050 per adult. Each non-cash payment costs on average $2.60.

Finally, although the check system is deeply entrenched in the U.S., the new electronic payments means are substantially cheaper, costing a third to a half as much as checks.

Looking across countries, the average American made 326 non-cash transactions in 1996. The corresponding figure for Canada is 151, for Europe it is 126 and for Japan a paltry 40. Of these, checks accounted for 244 of the transactions in the U.S., 62 in Canada, 31 in Europe and only two in Japan. Thus check clearing is much more common the U.S. than even in Canada. These other countries rely more heavily on debit cards and other electronic means (although the total number of non-cash transactions abroad is still well below the level in the U.S.).

For most of the post-war period, the U.S. has led the world in payments system technology. In other countries it is still quite common to pay even very large bills with cash. However, the backbone of the domestic payments system, the check clearing system, is showing its age. Other countries are already using sophisticated electronic systems such as debit cards and electronic check presentation, that have yet to become common here. Although most economists do not consider payments systems directly relevant to intermediation and the conduct of monetary policy, there is no question that they have a large and direct impact on all households in an economy.

17.2 A Model with Costly Audits

In this section we will consider a model that replicates several of the important features of financial intermediation. This model is originally due to Williamson (1987). Intermediaries will arise endogenously to evaluate the credit worthiness of borrowers, to bundle small deposit together to make a large loan, to minimize the cost of monitoring borrowers and to spread default risk across several lenders, so they are fully insured against default by borrowers.

This model is exciting because we are, for the first time, going to calculate a competitive equilibrium in the capital market with heterogeneous agents. In addition, this capital market is going to suffer from a realistic problem: audit costs. As a result, in equilibrium, some agents will be unable to get loans no matter what interest rate they promise to pay, so there will be *credit rationing* in equilibrium. Also, the audit costs are going to produce very familiar credit contracts, the fixed-obligation loan.

Agents

In this model there will be two kinds of agents. Type-1 agents, who form a proportion α of the population, will be *workers.* Type-2 agents, who form the remaining $1 - \alpha$ proportion of the population, will be *entrepreneurs.*

There will be two periods. In the first period, type-1 agents work some amount, consume and then save. Intermediaries accept these savings as deposits and use them to make loans to type-2 agents. Type 2 agents will operate their risky technologies in the "morning" of the second period, realize their outcomes and repay the intermediaries. The intermediaries will then take these payments and repay their depositors. There is no way to store the consumption good between periods.

Type-2 agents are born with a number stenciled on their foreheads. This gives their *audit cost.* That is, if a type-2 agent is born with γ_i on his forehead, it costs γ_i for any other agent to observe the outcome of his technology (more on that below). Note that the audit cost γ is public. These audit costs are distributed uniformly on the interval $[0, 1]$. Thus 25% of all type 2 agents will be born with audit costs below 0.25, and the remaining 75% of type-2 agents will have audit costs greater than 0.25.

Preferences

Type-1 agents care about consumption in both periods. Their preferences over consumption c_0 and c_1 and labor effort exerted in the first period ℓ_0 are as follows:

$$U^1(c_0, \ell_0, c_1) = u(c_0, \ell_0) + c_1. \tag{17.1}$$

Here $u_1(\cdot,\cdot) > 0$, $u_2(\cdot,\cdot) < 0$ and $u_{22} < 0$.

Type-2 agents care only about consumption in the second period. Their preferences over consumption c_1 are simply:

$$U^2(c_1) = c_1. \tag{17.2}$$

So type-2 agents are risk neutral in consumption and supply no labor effort.

Technology

Type-1 agents may work up to h hours while young (in the first period of life), where $h < 1$. Labor effort is transformed directly into the consumption good (so that the implicit "wage rate" is just unity). The risk-free interest rate offered by the financial intermediaries is r, so type-1 agents face budget constraints of the form:

$$c_0 = \ell_0 - s, \text{ and:} \tag{17.3}$$

$$c_1 = (1 + r)s. \tag{17.4}$$

Where $\ell_0 \leq h$ and s is the savings of a type-1 agent.

All type-2 agents have access to the same technology, no matter what their audit cost is (that is, they differ only in their audit cost). In exchange for a capital input of $k = 1$ while young, type-2 agents will operate a risky technology that produces output of y, where, for a particular agent i, output y_i is:

$$y_i = 1 + \varepsilon_i.$$

The idiosyncratic shock term ε_i is distributed uniformly on the interval $[0, 1]$. So the mean or average shock is 0.5 and mean output is 1 + 0.5 = 1.5.

Notice that type-2 agents have no intrinsic wealth of their own–they must get a loan of size $k = 1$ to finance their projects. Since the absolute maximum that a type-1 agent can produce, by working full-time ($\ell_0 = h$) is h, which by assumption is less than unity, a type-2 agent's project can be financed only by a loan from several type-1 agents. Thus intermediaries are going to have to aggregate deposits from several savers to make a single productive loan, which matches the real world experience well.

Finally, the output y_i of a particular type-2 agent's technology is *private* to that agent. Only agent i may costlessly observe y_i. All other agents must pay the audit cost γ_i stenciled on agent i's forehead to observe y_i.

Intermediation

Although we will not explicitly model the industrial organization of banking in this paper (that is, we will not write down how banks are formed, who owns them and so on), it is easy to see what the optimal banking structure will look like. Since the only cost to forming a bank is auditing defaulting borrowers, any agent can declare himself a bank, accept deposits and make loans. The larger the bank, the more insurance that depositors will have against defaults. Imagine a bank that matched n depositors, each of whom deposits some amount $1/n$ with exactly one borrower. There is no insurance at all for those n depositors. If the borrower defaults, then all n depositors lose everything. Now imagine that a second bank opens across the street from this first bank. This bank seeks to match $2n$ depositors with two borrowers. If one borrower defaults, depositors may still recover something from the other borrower. The second bank will clearly attract more depositors than the first bank. In this way, larger and larger banks will form, until all the potential depositors go to the same bank, which also makes all of the loans.

Now let us turn our attention to the question of optimal loan contracts. The bank advances a type-2 agent a loan of $k = 1$. The agent then experiences output y, which is secret. The bank is interested in minimizing the expected audit cost it will have to pay. One contract would be for the borrower to report his output, and for the bank to take, say 10% of that output as payment for the loan, and the bank never audits. The problem with this contract is that borrowers will always announce that $y = 1$, the minimum possible output. Since the bank is not auditing, there is no way for it to dispute this claim.

The bank wishes to minimize the cost of audits while at the same time ensuring that borrowers tell the truth. It turns out that the best contract for this purpose is one in which the bank announces some required repayment level X. No matter what output is, the bank insists on getting its amount X. Any borrower who announces output $y < X$ is declared bankrupt and audited. Whatever output that agent had produced (no matter how large or small) is seized by the bank. Borrowers with outcomes $y < X$ have no incentive to announce anything other than the truth, while borrowers with outcomes $y > X$ know they must always pay X or be audited, so they merely announce y, pay X and consume the residual $y - X$.

It is interesting to note that is exactly the kind of debt contract that is actually written. Lenders make loans at some interest rate, and expect to be repaid the loan amount plus interest. If borrowers do not repay this amount, they are declared bankrupt and legal proceedings begin. Of course in our society bankrupt agents do not consume zero, and in fact their minimum consumption varies from state to state, but the underlying principle is the same. Risk averse agents would prefer more complicated credit contracts, ones that provided a certain amount of insurance along with the loan. For example, most homeowners would be willing to pay above-market interest rates on their mortgages if their bank agreed to cut their mortgage payments if something bad happened to the homeowner (loss of job, injury etc). The reason such contracts are not more common is because output is largely hidden. Lenders want to be assured of repayment without having to closely monitor their borrowers (an expensive proposition).

Let us return to the problem of a financial intermediary making a standard loan of size $k = 1$ to a particular borrower with known audit cost γ, requiring repayment X and auditing the borrower if output falls below X. Write X as $1 + x$, so x is the net interest rate paid by borrowers. In addition, recall that output y is always between 1 and 2, depending on the value of the shock term ε. The actual revenue of the intermediary conditional on the repayment amount x, the audit cost γ, and the shock term ε is:

$$\pi(x,\gamma,\varepsilon) = \begin{cases} 1+x, & \varepsilon \geq x \\ 1+\varepsilon-\gamma, & \varepsilon \leq x. \end{cases}$$

The bank knows that the output shock ε is distributed uniformly on the interval $[0,1]$. Assuming that $0 \leq x \leq 1$, the expected revenue of the intermediary conditional only on x and γ is:

$$\begin{aligned} \pi(x,\gamma) &= \int_0^1 \pi(x,\gamma,\varepsilon)\, d\varepsilon \\ &= 1 + \int_0^x (\varepsilon-\gamma)\, d\varepsilon + x\int_x^1 d\varepsilon \\ &= 1 + \left(\frac{x^2}{2} - \gamma x\right) + x(1-x) \\ &= 1 + (1-\gamma)x - \frac{x^2}{2}. \end{aligned}$$

It is easy to see that expected revenues $\pi(x,\gamma)$ as a function of x are a parabola with a peak at $x^\star(\gamma) = 1-\gamma$. No bank would ever charge a repayment amount x greater than $x^\star(\gamma)$, since revenue is declining in x beyond that point, and borrowers are worse off.

Let $\pi^\star(\gamma)$ be the absolute maximum amount of revenue that the bank can accumulate as a function of the audit cost, γ. That is:

$$\pi^\star(\gamma) = \pi\left[x^\star(\gamma),\gamma\right] = 1 + \frac{1}{2}(1-\gamma)^2.$$

Thus for agents with $\gamma = 1$, the bank's maximum revenue occurs when $x^\star(\gamma = 1) = 0$ and the bank never audits, producing a revenue of $\pi^\star(\gamma = 1) = 1$ with certainty. For agents with $\gamma = 0$, the maximum revenue that the bank can extract occurs when $x^\star(\gamma = 0) = 1$ and the bank's expected revenue is $\pi^\star(\gamma = 0) = 3/2$.

Now let us turn our attention to the liabilities of the bank. The bank owes its depositors an amount $1 + r$ on a unit loan. Thus for each borrower of audit cost γ, banks will pick the lowest value of x such that $\pi(x,\gamma) = 1 + r$. There will be some agents, with relatively high values of γ, for which banks will never be able to realize an expected return of $1 + r$. That is, those agents with audit costs $\gamma \geq \gamma^\star(r)$ such that:

$$\pi^\star[\gamma^\star(r)] = 1 + r.$$

will be *credit rationed.* Banks will never make them a loan, they will be squeezed out of the credit market and their projects will not be funded. It is easy to see that $\gamma^\star(r)$ is given by:

$$\gamma^\star(r) = 1 - \sqrt{\frac{r}{2}}.$$

Thus when $r = 0$, $\gamma^\star = 1$ and no agents are credit rationed. As r increases, more agents are credit rationed, and at the astronomical interest rate of $r = 2$, all agents are credit rationed.

Equilibrium

It is easy to see now how equilibrium in this economy will be achieved. There will be a demand for capital, which is decreasing in r, and a supply of capital, which is increasing in r. The demand for capital is given by the number of agents who are not credit rationed at the interest rate r. That is, the proportion of type-2 agents (remember, they make up $(1-\alpha)$ of the total) with $\gamma \leq \gamma^\star(r)$. Thus the aggregate demand for capital is given by:

$$K^d(r) = (1-\alpha)\gamma^\star(r) = (1-\alpha)\left(1 - \sqrt{\frac{r}{2}}\right).$$

As r decreases, more and more type-2 agents may finance their projects. Each type-2 agent always wants exactly one unit of capital, so individual borrowing is constant but aggregate borrowing increases.

The supply of capital comes from the saving schedules of type-1 agents. As r increases, type-1, or worker, agents each save more. Say that the savings schedule of workers is given by $S(r)$. Thus the supply of capital is:

$$K^s(r) = \alpha S(r).$$

Equilibrium in the capital market will occur at the interest rate $r^\star$ at which:

$$K^s(r^\star) = K^d(r^\star).$$

That is, where the supply of capital from type-1 agents' savings equals the demand for capital from intermediaries lending to type-2 agents, so they can finance their projects.

An Example

Imagine that type-1 agents had preferences over consumption while young c_0, labor effort while young ℓ_0, and consumption while old c_1 given by:

$$U^1(c_0, \ell_0, c_1) = u(c_0, \ell_0) + c_1 = 2\sqrt{c_0} + 2\sqrt{h - \ell_0} + c_1.$$

Recall from the budget constraints, equations (17.3) and (17.4), that, by substituting in the savings term s, this becomes:

$$U^1(\ell_0, s) = 2\sqrt{\ell_0 - s} + 2\sqrt{h - \ell_0} + (1+r)s.$$

We take derivatives with respect to ℓ_0 and s to find the first-order necessary conditions for maximization:

$$\frac{1}{\sqrt{\ell_0 - s}} - \frac{1}{\sqrt{h - \ell_0}} = 0, \text{ and:}$$
$$-\frac{1}{\sqrt{\ell_0 - s}} + (1 + r) = 0.$$

We can solve these to find the savings schedule of type-1 agents:

$$s(r) = h - \frac{2}{(1 + r)^2}.$$

The aggregate supply of capital is thus:

$$K^s(r) = \alpha \left[h - \frac{2}{(1 + r)^2} \right].$$

Notice that $K^s(r_0) = 0$ where $r_0 = \sqrt{2/h} - 1$.

For equilibrium to occur with this specification of preferences, the interest rate at which there is zero demand for capital must exceed the interest rate at which there is zero supply. The zero-demand interest rate we know from above to be $r = 2$. The zero-supply interest rate is $r = r_0$. Thus for equilibrium:

$$\sqrt{\frac{2}{h}} - 1 < 2.$$

It is easy to see that this means that $h > 2/9$. Indeed, as h gets closer and closer to $2/9$, the supply curve shifts upward. This in turn causes the equilibrium interest rate to rise and the equilibrium supply and demand of capital to fall, decreasing the number of projects that are undertaken.

17.3 A Model with Private Labor Effort

The model of the previous section was very useful in thinking about equilibrium in the credit market. However, since in the real world we do not necessarily recognize "type-1" and "type-2" agents from birth, and we do not readily observe different audit costs, it makes sense to consider a different model. In this model all agents will be identical except for their wealth level. Some agents will be richer, others poorer. There will be no audit cost, and output will be public. However, agents are going to have to work to make the project succeed, and the amount of their labor effort will be *private* (that is, hidden). It can never be known (cannot be audited). This is another example of moral hazard.

Technology, Endowments, Preferences

All agents will have access to a common "back-yard" technology which will map capital k and labor effort ℓ into a probability that the project succeeds. If the project succeeds, output is high, at q. If it fails, output is zero. Capital k can take on only two values: $k = 0$ or $k = 1$. If $k = 1$ then the probability of the high output is just given by the level of labor effort, ℓ. If $k = 0$ then the low output occurs with certainty, no matter what the effort level was.

Agents are all endowed with some level of wealth a. For a particular agent, if $a < 1$, then that agent must get a loan of size $1 - a$ to operate the technology. If $a > 1$, then the agent can finance the technology alone and lend the surplus $a - 1$.

Agents have preferences over consumption c and labor effort ℓ given by:

$$c - \frac{q}{\alpha}\frac{\ell^2}{2}.$$

Here $0 < \alpha < 1$ and q is just the high output level.

Rich Agents

Assume that there is some riskless rate of return r on wealth. A rich agent, one with $a > 1$, can finance the project from her own wealth and lend the remainder at this interest rate r. How much effort does she supply? If the project fails and output is zero, she consumes $c = 0+(1+r)(a-1)$. If the project succeeds and output is q, she consumes $c = q+(1+r)(a-1)$. The project succeeds with probability ℓ, her labor effort. Thus her maximization problem is:

$$\max_{\ell}\left\{\ell[q + (1 + r)(a - 1)] + (1 - \ell)[(1 + r)(a - 1)] - \frac{q}{\alpha}\frac{\ell^2}{2}\right\}.$$

The first-order condition for maximization with respect to ℓ is:

$$q - \frac{q}{\alpha}\ell = 0.$$

This implies that $\ell^{\star}(a) = \alpha$ for $a \geq 1$. That is, agents wealthy enough to finance the project out of their own funds all supply the same labor effort, α.

Given that all rich agents supply $\ell = \alpha$ regardless of their wealth, we can easily calculate their expected utility by plugging back in:

$$\alpha q + (1 + r)(a - 1) - \frac{q}{\alpha}\frac{\alpha^2}{2}, \text{ or:}$$

$$\frac{\alpha q}{2} + (1 + r)(a - 1).$$

Thus the value from operating the technology is $\alpha q/2$, and the value from lending any excess capital is $(1 + r)(a - 1)$. The opportunity cost of the capital used in the productive process is $(1 + r) \cdot 1$ (since production requires one unit of capital). Thus for any agent to undertake the project, it must the case that:

$$\frac{\alpha q}{2} \geq 1 + r.$$

For the rest of this section we are going to assume that the interest rate r is below $\alpha q/2 - 1$, so agents will want to use the productive technology.

Poor Agents

Now consider the much more interesting case of a poor agent, with wealth $a < 1$, who seeks a loan from a financial intermediary of size $1 - a$ in order to finance the project. For now we will assume that the intermediary charges the borrower an amount X if the project succeeds and zero if it fails. This is again the root of the moral hazard problem: the borrower only repays the bank if her project succeeds. This decreases the incentive for the borrower to work.

Now if the project succeeds, the agent consumes $c = q - X$, that is, output q net of repayment X. If the project fails, the agent consumes $c = 0$. Thus the agent is repaying only in the state when the project succeeds. The agent's choice of effort comes from the maximization problem:

$$\max_{\ell} \left\{ \ell(q - X) + (1 - \ell)(0) - \frac{q}{\alpha}\frac{\ell^2}{2} \right\}.$$

The first-order necessary condition for maximization with respect to effort ℓ is:

$$(q - X) - \frac{q}{\alpha}\ell = 0.$$

Solving for $\ell^\star(X)$ (effort as a function of repayment) produces:

$$\ell^\star(X) = \alpha\left(1 - \frac{X}{q}\right).$$

Notice immediately that $\ell^\star(0) = \alpha$ and that $\ell^\star(X)$ is decreasing in X. Hence poor agents, borrowers, will work less hard than rich agents, lenders.

Intermediaries

Now consider the problem of the intermediary raising deposits and making loans to poor agents. If this intermediary makes a loan of size $1 - a$, it must pay its depositors an amount

$(1+r)(1-a)$ on this loan. Thus its expected return on the loan must be equal to $(1+r)(1-a)$, the risk-free cost of capital. If the project succeeds the bank makes X, if it fails, the bank makes 0. The probability of success is ℓ, which cannot be observed or controlled directly by the bank. The bank must take the agent's effort choice as a function of repayment $\ell^\star(X)$ as given. So the bank's zero profit condition is:

$$\ell^\star(X) \cdot X + [1 - \ell^\star(X)] \cdot 0 - (1+r)(1-a) = 0.$$

Substituting the borrower's choice of labor effort conditional on repayment, $\ell^\star(X)$, in from above, the bank's zero-profit condition becomes:

$$X\alpha \left(1 - \frac{X}{q}\right) - (1+r)(1-a) = 0.$$

This is a quadratic equation in X and may be written as:

$$X^2 - qX + \left(\frac{q}{\alpha}\right)(1+r)(1-a) = 0.$$

Using the quadratic formula, we obtain two possible values for $X(a)$, that is, repayment as a function of wealth:

$$X(a) = \frac{q \pm \sqrt{q^2 - 4\left(\frac{q}{\alpha}\right)(1+r)(1-a)}}{2}$$

Since competitive pressures will force intermediaries to charge the lowest possible value of X, we concentrate on the lower branch. Notice that $X(1) = 0$.

Consider the term inside the radical in the definition of $X(a)$ above:

$$q^2 - 4\left(\frac{q}{\alpha}\right)(1+r)(1-a).$$

Notice that if a is small and r large, then this term might be negative. This means that, for poor borrowers, there is no value of X for which the intermediary can recover the cost of making the loan. Call this critical wealth $a^\star(r)$. Notice that $a^\star(r)$ satisfies:

$$a^\star(r) = 1 - \frac{\alpha q}{4(1+r)}.$$

Agents with wealth below $a^\star(r)$ are *credit rationed.* What will these agents do? They will become so-called *poor savers* and invest their meager funds in the economy-wide mutual fund.

Thus there will be three classes of agents in this model: the poor, who wish to borrow but cannot, and so save; the middle-class, who cannot self-finance but can borrow; and the rich, who self-finance and save the remainder of their wealth. Thus the demand for loans comes entirely from the middle-class, while the supply of funds comes from both the rich and the very poor. Increases in the interest rate will help the rich and the poor, and hurt the middle class.

17.4 A Model of Bank Runs

Now we turn our attention to the celebrated model of bank runs from the paper "Bank Runs, Deposit Insurance, and Liquidity" by Diamond and Dybvig. As Diamond and Dybvig say:

> Bank runs are a common feature of the extreme crises that have played a prominent role in monetary history. During a bank run, depositors rush to withdraw their deposits because they expect the bank to fail. In fact, the sudden withdrawals can force the bank to liquidate many of its assets at a loss and to fail. In a panic with many bank failures, there is a disruption of the monetary system and a reduction in production.

This neatly sums up the important features of the discussion. Bank runs can be a self-reinforcing phenomenon: if one agent sees others running for the bank, she must also join the run or face the certain loss of her wealth if the bank should fail. This sudden demand for cash (also called the demand for *liquidity*) causes the bank to sell (or *liquidate*) its assets (loan portfolio) at a loss, so it may be unable to satisfy the demands of its depositors. The real loss caused by premature liquidation is the fundamental reason why bank runs are bad.

How can bank runs be stopped? The authors consider two possibilities. First, *suspension of convertibility*, in which a bank temporarily refuses to cash out deposits. This is also known as a *bank holiday* and was quite common in the financial panics of the Great Depression. We shall see that bank holidays will only work under special conditions. Second, the authors consider *deposit insurance*, in which the government taxes all agents in order to honor banks' obligations. Deposit insurance will prevent bank runs even under very general conditions, and so we conclude that they are a more robust way of preventing bank runs. The authors do not consider the possibility of replacing the bank with a *mutual fund*, but, as we shall see, this too would prevent bank runs.

In this section we will consider a very different reason for banks to exist than in the previous two sections. Earlier, we viewed banks as institutions for getting capital from rich savers to poor borrowers (roughly speaking). In this section, all agents will have identical wealth and productive opportunities, but they will differ in the timing of their demands for consumption. Some agents will be content to consume later, while others will want to consume immediately. Agents will not know their type when they invest. Think of it this way: all agents are perfectly identical, except that some have cancer. Cancer diagnoses are announced only after all agents have made their investment decisions. When they are diagnosed, the cancerous agents want to consume immediately, while the noncancerous agents are content to wait until the following period to consume. The productive technology is (as we shall see) illiquid, so the cancerous agents are forced to prematurely liquidate. Banks will convert their illiquid assets into liquid liabilities. In doing so, the bank will leave itself open to the possibility of a bank run. If there is a bank run, then all assets are prematurely liquidated and there is real economic harm done.

As in the previous two models that we have considered, there is an informational problem. Here, there will be no way for banks to distinguish agents who truly have urgent consumption needs (our "cancerous" agents) and those who do not (the "noncancerous" agents).

Technology, Endowments and Preferences

There are three periods, $t = 0, 1, 2$. There is a single homogeneous good, and agents are endowed in period $t = 0$ with one unit of this good. There is a common, riskless, technology which converts a unit invested in $t = 0$ to $F > 1$ units in period $t = 2$. If the technology is interrupted in the middle period, $t = 1$, the *salvage value* is just the unit again. Think of this as a "growing turnip" technology. All agents are endowed with a turnip at birth in period $t = 0$, which they plant. If they uproot the turnip in the second period of life, $t = 1$, they just get their original turnip back. If they leave the turnip in the ground all the way to the harvest date of $t = 2$, it will have grown to $F > 1$ turnips.

Agents have no desire to consume in period $t = 0$. Let c_1 be consumption in period $t = 1$ and c_2 be consumption in period $t = 2$. Agents will have preferences of the form:

$$U(c_1, c_2; \Theta) = \begin{cases} \ln(c_1), & \text{if } \Theta = 1 \\ Q\ln(c_1 + c_2), & \text{if } \Theta = 2. \end{cases}$$

Here $1 \geq Q > F^{-1}$. The term Q is less than one. As a result, agents with $\Theta = 2$ (noncancerous) have a lower marginal utility than agents with $\Theta = 1$ (cancerous). So not only do agents with $\Theta = 1$ have to consume in period $t = 1$, they have a high marginal utility to boot. This is sometimes known as being "urgent to consume".

Agents of type $\Theta = 1$ have no desire to consume in period $t = 2$ at all, while agents of type $\Theta = 2$ are indifferent between consumption in $t = 1$ and $t = 2$. An agent with $\Theta = 1$ has cancer and must consume while young while one with $\Theta = 2$ does not have cancer and is indifferent between consumption while young or old.

Assume that there is some probability θ of having $\Theta = 1$ (that is, θ is the probability of getting cancer). With probability $1 - \theta$, $\Theta = 2$. Assume further that there is a continuum of agents, so a *proportion* θ will get cancer and the remaining $1 - \theta$ will not.

Optimal Insurance Contracts

Let c_t^i be the consumption of an agent of type i in period t. Without banks, because one's type is private, there can be no insurance contracts, so all agents with $\Theta = 1$ uproot their turnip in $t = 1$ and consume $c_1^1 = 1, c_2^1 = 0$. All agents with $\Theta = 2$ leave the turnip in the ground until $t = 2$ and consume $c_1^2 = 0, c_2^2 = F$.

Since agents are risk averse, they would prefer to be assured of some consumption between the low level of 1 and the high level of F. If Θ were public (that is, commonly observed), zero-cost insurance companies would provide agents with insurance contracts. Optimal insurance contracts would have the feature that $c_2^1 = c_1^2 = 0$ since agents with $\Theta = 2$ are content to wait. The budget constraint of the insurance company (equation 1c in Diamond and Dybvig) is a little hard to understand at first glance. Think of it like this: a proportion θ of the population will get c_1^1. This leaves $1 - \theta c_1^1$ in the ground between period $t = 1$ and $t = 2$, where it grows to $F(1 - \theta c_1^1)$. This is then spread between the remaining $1 - \theta$ of the population. Thus c_2^2 must satisfy:

$$c_2^2 \leq F \frac{1 - \theta c_1^1}{1 - \theta}.$$

This is equivalent to:

$$\theta c_1^1 + \frac{(1-\theta)c_2^2}{F} \leq 1. \tag{17.5}$$

The insurance companies' Lagrangian is:

$$\mathcal{L}(c_1^1, c_2^2, \lambda) = \theta \ln(c_1^1) + (1-\theta) Q \ln(c_2^2) + \lambda \left(1 - \theta c_1^1 - \frac{(1-\theta)c_2^2}{F}\right).$$

This has first-order conditions of:

$$\frac{\theta}{c_1^1} - \lambda\theta = 0,$$

$$Q\frac{1-\theta}{c_2^2} - \lambda \frac{1-\theta}{F} = 0, \text{ and:}$$

$$\theta c_1^1 + \frac{(1-\theta)c_2^2}{F} = 1.$$

We can solve these equations for the *optimal consumptions,* call them $c_1^{1\star}$ and $c_2^{2\star}$. We find:

$$c_1^{1\star} = \frac{1}{\theta + Q(1-\theta)}, \text{ and:} \tag{17.6}$$

$$c_2^{2\star} = \frac{QF}{\theta + Q(1-\theta)}. \tag{17.7}$$

Note that by assumption $QF > 1$, so $c_2^{2\star} > c_1^{1\star}$. In a perfect insurance contract, type-2 agents (agents with $\Theta = 2$) consume more than type-1 agents (unlucky agents with $\Theta = 1$).

Demand Deposit Contracts

The optimal insurance contract can be recaptured by a demand deposit contract provided by banks. Banks will accept deposits from agents in period $t = 0$. At period $t = 1$, some depositors will be of type 1, and will approach the bank to withdraw their deposits early. Their turnips have not matured, but the bank will rip up other agents' turnips to provide type-1 agents with more than just their unit turnip in return. In period $t = 2$, the remaining (type-2) depositors will get whatever is left over.

The bank has no way of telling which agents are type 1 and which are type 2, so it structures contracts like this: Deposits placed at $t = 0$ will earn an interest rate of r_1 if withdrawn in period $t = 1$ and r_2 if withdrawn in period $t = 2$. Of course, since $r_1 > 0$, it is technologically impossible for the bank to pay off all agents the amount $1 + r_1$ in period $t = 1$, since at that time no turnips have actually matured. However, the bank has a technical legal liability to pay off $1 + r_1$ to any depositor who appears at its door in period $t = 1$.

The banking system as a whole faces a *sequential service constraint.* This constraint is fundamental to the operation of banks in this model. It requires that depositors be honored in the order in which they show up at the bank. Even though the bank can look out the window and see a line that clearly exceeds its capacity to pay, it must pay out $1 + r_1$ on a first-come, first-served basis.

These demand deposit contracts will have two equilibria. The first equilibrium will be the "good" equilibrium and will not feature bank runs. The second will be the "bad" equilibrium and will feature a bank run.

Begin with the first equilibrium. Agents of type 1 (and only those agents) appear at the bank in period $t = 1$ requesting their deposits plus interest, withdrawing $1 + r_1$ each. In period $t = 2$, the remaining agents split what is left (remember that all the turnips left in the ground from period $t = 1$ to $t = 2$ will have grown by a factor of F). Thus c_1^1 and c_2^2 are related by:

$$c_1^1 = 1 + r_1, \text{ and:}$$
$$c_2^2 = F\frac{1 - \theta(1 + r_1)}{1 - \theta}.$$

Notice that although the bank announces the interest rates $\{r_1, r_2\}$, they are not independent. Choosing a value of r_1 automatically fixes r_2. For the rest of this section we will not calculate r_2 explicitly.

For this equilibrium to work, type-2 agents must not prefer the contract offered to type-1 agents. Since, for type-2 agents, consumption in periods $t = 1$ and $t = 2$ are perfect substitutes, it must be the case that, for the banking equilibrium to work:

$$c_2^2 \geq c_1^1.$$

Substituting in from above, this translates to:

$$F\frac{1-\theta(1+r_1)}{1-\theta} \geq 1+r_1.$$

We can rearrange this find:

$$\theta(1+r_1) + \frac{(1-\theta)(1+r_1)}{F} \leq 1. \tag{17.8}$$

Notice that this is exactly the same as equation (17.5) above, the budget constraint on the optimal insurance contract.

How do banks set r_1? They maximize the expected utility of their borrowers:

$$\theta \ln(1+r_1) + (1-\theta)Q \ln\left(F\frac{1-\theta(1+r_1)}{1-\theta}\right). \tag{17.9}$$

But this is exactly the same objective function as we used above, in the optimal insurance contract. Thus $1+r_1$ will satisfy:

$$1+r_1 = c_1^{1\star}.$$

Here $c_1^{1\star}$ is from equation (17.6) above. Thus demand deposits can replicate perfectly the optimal insurance contract.

If all type-2 agents stay at home in period $t = 1$, everything works perfectly. Type-2 agents will only be willing *not* to go to the bank in period $t = 1$ if they are certain that no other type-2 agent will be withdrawing in period $t = 1$.

This is the key to the bad equilibrium in this model, the bank run. If type-2 agents suspect that other type-2 agents are withdrawing from the bank in period $t = 1$, their consumption in period $t = 2$ will diminish. If enough type-2 agents attempt to withdraw in period $t = 1$, there will be nothing left in period $t = 2$. Thus if a type-2 agent believes that other type-2 agents are going to the bank to withdraw in period $t = 1$, their optimal response is also to withdraw in period $t = 1$. In a bank run equilibrium, the entire population appears at the bank in period $t = 1$ demanding $1+r_1$. Since $1+r_1 > 1$, and there is only one unit of the consumption good present in the bank, the bank liquidates its entire stock of consumption good to satisfy the first $1/(1+r_1)$ proportion of agents in line. All other agents get nothing in $t = 1$ and, of course, nothing in $t = 0$.

Notice that the bank run has a real economic cost: by liquidating the turnip crop in period $t = 1$, none is left to grow in period $t = 2$ and total economy-wide output goes down. Moreover, in a bank run, some agents lose the entire value of their endowment. Although Diamond and Dybvig don't model it, we would expect this outcome to lead to social unrest.

Suspension of Convertibility

During the financial panics of the of the Great Depression, banks would often close their doors when faced with a bank run. Remember that in those days deposits were entirely uninsured, so depositors were desperate to realize any part of their deposits. As more and more banks refused to honor their deposits, the Federal government declared several *bank holidays,* during which no banks (solvent or insolvent) could open their doors. The banks were, in effect, suspending the ability of their depositors to convert their deposits to cash.

Diamond and Dybvig's model provides us a way to think about how suspension of convertibility works. It turns out to be an effective deterrent against bank runs only if θ is known in advance.

The complete derivation of this result is beyond the scope of this section, but we can sketch it out here. Imagine that θ is known with certainty. The bank announces that only the first θ depositors in line in period $t = 1$ will be served. A type-2 agent faces no penalty for staying at home in period $t = 1$ even if other type-2 agents are going to the bank. He is secure that there will be no excessive liquidation, and that his deposits will mature as expected in the next period. Indeed, it is to his benefit to have other type-2 agents withdraw early, in period $t = 1$, since the total number of withdrawals is capped at θ, the more type-2 agents who withdraw early, the fewer type-2 agents will be left in period $t = 2$ to share the value of the remaining deposits.

What if θ is not known with certainty? The first thing to establish that, in principal, nothing is different. Imagine that are two possible values of θ: high, with $\theta = \theta_1$ and low, with $\theta = \theta_0$. Say that the high-θ outcome occurs with probability ζ.[3] Then the expected utility of an agent who consumes c^1 if type 1 and c^2 if type 2, is:

$$\zeta\left[\theta_1 u(c^1) + (1-\theta_1)Qu(c^2)\right] + (1-\zeta)\left[\theta_0 u(c^1) + (1-\theta_0)Qu(c^2)\right].$$

This can be rearranged as:

$$[\zeta\theta_1 + (1-\zeta)\theta_0]u(c^1) + [\zeta(1-\theta_1) + (1-\zeta)(1-\theta_0)]Qu(c^2).$$

Define $\bar{\theta}$ to be the expected value of θ:

$$\bar{\theta} = \zeta\theta_1 + (1-\zeta)\theta_0.$$

The expected utility may be rewritten using $\bar{\theta}$ as:

$$\bar{\theta}u(c^1) + (1-\bar{\theta})Qu(c^2).$$

When forming expectations, agents use the expected value of θ.

Imagine that θ_1 is quite high, approaching one. The bank cannot suspend convertibility at any proportion below θ_1, because it cannot know the true value of θ. In fact, no one in the

[3] The Greek letter ζ is called "zeta".

economy knows the true value of θ. Imagine the plight of a type-2 agent watching agents in line in front of the bank. Is this a bank run? Are there type-2 agents in that line? Or is it simply the case that the high-θ outcome has been realized? If there are type-2 agents in that line, the optimal response for the type-2 agent is also to get in line, since there is a probability that in fact the low-θ outcome has been realized, and real economic damage is being done. This story bears a striking resemblance to what actually happened during the financial panics of the Great Depression: in the midst of confusion about the true state of liquidity demand, banks kept their doors open, forcing other agents to run to the bank.

Deposit Insurance

Deposit insurance will completely cure bank runs, even if θ is not known. In this model, deposit insurance is more than a promise by the government to honor all deposits. Since the stock of turnips is limited, the government must also tax agents to honor deposits. Deposit insurance works like this: in period $t = 1$ a certain number of agents apply to withdraw their deposits and realize $1 + r_1$. If the banks can honor these deposits and still invest enough between $t = 1$ and $t = 2$ to honor the remaining deposits, the government does nothing. If there is excess demand for withdrawals, the government begins taxing depositors to honor all the demand deposits in $t = 1$ and to ensure that deposits are honored in period $t = 2$. Agents (of both types) who withdraw their deposits in $t = 1$ will, if there is a bank run, consume less than $1 + r_1$, because of the taxes used to finance the deposit insurance.

From the point of view of a type-2 agent, even if other type-2 agents are running to withdraw in period $t = 1$, he is assured that there will be enough invested to honor his deposit in period $t = 2$. Thus there is no benefit to joining in the run. Indeed, because of the excess demand for withdrawals in period $t = 1$ precipitated by a bank run, all agents (type 1 and type 2) who rush to cash out their deposits in period $t = 1$ will realize less than the $1 + r_1$ they are owed because they are taxed by the government.

Mutual Funds

The multiple equilibria in this model of banking depend critically on the presence of the sequential service constraint. By relaxing this constraint, we can overcome the bad equilibrium.

A sequential service constraint is an integral part of a banking system with fixed-obligation deposit contracts. That is, if a bank is going to promise $1 + r_1$ to anyone who walks through the door in period $t = 1$, it is bound to serve its customers sequentially. Doing away with the sequential service constraint means doing away with banking entirely.

As an alternate system, consider a mutual fund. This is exactly the kind of institution that would replace banks in a narrow-banking system. All agents trade their turnips in period $t = 0$ for a single share in the mutual fund. In period $t = 1$ there will be a market for shares in the mutual fund: agents will be able to cash them out at some price p_1 for the consumption good. If all agents decide to cash out their shares, this price will be unity. In period $t = 2$, the remaining shareholders will split the remaining assets of the mutual fund. If some proportion α of the population wish to trade in their shares at some price p_1, the remaining proportion of population will consume p_2 in period $t = 2$, where p_2 is given by the familiar equation:

$$p_2 = F\frac{1 - \alpha p_1}{1 - \alpha}.$$

It must be the case that $p_2 \geq p_1$ or no agents (not even type-2 agents) will be willing to hold on to the mutual fund until period $t = 2$. This can be rewritten as:

$$\alpha p_1 + \frac{(1 - \alpha)p_1}{F} \leq 1.$$

The competitive equilibrium in mutual fund shares will have the highest possible value for p_1, but p_2 will still be greater than p_1. As a result, only type-1 agents will sell out in period $t = 1$ and all type-2 agents will wait until period $t = 2$ to consume. This arrangement is not susceptible to runs. Imagine a type-2 agent in period $t = 1$ when other type-2 agents are "running" (in this case, selling out early). Since there must always be a competitive equilibrium, p_1 falls, and p_2 is always greater than p_1. As a result, our type-2 agent sees no benefit in joining the run, waits until period $t = 2$ and consumes $p_2 \geq p_1$. The key is that the sequential service constraint has been replaced by a competitive market in shares. Unusually high demand for consumption in period $t = 1$ is met by an unusually low price for shares in that period, p_1. In all cases, $p_2 \geq p_1$.

Variable	Definition
H_i	Cash loan or transfer to household i.
M', M	The new stock of money, the old stock of money.

Table 17.1: Notation for Chapter 17.

Exercises

Exercise 17.1 (Easy)
In the model of bank runs, explicitly calculate the interest rate on deposits held until period $t = 2$, r_2, when the interest rate on deposits held until period $t = 1$ is r_1.

Exercise 17.2 (Moderate)
For this problem, assume that there are only two types of potential borrowers: Safe (who comprise α of the population) and Risky (who comprise the remaining $1 - \alpha$ of the population). Banks cannot tell the difference between them, and with probability α, a borrower is safe and probability $1 - \alpha$ a borrower is risky. Safe borrowers have access to safe projects, which pay off π_S if they succeed and 0 if they fail. Safe projects succeed with probability p_S. Risky borrowers have access to risky projects, which pay off π_R if they succeed and zero if they fail. Risky projects succeed with probability p_R.

Risky and safe projects have the same expected payoff:

$$p_S \pi_S = p_R \pi_R,$$

but the probability of success is lower for risky projects, so $p_R < p_S$, and the payoff from succeeding is greater, so $\pi_R > \pi_S$. Both risky and safe projects have public failure, that is, there is no need to audit agents who claim that their project failed.

To finance the projects borrowers need a unit of capital from a bank. The bank in turn announces a repayment amount x in the event that the borrower's project does not fail. If the project fails, borrowers owe nothing (they declare bankruptcy). If the project succeeds, borrowers consume their output minus x, if the project fails, borrowers consume zero. Assume that borrowers are *risk neutral* so that their utility function is just their expected consumption.

There is a risk-free interest rate of r that banks must pay to their depositors (thus they have to realize at least $1 + r$ in expected value on their loan to meet their deposit liability).

1. Write down a bank's balance sheet (in terms of x, r, p_S, and p_R) assuming that, with probability α the borrower is safe and with probability $1 - \alpha$ the borrower is risky.
2. Assume that banks compete by offering the lowest value of x that gives them non-negative profits in expectation. Determine the equilibrium interest rate $x^\star(r, \alpha)$ as a function of the interest rate r and the proportion of safe agents α.
3. Find the expected utility of a safe agent who borrows, $V_S(r)$, as a function of the interest rate r when x is given by $x^\star(r, \alpha)$. Repeat for a risky agent.
4. Agents stop borrowing if the expected utility of being a borrower falls below zero. Show that if a safe agent decides to borrow, a risky agents will too. Find the critical interest rate $r^\star$ at which safe agents stop borrowing. At interest rates greater than or equal to this critical value, $r \geq r^\star$ all safe agents leave the pool, so $\alpha = 0$. What happens to the equilibrium payment x?

Exercise 17.3 (Moderate)
Consider the model of costly audits again. Now suppose that intermediaries gain access to a technology which allows them to extract more from each borrower (that is, for each value of announced repayment x and audit cost γ, suppose $\pi(x, \gamma)$ shifts up). What happens to

the demand schedule of capital? What happens to the supply schedule of capital? What happens to the equilibrium interest rate? What happens to equilibrium economy-wide output? Are agents made better off or worse off?

Exercise 17.4 (Moderate)
Yale University costs 1 dollars to attend. After graduation, Yalies (that is, graduates of Yale) either land good jobs paying w or no job at all, paying nothing. The probability of landing the good job is π where π is hidden effort exerted by the Yalie. Yalies are born with wealth $a \geq 0$, and those Yalies born with wealth $a < 1$ must have a loan of $1 - a$ to attend. Yale University will act as a lender to those students. Yale must borrow at the risk-free gross interest rate $r > 1$ to finance the loans. Student borrowers who get the good job must repay Yale University some amount x out of their wages w. Student borrowers who do not land the good job pay nothing. All students have preferences over lifetime expected consumption $E(c)$ and private labor effort π of:

$$V(E(c), \pi) = E(c) - \frac{w}{\alpha}\frac{\pi^2}{2}.$$

Assume $0 < \alpha < 1$.

1. Start with a rich Yalie, with $a > 1$. Show that her optimal effort $\pi^\star$ is α.
2. Now consider poor Yalies, with $a < 1$, who must borrow to finance their education. Calculate a borrower's optimal effort $\pi(x)$ as a function of x.
3. Write down Yale University's expected profit on a loan to a student with wealth $a < 1$ as a function of x, assuming that Yale University knows $\pi(x)$ from Exercise (2).
4. Assume that Yale University operates a "fair lending policy" in which borrowers of wealth a must repay an amount $x(a) = r(1 - a)/\alpha$ if they get the good job. What is "fair" about this lending policy? Given this policy and your answer to Exercise (2) above, calculate a borrower's optimal effort as a function of their wealth. That is, write down $\pi[x(a)]$, and call it $\pi(a)$.
5. Show that, given Yale University's "fair lending policy", all Yalie borrowers exert less effort than rich Yalies, that is, for Yalies with wealth $0 \leq a < 1$, show that $\pi(a) < \pi^\star$ and that $\pi(a = 1) = \pi^\star$, where $\pi^\star$ is from Exercise (1) above.
6. Finally, show that given its fair lending policy, that Yale loses money on student loans, and that the loss is increasing in loan size. Why does the fair lending policy cost Yale money?

Variable	Definition
α	Proportion of population who are type-1 workers.
γ_i	The audit cost of agent i.
c_0	Consumption in the first period of life (type-1 agents only).
ℓ_0	Labor effort in the first period of life (type-1 agents only).
c_1	Consumption in the second and last period of life (both types of agent).
$U^1(c_0, \ell_0, c_1)$	Preferences of type-1 agents.
$U^2(c_1)$	Preferences of type-2 agents (risk neutral).
k	Capital input to type-2 agent's project, can take on only two values, $k = 0$ or $k = 1$.
y_i, y	Output of agent i's project, or just output.
ε	Shock to output, distributed uniformly on $[0, 1]$.
h	Maximum labor effort by a type-1 agent, $h < 1$.
$s, S(r)$	Savings of a type-1 agent, or of the representative type-1 agent (aggregate supply of capital).
r	Economy-wide equilibrium interest rate on capital.
X, x	Repayment amount, $X = 1 + x$.
$\pi(x, \gamma, \varepsilon)$	Revenue of bank on a loan (gross of the borrowing cost r) to an agent with audit cost γ, when the repayment amount is x and the production shock is ε.
$\pi(x, \gamma)$	Expected revenue (gross of the borrowing cost r) on a loan to an agent with audit cost γ, when the repayment amount is x. Expectation taken over the production shock ε.
$\pi^\star(\gamma)$	The highest possible expected revenue (gross of the borrowing cost r) on a loan to an agent with an audit cost of γ.
$x^\star(\gamma)$	The repayment amount that results in the highest revenue to the bank on a loan to an agent of audit cost γ.
$\gamma^\star(r)$	Largest value of the audit cost γ at which the bank can make enough revenues to cover the cost of borrowing, r.
$K^d(r)$	Aggregate demand for capital.

Table 17.2: Notation for the model of audit costs in Section 17.2

Variable	Definition
k	Capital input to productive technology, can take only two variables, $k = 0$ or $k = 1$.
a	Wealth of agent.
ℓ	Private labor effort of agent.
c	Consumption of agent.
q	High output of technology (the low output is zero).
α	Disutility of effort in agent's preferences.
r	Economy-wide risk-free rate on capital.
X	Repayment amount.
$a^\star(r)$	Threshold credit rationing wealth.

Table 17.3: Notation for model with moral hazard in Section 17.3

Variable	Definition
F	Technology parameter: growth of asset between periods $t = 1$ and $t = 2$.
$U(c_1, c_2; \Theta)$	Utility function over consumption in period $t = 1$, $t = 2$ and shock term Θ.
Θ, θ	Shock term: $\Theta = 1$ means that the agent is urgent to consume (probability θ).
Q	Preference parameter: marginal utility of type-2 agents, $Q < 1$.
r_1, r_2	Interest rate promised by the bank on deposits held until period $t = 1$ (r_1) or period $t = 2$ (r_2).
c_j^i	Consumption by agent to type i in period j.
$c_1^{1\star}, c_2^{2\star}$	Optimal consumption in period $t = 1$ by type-1 agents and the optimal consumption in period $t = 2$ by type-2 agents.

Table 17.4: Notation for model of bank runs in Section 17.4