Economics 202A
Lecture Outline #13 (version 2.0)

Maurice Obstfeld

Financial Instability

In this lecture I describe the Diamond-Dybvig model of bank runs, from *Journal of Political Economy*, June 1983. In this model, bank-like financial intermediaries promote risk sharing among individuals, but they are subject to arbitrary panics.

The model

There are three periods, $T = 0, 1, 2$.

There are two possible technologies on date 0, short and long.

Investment of 1 unit of output in the short technology at $T = 0$ yields 1 unit of output in period 1 and 0 in period 2.

Investment of 1 unit of output in the long technology at $T = 0$ yields 0 units of output in period 1 and $R > 1$ units in period 2.

Individuals need not specify the technology they are choosing *ex ante*. They opt for the short or long technology simply by "harvesting" the yield either on date 1 or 2, respectively.

The idea is that more roundabout technologies are more productive.

At time 0, a depositor does not know his/her "type," patient or impatient. Depositors are indexed by the unit interval, $[0, 1]$. at the start of period 1, a fraction p is revealed to be of type 1, or impatient. The rest (of measure $1 - p$) are of type 2, patient. An agent has an endowment 1 in period 0 and consumes in period 1 and/or 2. The utility functions of types 1 and 2 are

$$\begin{aligned} U(c_1, c_2; 1) &= u(c_1), \\ U(c_1, c_2; 2) &= u(c_1 + c_2), \end{aligned}$$

where $\lim_{c \to 0} u'(c) = \infty$, $\lim_{c \to \infty} u'(c) = 0$, and $-cu''(c)/u'(c) > 1$.

First consider an autarkic individual. That person will pick $c_1 = 1$ if he/she turns out to be impatient, $c_2 = R$ if patient. That person's *ex ante* expected utility is an average over the utilities of the two types:

$$\mathrm{E}U = pu(1) + (1-p)u(R).$$

People can do better than this, however, if there are financial intermediaries.

Social optimum

A benevolent and omnipotent planner would withdraw an amount $1 - x$ from investment on $T = 1$ so as to maximize the expected utility of a representative individual

$$pu\left(c_1^1\right) + (1-p)u(c_1^2 + c_2^2)$$

subject to the aggregate resource constraints

$$\begin{aligned} pc_1^1 + (1-p)c_1^2 &= 1 - x, \\ (1-p)c_2^2 &= Rx. \end{aligned}$$

Here, c_j^i is the amount type i consumes in period j. Of course, it is always optimal that $c_1^2 = 0$.

So we are left with the simpler problem:

$$\max_{c_1^1, c_2^2} \; pu\left(c_1^1\right) + (1-p)u(c_2^2)$$

subject to

$$pc_1^1 + (1-p)\frac{c_2^2}{R} = 1.$$

If λ is the Lagrange multiplier on the resource constraint, the first-order conditions for a maximum are

$$\left.\begin{array}{r} u'\left(c_1^1\right) = \lambda \\ u'\left(c_2^2\right) = \lambda/R \end{array}\right\} \implies$$

$$u'\left(c_1^1\right)/u'\left(c_2^2\right) = R.$$

This social optimum implies that an impatient person gets to consume more than $c_1^1 = 1$, the autarky value. Why? The budget constraint of the planner is

$$c_2^2 = \frac{R}{1-p} - \frac{pR}{1-p} c_1^1.$$

At the autarky allocation, however, because relative risk aversion exceeds 1, the absolute-value slope of the social indifference curve satisfies

$$\frac{pu'(1)}{(1-p)u'(R)} > \frac{pR}{1-p},$$

which means that it exceeds the absolute-value slope of the planner's budget line. (Please refer to the diagram on the next page so that you can visualize this.) For example, if $u(c) = c^{1-\rho}/(1-\rho)$, this condition is $u'(1) > Ru'(R)$, or $1 > R^{1-\rho}$, which holds for $\rho > 1$ (because $R > 1$). In this case of high risk aversion, the social optimum "insures" agents against being impatient and ending up with relatively low consumption. I denote the social optimum consumption levels by c_1^{1*} and c_2^{2*}. Observe that c_1^{1*} must be stricly less than c_2^{2*} (as is also indicated in the diagram).[1]

Banks and bank runs

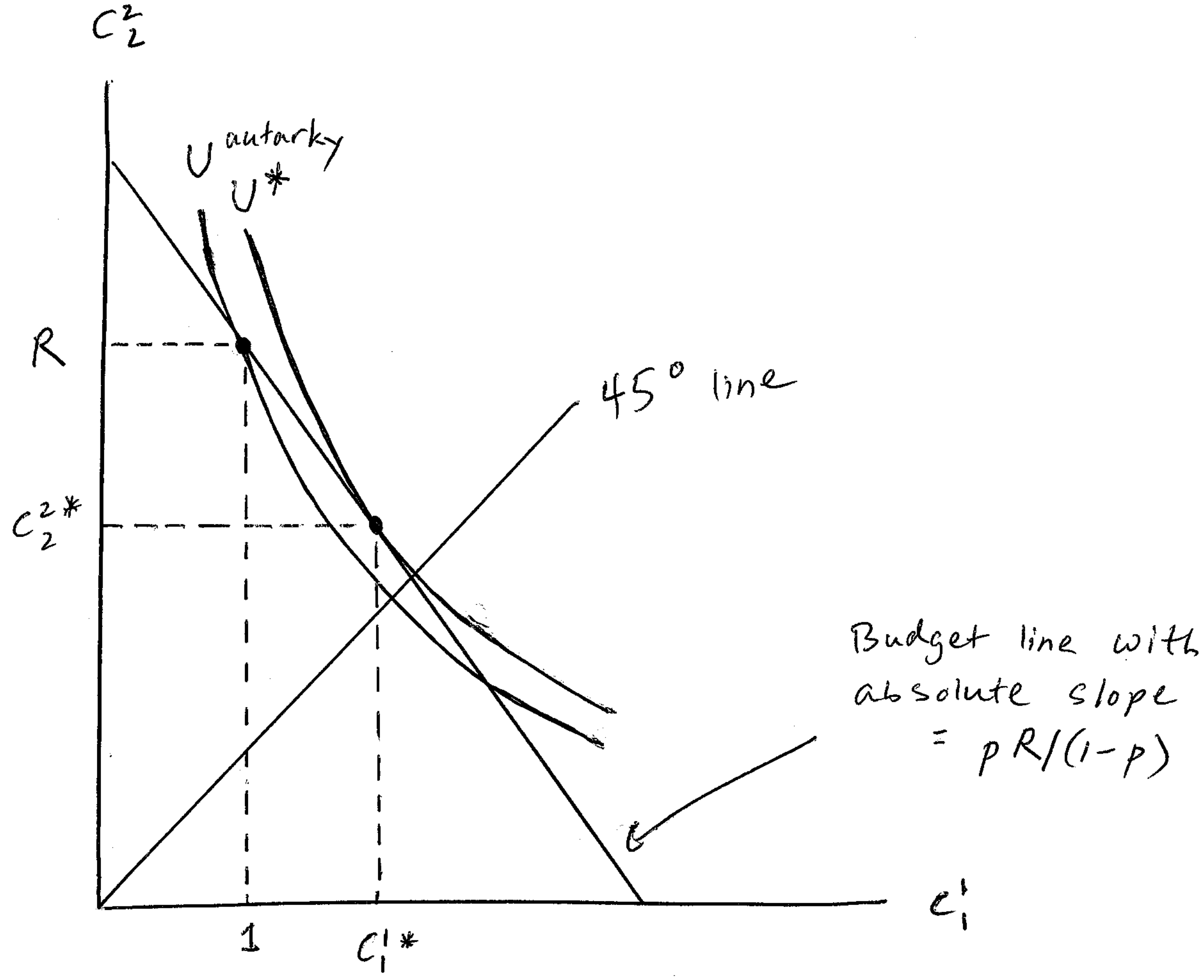
To make the model interesting, assume that an individual's type and consumption cannot be verified. Imagine there were contracts that would insure people upon learning they were impatient. The payments would have to come from patient types liquidating part of their investment.

Such contracts would never work. You would have an incentive to pretend to be impatient, reaping an insurance payment, say x, that you could

[1] The slope of an indifference curve U at the consumption pair $\left(c_1^1, c_2^2\right)$ is

$$\left.\frac{\mathrm{d}c_2^2}{\mathrm{d}c_1^1}\right|_U = -\frac{pu'\left(c_1^1\right)}{(1-p)u'(c_2^2)}.$$

This implies that where $c_1^1 = c_2^2$, the absolute-value slope of any indifference curve is $p/(1-p)$. Because $R > 1$, that slope is strictly below the absolute-value slope of the planner's budget line, $pR/(1-p)$. As a result, c_1^{1*} is stricly less than c_2^{2*}



Autarky allocation and social optimum in the Diamond-Dybvig model consume in period 1 (making $c_1^2 = x$). Then you could leave your investment in place and still consume $c_2^2 = R$ in period 2.

So consider instead a bank contract. Everyone deposits their resources in the bank at time 0. Patient types can withdraw $r_1 > 1$ in period 1 — with their withdrawals monitored by the bank. Patient depositors get their pro rata share of what is left after period 1 withdrawals.

Banks have the potential to implement the optimum. If $r_1 = c_1^{1*}$ and a fraction p of the population (the impatient) withdraws deposits on date 1, then each of the patient consumes his or her pro rata share of the balance, $R(1 - pc_1^{1*})/(1 - p) = c_2^{2*}$. Because $c_2^{2*} > c_1^{1*}$, as observed above, no patient depositor has an incentive to withdraw early. So this setup clearly yields an equilibrium for periods 1 and 2. Furthermore, if on at $T = 0$ agents expect this equilibrium to prevail with probability 1, each of them, knowing that the expected utility from signing the contract exceeds the autarky level, will indeed sign and deposit his or her resources in the bank.

Things can go wrong however, because the preceding equilibrium for periods 1 and 2 is *not* the only one. To capture the reality of banking, the model assumes a *sequential service constraint*: essentially, this means that the bank services customers' claims, in the order in which they arrive, until its resources run out. Let V_1 be the payoff you get (depending on your place in line) is you withdraw in period 1, and V_2 the payoff you get in period 2 if you do not withdraw in period 1. If f_j denotes the number of depositors serviced before depositor j on date 1, and f is the total number of withdrawals on date 1, then

$$V_1(f_j, r_1) = \begin{cases} r_1 \text{ if } f_j r_1 < 1 \\ 0 \text{ if } f_j r_1 \geq 1 \end{cases}$$

and

$$V_2(f, r_1) = \max\left\{\frac{R(1 - r_1 f)}{1 - f}, 0\right\}.$$

In the first-best equilibrium, $f = p$ and so

$$V_2(f, r_1) = V_2(f, c_1^{1*}) = \frac{R(1 - pc_1^{1*})}{1 - p} = c_2^{2*}.$$

Alas, if r_1 were equal to 1, then we would have $V_2(f, 1) = \max\{R, 0\} = R$, and patient types would never have an incentive to withdraw in period 1. But then, banks would be no better than autarky. To do better, we need $r_1 > 1$, and in that case, there can be a depositor panic — a run on the bank.

For example, suppose you turn out to be patient but think that f will be $1/r_1$. In that case, you expect depositors to withdraw all the bank's resources at $T = 1$, making $V_2\left(\frac{1}{r_1}, r_1\right) = 0$. So it is individually rational for you to join the queue of depositors in front of the bank as quickly as you can, in the hope of getting your money out. Of course, everyone will do the same in this equilibrium, and some depositors will leave empty-handed. No one will get money back at $T = 2$. The bank will have failed.

So the first-best bank equilibrium looks inherently fragile. It depends on the confidence of depositors that the bank will not fail — a self-fulfilling prophecy, however it goes.

I note that in the event of a run, the bank is suffering from a pure *liquidity crisis*. Because all depositors want their money right away at $T = 1$, the bank is forced forgo the higher returns on long-term investments with which it would otherwise be able to repay patient depositors later on. It is the bank run itself – and no other factor – that causes the bank to fail. It can also happen that the bank simply makes unwise risky investments, and when these fail, it lacks the funds to make good on deposit obligations. In that case we would call the bank *insolvent* rather than illiquid. The distinction between illiquid and insolvent institutions is incredibly important for policy decisions, yet unfortunately, the distinction is much clearer in specific models than it ever is in the real world. One reason for this ambiguity is that lenders seldom panic, thereby making a bank illiquid, unless there is some *chance* of insolvency. In a *systemic financial crisis* such as the 2007-09 crisis, where many institutions simultaneouly face runs, the distinction becomes blurrier still. In that situation, market prices will be plummeting as banks and other short-term borrowers *all* try to sell illiquid assets to meet creditor demands for cash.

Back to the model. What happens at $T = 0$ if there is a nonzero chance of a bank run at $T = 1$? It depends. In this model, the possibility of a bank run on date 1 is competely exogenous – for example, it could be determined by the emergence of a random "sunspot" at the start of period 1. If the arbitrary probability π of the sunspot is big enough, and if the probability that a depositor gets to the bank too late is also high enough, no one would wish to sign bank contracts in the first place. They would be deterred by an excessive probability of ending up with zero consumption (in the event of a bank failure). It could be that certain contracts with $1 < r_1 < c_1^{1*}$ would be signed – because for these, the probability that the bank runs out of funds before you get there is smaller. However, such contracts also yield an expected utility level below the social optimum.

Remedies

The topic of financial instability is a critically important one — but it has returned to prominence as a result of events starting in August 2007. How can this threat be addressed through policy?

Deposit insurance. In the U.S. at the moment, the FDIC insures all deposits up to balance of $250,000 – scheduled to return to its previous value of $100,000 in 2014. So small depositors should have no incentive to run the bank. However, banks that lend to other banks in the interbank market are not insured — they can face default if the borrowing bank closes its doors. That is what happened in the 2007-09 crisis — the interbank market was beset by fears about other banks' solvency.

In Britain deposit insurance was minimal in the autumn of 2007 when there was a depositor run on Northern Rock bank, a big mortgage lender. This was the first British bank run since 1866, and it was a huge embarrassment for Her Majesty's government. The government stopped the run by promising to insure all deposits at all banks. (A great article on this episode is by Hyun Song Shin, "Reflections on Northern Rock: The Bank Run that Heralded the Global Financial Crisis," *Journal of Economic Perspectives*, Winter 2009.)

Essentially, as Diamond and Dybvig recount, deposit insurance works by having the government promise to levy taxes to repay depositors. This is just what the British government did, but after the fact. (Even that solution runs into problems when the bank is so big that its liabilties exceed GDP! Union Bank of Switzerland is a case in point.)

Lender of last resort (LLR). Since the central bank prints money, it can easily support any bank needing liquidity by providing cash. The Bank of England did this in the case of Northern Rock; it was the news of the Bank's LLR support that set off the run by small depositors! Some argue that ultimately, the government's fiscal powers must back up any banking guarantees. The fact that Northern Rock ended up owing the Bank of England large sums of money suggests the centrality of the government's fiscal powers to guarantees of financial stability. Sometimes, banks that seem merely illiquid at first glance may turn out to be insolvent, in which case the government is likely to take them over to protect depositors. The idea of an LLR originated late in the eighteenth century ("*le dernier resort*," as the Bank of England was called at one point), and was elaborated by writers such as Henry Thornton and Walter Bagehot. (Thornton, who incidentally anticipated much of modern monetary theory in 1802, also was instrumental in Britain's outlawing of the slave trade. He was a cousin of the abolitionist leader William Wilberforce and is portrayed in the recent film *Amazing Grace*).

Moral hazard. If everyone knows the government is standing ready to save the banks, the banks will take excessive risks and depositors will fail to monitor bank practices. This is the moral hazard problem. It can also lead to problems in monetary policy if central banks cut interest rates excessively to favor distressed financial institutions. The problem is similar to the other problems of dynamic inconsistency in monetary policy that we have already discussed — and the leading central banks find themselves in that situation now. One fix is to impose much stricter prudential supervision of banks as well as other financial institutions that might need public funds. Debate over enhanced safeguards is going on now in several countries including the United States. A full discussion of such regulatory issues would require a course in itself.