BARNARD COLLEGE, COLUMBIA UNIVERSITY

The Economics of Money and Banking

ECON V3265

Professor Perry Mehrling

Fall 2016
MW 6:10pm - 7:25pm
202 ALTSCHUL HALL

Introduction to the principles of money and banking. The intermediary institutions of the American economy and their historical developments, current issues in monetary and financial reform.

Contents

Section 1

1. The Four Prices of Money 2
2. Natural Hierarchy of Money 7
3. Money and the State, the US Case 15
4. The Money View, Micro and Macro 22
5. The Central Bank as a Clearinghouse 32
6. Federal Funds: Final Settlement 37
7. Repos: Postponing Settlement 45
8. Eurodollars: Parallel Settlement 51
9. The World that Bagehot Knew 59
10. Dealers: Liquid Security Markets 67
11. Banks: The Market for Liquidity 74
12. Lender/Dealer of Last Resort 81

Section 2

13. Chartalism, Metallism, and Key Currencies 92
14. Money and the State, International 101
15. Banks and Global Liquidity 107
16. Foreign Exchange 113
17. Direct and Indirect Finance 119
18. Forwards and Futures 128
19. Interest Rate Swaps 136
20. Credit Derivatives 142
21. Shadow Banking, Central Banking, and the Future of Global Finance 149
22. Touching the Elephant: Three Views 162

1. The Four Prices of Money

Welcome

Syllabus handout, release handout

Texts:

Lecture Notes on Courseworks, soon after each lecture
Stigum
--a practitioner's view of the money markets, recently updated
Reading Packet on Courseworks, already there
--wide range of perspectives, views
Frydman or Barofsky
--reading period assignment, opportunity to integrate, connect
Financial Times
http://ineteconomics.org/blog/money-view

Big Picture:

This course builds on **my own background** in the history of monetary economics and financial economics, as developed in my first two books, The Money Interest and the Public Interest, American Monetary Thought 1920-1970 and Fischer Black and the Revolutionary Idea of Finance. I locate myself in the distinctly American tradition of monetary thought that stretches from AllynYoung to Alvin Hansen, Ed Shaw, and Hyman Minsky, all of whom engage in one way or another with the great tradition of British central banking thought that stretches from Walter Bagehot to Ralph Hawtrey, Richard Sidney Sayers, and Charles Goodhart. I have included the writings of some of them in the reading list to introduce you to these traditions.

But this course is not about that past development. It is quite emphatically a course about the present, and indeed about the emerging future. It is a course about money in a world of **financial globalization**. The history of monetary and financial thought is a story about how each generation had to rethink the underlying issues for themselves in order to make sense of the conditions of their own time. So do we today. Our task is to work out the implications of financial globalization for how we understand money, banking, and central banking.

To do that we start not so much with the tools of economics but rather with the testimony of practical bankers themselves, whose business requires them to work out the implications of a changing world for their own practice. One thing that my historical studies have taught me is that the origins of monetary thought lie in **concrete banking practice**. There have always been some practitioners who respond to an inner urge to write down what they know in the form of a manual to teach others. Today these people can be found at central banks, the BIS and IMF, and also inside the global banking system that comprises the infrastructure of financial globalization. We will be reading and taking seriously these voices, but always with an idea to extract from them some nuggets of wisdom we can translate into economics.

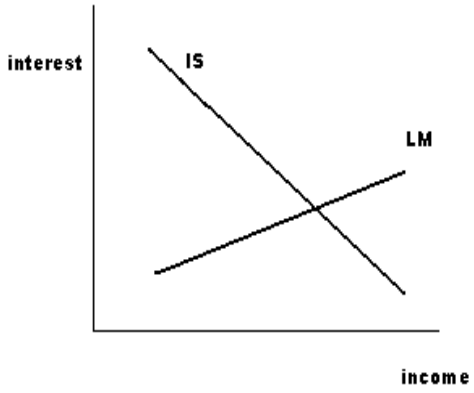
Teaching Assistants will be helping us with that translation

Discussion sections: weekly, on the reading for the week
Problem Sets: balance sheets, repo math, dealer economics, international money, derivatives
Exams--dates, style (analytical distinctions, T/F, identifications)

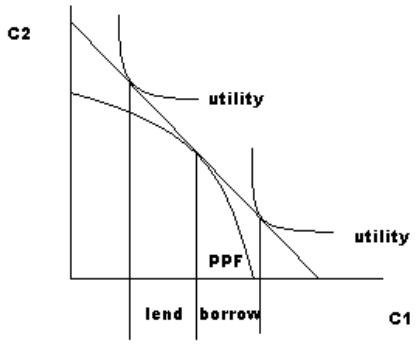
Prerequisites and Math

Intermediate Macro and Micro required for economics majors. What does that mean?

Macro means IS/LM—macro story about nominal interest rate
--the Fed, Taylor Rule, DSGE



Micro means Fisher diagram—micro story about real interest rate, slope of budget line=-1/(1+r)



It is possible to teach money and banking as a course in applied macro, or applied micro. Applied economic theory emphasis:

--micro, banks as firms, intermediation because of asymmetric info
--macro, foundations of the LM curve, money demand and supply
--monetary policy, credit channel and money channel

We will not however be making much explicit use of these analytical constructs until the end of the course. As I say, we will be listening to bankers, trying to build up our own sense of the system inductively.

The analytical backbone of the course is an accounting model, both micro and macro. For now it is sufficient to start with micro, and thinking of every agent in the economy as a kind of bank, in the sense that every agent needs to be attending to the daily inflow and outflow of cash balances. Every agent has to worry about this, but only actual banks make a business out of worrying about it. What is a bank?

Bank

Assets	Liabilities
Loans Securities Cash Reserves	Deposit Accounts Other borrowing Net Worth

Issue of Solvency vs. liquidity, issue of monetary control

"Shadow Bank"

Assets	Liabilities
Mortgage Backed Security Tranche Interest Rate Swap Credit Default Swap	Money Market Borrowing Repurchase agreement Asset Backed Commercial Paper Auction Rate Securities Eurodollar Borrowing

Issue of solvency vs. liquidity, issue of monetary control

Central Bank is also a Bank, at a higher level:
http://www.federalreserve.gov/releases/h41/Current/h41.pdf

Issue of solvency vs. liquidity, issue of monetary control

Centrality of the issue of liquidity—that is what makes banks special and different from other corporations.

Syllabus overview

A course in banking phenomena, not banks as legally defined institutions.

Four **functions** of banking
Clearing, market making, advance clearing, intermediation

I believe that understanding the payments system is the best starting place for understanding the phenomenon of liquidity. Once we understand the phenomenon of liquid assets we are ready to consider the phenomenon of liquid markets. Thus clearing and market making are the two central functions that we will be concerned about before the midterm.

Four **prices** of money
Par—different types of money (currency, deposits)
Interest—future money (fed funds, Eurodollars, repo)
Exchange rate—foreign money
Price level—commodities

Financial Globalization Redux

Why don't I just teach the standard stuff?

The world has changed, and even the best textbooks lag behind. Existing theory was developed to explain a very different world: postwar (govt debt) and post Depression (govt guarantees, regulation). This led to emphasis on intermediation story, and monetary policy story. Today we live in a different world. Private financial markets, and globalization, are the key facts of our time. Practical central bankers know this better than academic economists, and traders in the money markets know this best of all.

In terms of **globalization** (trade and integration), the current world is more similar to the pre-WWI world than the post-WWII world, the world when the pound sterling was king and the Bank of England operated as central bank for the entire world. Longer history thus gives more perspective on our current situation than does more recent history. Bagehot helps us more than Irving Fisher.

In terms of **finance**, the current world is in many ways new. There are new financial instruments, much more wide usage of existing instruments, electronics and telecommunications, and new theories. This is not a course about finance, but it is deeply informed by finance practice, and also finance theory. The strong finance view (such as Fischer Black) says there should be no separate theory of money, that liquidity is a free good in efficient markets. Obviously this is a strong challenge to a course like this one, and I don't think the standard economics view offers a very adequate response. For me, the place where money meets finance is in a kind of financial derivative called a swap since all banking is essentially a swap of IOUs.

Next week:
Lectures 2 and 3: Natural hierarchy, Monetary history
Reading: Young, prof at Harvard, American institutionalist, chapters in an encyclopedia written in 1920s, new Fed, end of Sterling hegemony. Read it once over the weekend and then again after next week lectures and discussion section.

2. Natural Hierarchy of Money

Always and everywhere, monetary systems are hierarchical.

One way that economists have tried to get an analytical grip on this empirical fact is to distinguish *money* (means of final settlement) from *credit* (promise to pay money, means of delaying final settlement).[1] This is fine so far as it goes. But in one sense it doesn't go far enough because it posits only two layers of the hierarchy. And in another sense it goes too far because what counts as final settlement depends on what layer we are talking about. What looks like money at one level of the system looks like credit to the level above it.

I.

To see this point more clearly, think about the monetary system under a gold standard and think not about money and credit in the abstract but rather about the concrete financial instruments gold, currency, bank deposits, and securities.

A Simple Hierarchy

Money	Gold
↑	Currency
↑	Deposits
Credit	Securities

In such a world *gold* is the ultimate money because it is the ultimate international means of payment. National *currencies* are a form of credit in the sense that they are promises to pay gold. National currencies may be "*backed*" by gold, in the sense that the issuer of currency holds some gold on hand, but that doesn't mean that these currencies represent gold or are at the same hierarchical level as gold. When a currency is backed by gold reserves, it is still a promise to pay, just a more credible promise to pay because the presence of reserves makes it more likely that the issuer of currency can fulfill on the promise if called upon to do so.[2]

Farther down the hierarchy, bank *deposits* are promises to pay currency on demand, so they are twice removed promises to pay the ultimate money, and *securities* are promises to pay currency over some time horizon in the future, so they are even more attenuated promises to pay. The credibility of these promises is an issue here, just as in the case of national currencies, and here as well reserves of the various instruments that lie higher up in the hierarchy can help to enhance credibility.

---

[1] Just so, Ralph Hawtrey's Currency and Credit (1923).

[2] National currency is not in general a "cloakroom ticket" representing ownership of gold that is being held somewhere on behalf of the currency holder. This is true even in the extreme case of 100% reserves (or a currency board arrangement). Even in such an extreme case, there is still a promise to pay, a promise that can be broken.

In this hierarchy, where is the dividing line between money and credit? It is tempting to draw the line between currency (and everything above it) as money, and deposits (and everything below it) as credit. The source of this temptation is the institutional fact that currency is the final means of settlement for domestic payments. Just so, for a bank settling its accounts at the end of the day, currency or "high-powered money" is certainly the means of settlement.

But things look different farther down the hierarchy. For ordinary people like ourselves, bank deposits are the means of settlement. Hence we might be inclined to view deposits (and everything above them) as money, and securities as credit. This is more or less what most modern textbooks mean when they speak of the money supply, although even here there is some ambiguity which is reflected in the various definitions of money: M1, M2, M3 and so forth.

And things look different farther up the hierarchy as well. For a country settling its accounts at the end of the day, national currency is of limited value. What other countries want is their own currency, or the international means of settlement, which means gold in the case of a gold standard, perhaps SDRs (Special Drawing Rights at the IMF) in the modern case. (The US is an exception because of the international role of the dollar.)

The point to hold on to here is that what counts as money and what counts as credit depends on your point of view, which is to say it depends on where in the hierarchy you are standing. Are you thinking of the problem of international settlement, of bank settlement, of retail settlement, or what? Best therefore not to reify the concepts of money and credit, and to rest instead with the more general idea that the system is hierarchical in character.

That point established, I need to remind you that, even with four layers, the hierarchy we've been talking about is much simpler than that in the real world. In the real world we see many more layers, and finer gradations in the hierarchy. Just so, a minute ago I used currency and high-powered money as synonyms (four paragraphs up), but they are not actually the same thing. (High powered money includes not only currency but also deposits at the central bank.) I also treated the category of bank deposits as homogeneous but it is not--there are different kinds of deposits, and also some deposit-like things (MMMF accounts) that are not the liability of any bank at all. The category of securities is, if anything, even more heterogeneous, encompassing promises of various maturities, credit quality, and so forth. All this just reinforces the point that we want to avoid sterile debates about what is money and what is credit, and stand instead on the point that the system is hierarchical in character.

II.

So far we have been thinking about the hierarchy as a matter of the qualitative difference between various financial *instruments*. It is illuminating to shift now and see the hierarchy from the standpoint of the various financial *institutions* that issue the instruments.

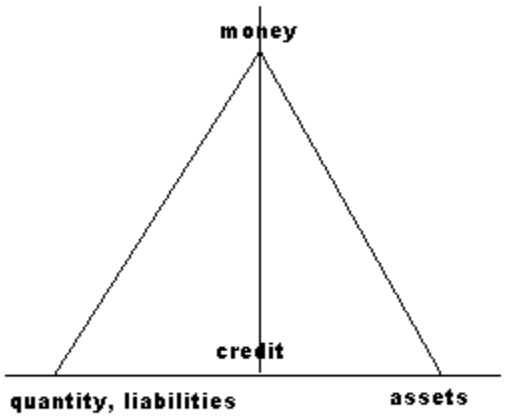
The Hierarchy in Balance Sheets

Central Bank		Banking System		Private Sector	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
Gold	Currency	Currency	Deposits	Deposits Securities	Securities

To keep things simple, I have noted only the instruments we have already been talking about, so there are important entries missing: government debt as the most important asset of the central bank, and loans as the most important asset of the banking system. For present purposes, the important point to appreciate is that all of the instruments except gold appear as both assets and liabilities. They are thus clearly all forms of credit. If we were to consolidate all three balance sheets in order to treat the economy as a single aggregate entity, all forms of credit would appear as both assets and liabilities, and hence cancel. Only gold would remain because only gold is an asset that is no one's liability.

More generally, the difference between gold and other forms of money is the difference between "outside" money and "inside" money, an analytical distinction first proposed by Gurley and Shaw in their seminal 1960 Money in a Theory of Finance. Actually, Gurley and Shaw treated currency as outside money and deposits as inside money because they aggregated only over the private economy, not including the government sector. So from their point of view currency as well as gold appears to be an asset that has no liability counterpart. In this course, by contrast, we will typically be thinking about the entire economy, even the entire world economy, so all financial assets will be inside, including currency. Once again, what counts as money and what counts as credit depends on your point of view. In this course we are taking a global view.

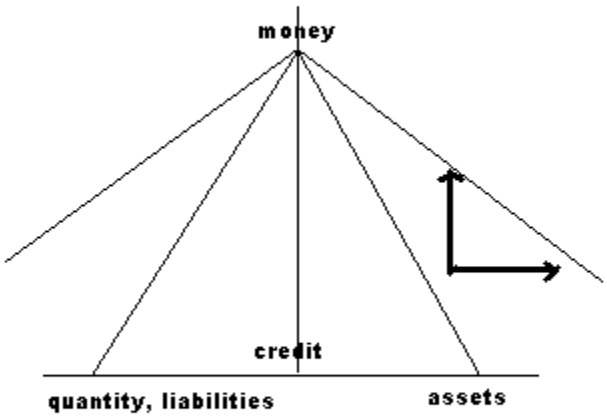
To consolidate the idea of an "inside" asset, it may help to visualize the hierarchy as a symmetric pyramid rising on a credit-to-money axis from a line centered on zero, so that net outstanding credit at any level is zero. I place the peak of the pyramid at zero even though there is a positive quantity of gold, simply to emphasize that that quantity is vanishingly small compared to the vast edifice of credit below. From the point of view of the system as a whole, every liability is someone else's asset. These credit forms cancel if we consolidate, but such consolidation misses the entire point. Macroeconomic variables like interest rates and GDP are affected not only by the outstanding gross quantity of inside credit, but also by who is issuing it, who is holding it, and where that credit lies in the larger money-credit hierarchy. (Standard macro models simplify by focusing entirely on some measure of the outstanding quantity of money, which they usually treat as an outside asset.)



III.

If we focus our attention on the hierarchy for any period of time, one thing becomes immediately clear, which is that the hierarchy is **dynamic**. At almost any time scale you care to examine, it is a system in motion. Focus your attention on daily clearing and settlement, on the business cycle frequency, or on the longer term secular scale, and you'll see constant flux: daylight overdrafts, credit cycles, wars and depressions. At every time scale, we see expansion and contraction of the hierarchy. As it expands, the hierarchy flattens and the qualitative difference between credit and money becomes attenuated, but then the system contracts and the hierarchy reasserts itself. At the business cycle frequency, the phenomena surrounding this contraction and reassertion are grouped under the headings "irrational exuberance" in the expansion phase and "financial crisis" in the contraction phase.

We can distinguish two dimensions of this fluctuation. First, and most simple, is the expansion and contraction of the *quantity* of credit, which takes place at all levels of the system. Second, and more subtle, is the fluctuation of the "moneyness" of any given type of credit. In this respect, the *quality* of credit tends to increase during an expansion, and to decrease during a contraction. To some extent we can observe this qualitative fluctuation directly as fluctuation in the availability of credit to marginal borrowers. More generally we observe fluctuating credit spreads between the rates charged to qualitatively different borrowers, i.e. price.



Whatever the underlying cause of fluctuation, we can usefully think of it as involving a swing from *scarcity* to *elasticity* and back again. At all times, the monetary system can be characterized by the balance between these two dimensions.

The history of monetary theory is to a large extent comprised of a dialogue between two points of view, often distinguished as the Currency School versus the Banking School, which emphasize respectively the importance of scarcity and the importance of elasticity. From the point of view I have been developing, both have part of the truth but neither has it all. Thus, liquidity is at the same time both naturally scarce and naturally elastic. How can this be so?

The Scarcity of (ultimate) Money	↔	**Currency Principle**
The Elasticity of (derivative) Credit	↔	**Banking Principle**

The natural scarcity comes from the fact that agents at any particular level in the hierarchy cannot by their own actions increase the quantity of the forms of money at a higher level than themselves. Just so, governments cannot increase the quantity of gold, and banks cannot increase the quantity of government currency. The availability of money thus serves as a constraint that holds the system back in its attempts to expand.

The natural elasticity comes from the fact that agents at any particular level in the hierarchy can, by their own actions, increase the quantity of forms of credit at their own level, and possibly also below them. If you and I want to make a trade and you are willing to accept an IOU from me, then we can trade and what makes the trade possible is an expansion of credit. The elasticity of credit thus serves as an element of freedom that facilitates breaking loose from any constraint that may be standing in the way of expansion.

This natural elasticity applies to banks as well. By trading among themselves, banks can and do break loose of the constraint of central bank reserves. The important point is that the system involves at all times a balance between discipline and elasticity, with sometimes one and sometimes the other serving as the more dominant feature.

IV.

I have used the word "natural" in my title, and now I want to explain why. I use it to emphasize that the hierarchical character of the system, and its dynamic character over time, are deep features of the system. The institutional organization of the monetary system is hierarchical because of this underlying feature, not vice versa. That is to say, the hierarchy is not something imposed from the top down, e.g. by the government or the central bank. Monetary systems are naturally hierarchical, from the ground up. This is probably a controversial point of view, so I had best be careful to explain what I mean. (I do not mean to suggest that the monetary system is self-organizing in the sense meant by some Austrian thinkers, as Menger).

Rather I think of the institutional organization of the monetary system as inherently involving a system of **market makers** at different levels of the natural hierarchy. The term "market-maker" may be familiar to you from finance since a security dealer is a kind of market maker. A security dealer stands ready to buy or sell a security at a given price (actually two prices, the buy-sell spread) in terms of money. He does this by holding an inventory of both securities and money (actually an inventory of credit instruments that provide access to inventories of securities and deposits held elsewhere, i.e. *repos* and *reverse repos*).

I propose to think of banks as a special kind of security dealer that stands ready to buy or sell a deposit at a given price (now only one price) in terms of currency. And I propose to extend the idea also to the central bank which, under a gold standard, stands ready to buy or sell currency in terms of gold. Both banks and central banks are thus like specialized types of security dealers. We'll develop this point in much more detail later on in the course.

For now, thinking of our simple hierarchy of money, the point is that there is a simple hierarchy of market makers to go along with the hierarchy of instruments. And for each market maker, there is an associated price of money. The prices in the simple hierarchy are three: the exchange rate (the price of currency in terms of gold), par (the price of deposits in terms of currency), and the rate of interest (the price of securities in terms of deposits or currency, assuming par). These prices are the *quantitative* link between layers of *qualitatively* differentiated assets. At any moment in time, market makers are straddling two layers of the system, with one leg in one layer and the other in another. More accurately, the two legs are the two sides of the market- maker's balance sheet.

Simple Hierarchy of Market Makers

Asset	Market Maker	Price
Gold	Central Bank	Exchange Rate
Currency	Banking System	Par
Deposits	Security Dealers	Interest Rate
Securities		

If the market makers do their job well, we will observe continuous markets at the various prices of money. In other words, the qualitatively differentiated hierarchy will appear as merely a quantitative difference between various financial asset prices. It is this transformation from quality to quantity that makes it possible to construct theories of economics and finance that abstract from the hierarchical character of the system (as most do). But the hierarchical character remains, and shows itself from time to time, especially when the market makers are not doing their job well, such as during periods of financial crisis or under the extreme stress of war finance.

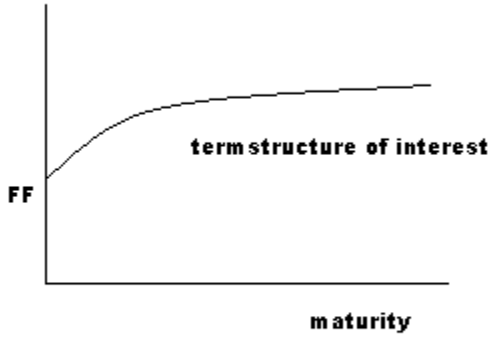
Even in less extreme times, fluctuation in the natural hierarchy shows up as strain on the market making institutions. A narrowing of the hierarchy means an increased qualitative differentiation between credit and money, and in the course of that differentiation there is bound to be pressure on the quantitative price of credit in terms of money. Just so, interest rates can and do change to reflect that stress, but that quantitative change is not necessarily equilibrating, and might just make the strain worse. Further, the whole point of par is that it is a price that does not change, and under a fixed exchange rate system the same is true internationally. The banking system thus comes under stress as the natural hierarchy fluctuates because it is bound to defend the fixed price between different layers of the hierarchy. Reserves are one line of defense but, as we shall see, not at all the only line of defense.

V.

So far, I have been talking about market makers as a reactive bunch who respond to fluctuations in the natural hierarchy that are outside their control. I have also been implicitly assuming that their behavior is driven by the dictates of **profit maximization**. The whole idea of *monetary policy* however is to intervene actively with some objective in mind other than profit maximization. We can understand monetary policy as an attempt to manage the natural fluctuation of the system for the general good, rather than for the profit of the central bank.

So for example at a time when the natural hierarchy is flattening, instead of waiting for a financial crisis to reassert the scarcity of money, the monetary authorities might try themselves to reassert the scarcity by raising the Fed Funds rate ("taking away the punch bowl" in the immortal words of William McChesney Martin). Or, at a time when the natural hierarchy is very steep, the monetary authorities might try to reassert elasticity by lowering the Fed Funds rate ("pushing on a string" in the immortal words of John Maynard Keynes).

Note here that the main policy instrument of the central bank is the overnight borrowing rate. There can be considerable slippage between that rate and the longer term interest rates on which important investment and spending decisions depend. One of the purposes of this course will be get a better sense of the sources of that slippage.



I've said that the central bank intervenes for the general good. What does it mean, "the general good"? In this regard the ambitions of the monetary authority have shifted over time. The very first central banks focused their attention simply on "managing the hierarchy" to protect par and the exchange rate. If ordinary banks were having difficulty defending par, the central bank could help them out by providing additional reserves—this is the essence of the classic "lender of last resort" function. On the other hand, if the central bank itself was having difficulty defending the exchange rate, it could raise the interest rate in order to attract international reserves (gold). Clearly both of these interventions are natural objectives for a market maker that conceives of itself as safeguarding the monetary system.

Modern central banks have larger ambitions. They are typically concerned with "managing the economy." But their actions still have their impact only to the extent that they succeed in influencing the natural hierarchy. Modern central banks are perhaps not so much concerned with the shape of the hierarchy per se as they are with how that hierarchy articulates with the real economy, specifically aggregate demand and aggregate supply. That's fine, but it is vital not to lose sight of the underlying mechanisms of money and credit. As you might expect, the attempt to use monetary policy for non-monetary purposes can put strain on both par and the exchange rate, and more generally on the institutions charged with maintaining quantitative equivalence between qualitatively different levels of the hierarchy. Standard analytical frameworks in macroeconomics (such as IS-LM) tend to abstract from such strains, and hence from the limits of monetary policy. In this course, we will not be abstracting from such effects, but rather bringing them up to the center of our attention.

Financial globalization poses many challenges, but one of the most important is to the autonomy of central banks. Perhaps we are back to managing the hierarchy.

3. Money and the State, the US Case

Last time I painted a picture of private money and private credit, a picture in which the central bank appears as a **banker's bank**. Today I want to bring the state back in, and with it the conception of the central bank as a **government bank**. And I'm going to do it taking American monetary history as my subject, and using the balance sheet apparatus we have been starting to build up, so that we understand the government as just another balance sheet. American monetary history is a fairly wild story; I will largely leave aside for today the whole issue of the money standard—gold, silver, or fiat—but we will come back to it after the midterm.

I want to begin with the picture Young provides of the structure of the national banking system in the years before the establishment of the Fed (p. 302).[1] We are going to ask where this system came from and why and how it got replaced by the Federal Reserve system. As we shall see, it is a story of war finance in the Civil War, when the North fought the South, as well as lots of political constraints given American antipathy for central banking. Note that before the Fed there was no official central bank, but there was a collection of New York banks that operated informally as such for their members. Behind the scenes are Big Finance and Big Government, the two big bogeymen of American monetary experience, and the populist agitation for easy money for themselves, not for Wall Street and not for Washington.

It is a story in three acts:

Act 1:	**Civil War Finance**	**Greenback Era**
Act 2:	**National Banking System**	**currency principle, scarce money**
Act 3:	**Federal Reserve System**	**banking principle, elastic credit**

During wartime, the problem of any government is to raise as much money as it can, first through taxation and then by borrowing in anticipation of future tax receipts. Borrowing means selling new government bonds to the private sector. The way this happens is as follows:

Government		Private Sector		Banking Sector	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+deposits, G	+bonds	-deposits, PS +bonds			-deposits, PS +deposits, G

Here I am showing the private sector buying bonds from the government by writing a check on their bank accounts. The effect of that check is to transfer the deposit account from the private sector to the government, so that at the end of the day the government has more money to spend for the war. Note how every entry in the table appears twice. The new bonds enter as a new

[1] For details see Wesley C. Mitchell, The Greenbacks; and Allyn Young, Chaps. 31-34 in "Commerce: The Marketplace of the World", 1924. Reprinted as pp. 265-321 in Mehrling and Sandilands, ed. Money and Growth, Routledge 1999.

liability of the government and a new asset of the private sector. Deposits are similarly an asset of the government or private sector and a liability of the banking sector, so payment for the bonds enters as a reduction in both private sector assets and banking sector liabilities, and as an increase in both government assets and banking sector liabilities. In the banking sector, total deposit liabilities remain unchanged; all that has happened is the bank now owes the government what it used to owe the private sector.

That is how things work when the government is able to issue bonds and sell them to the private sector. But in war time, that might not be enough. The private sector may simply refuse to lend, or lend only at an unacceptably high rate. In that case, the temptation is always to sell the bonds directly to the bank, bypassing the private sector entirely. After all, what the government needs is a bank deposit, so why not simply swap IOUs with the banking system as follows:

Government		Private Sector		Banking Sector	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+deposits, G	+bonds			+bonds	+deposits, G

In this case the banking sector is directly funding the government, not just transferring funds it is holding for the private sector. And it is funding the government by expanding the supply of money. How so? When the banking system swaps IOUs with the government, it expands both sides of its balance sheet, so total deposit liabilities increase.

The story of Civil War finance begins with just such an operation. The overriding theme of Young's discussion is the maximal use of existing money and credit institutions in order to finance pressing war needs, so testing the limits of those institutions. The resulting stress proved too much, with the result that the entire banking system broke free of gold and only returned 17 years later in 1878. One way to understand the economics involved is to follow closely the moves made by Salmon P. Chase, Secretary of the Treasury.

Act 1: War and Aftermath. First, in August 1861, he took out a big bank loan and then withdrew the proceeds, hence gaining control of the banking system's gold for his own war purposes (see p. 281). This allowed the North to buy needed material from abroad, while the South was forced to rely on barter for cotton exports. But it also forced suspension of specie payments domestically. Let's see how that worked.

The Loan

Banking System		Treasury	
Assets	Liabilities	Assets	Liabilities
Gold +Loan, $150	+Deposit, $150	+Deposit, $150	+Loan, $150

Note here how the government loan is made initially by expanding the bank balance sheet on both sides, and the government balance sheet as well. I like to call this "swapping IOUs", and

consider such a swap to be the essence of banking, at the heart of the apparently alchemical ability of banking to create money from thin air.

What Salmon P. Chase did next was not to spend the deposit, which would merely have transferred it to someone else, but rather to withdraw it in gold.

The Withdrawal

Banking System		Treasury	
Assets	Liabilities	Assets	Liabilities
Loan -Gold	-Deposit	+Gold -Deposit	Loan

Note that when the government withdraws the deposit, the banking system loses its accumulated gold. Once the banking system has no more gold, its promises to pay gold (its deposits) lose their credibility, and banks accordingly "suspend convertibility". Deposits are no longer promises to pay gold, so what are they?

The answer comes in Salmon Chase's next move; bank deposits become promises to pay government issued legal tender. In 1862, because of the weak market for government bonds (on account of risk, possibly overestimated, but then the banks had been burned once already), the government decided not to try to issue marketable bonds but rather to rely on legal tender note issues.[2]

Government		Business		Banking System	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+war goods	+legal tenders, $400 MM	-war goods +legal tenders -legal tenders, $100 +deposits, $100		+legal tenders, $100 MM	+deposits, $100

Essentially what the government did was to insert legal tenders between gold and bank deposits in the hierarchy of money. At the same time it broke the connection between the dollar and gold, since the legal tenders were not a promise to pay any specific quantity of gold. Over

[2] I am deliberately using the term "legal tender" rather than "fiat money" in order to emphasize that the notes were liabilities that were expected to be redeemable at some future date in gold. The economics of the Greenback Era was analyzed by Wesley Clair Mitchell in his famous book A History of the Greenbacks (1903). Mitchell was a student of Laughlin, friend of Young, prof at Columbia where he taught a famous course "Types of Economic Theory".

the course of the war, the legal tenders depreciated against gold.[3] That means it took more and more dollars to buy the same quantity of gold, and hence also more and more dollars to buy gold-priced imports. At the same time, inside the US, it took more and more dollars to buy the same quantity of domestic goods (inflation). Legal tender money lost value not only against the better gold money, but also against non-money commodities.

Even after the war was over, it was many years before resumption of the gold standard. Eventually the North possessed the creditworthiness required to float bonds. Once it did, bonds could be issued to buy gold, and that gold was used to retire legal tenders. The retirement of legal tenders brought the dollar back to its prewar gold parity, and also deflated domestic prices to prewar levels.

Young makes a big point of how expensive it was to finance the war in this way. The legal tenders bought only 50 cents worth of gold during the war but were redeemed at par afterward. It would have been much cheaper if we could have financed the war by issuing bonds. (We also could have avoided the politically and economically damaging effects of shifts of wealth between creditors and debtors caused by inflation and then deflation.) In this respect, civil war finance compares unfavorably to the finance of WWI and WWII. In both subsequent wars, the government once again used the money and credit system maximally to finance its operations, but in each succeeding case the system was more developed than in the last. Thus in each succeeding case the system not only provided more finance but also did so without so much inflation.

Act 2: National Banking. Here the theme is the stress caused by seasonal expansion and contraction of private credit on a fixed note basis.

The National Banking System was established in 1863 during the Civil War as another of Chase's attempts to secure wartime finance. The overriding idea was to strengthen the market for government bonds. But the lasting effect was to fix the money supply.[4]

Chase's idea was to issue a special class of bonds that could be used to back the bank note issue, while at the same time putting in place measures (taxes) that eliminated other potentially competitive note issues. Let's see how that worked in detail. The initial bond issue can be thought of simply as a swap of IOUs with the banking system. This time however it would do no good to withdraw deposits because the banking system has no gold reserves left. Instead the banks were given the right to issue bank notes. Here is the initial swap.

Before

Banking System		Government	
Assets	Liabilities	Assets	Liabilities
+2% Bond	+Deposit	+Deposit	+2% Bond

You might wonder why the banking system would willingly engage in this swap, since it is getting only 2% interest, well below the market rate. The answer is that it has the right to issue bank note liabilities that pay 0% interest, so the difference is profit. Now, when government

[3] Mitchell explains this depreciation as a matter of market expectations about the probability of future return to convertibility, expectations that shifted whenever the North won or lost a battle.

[4] Historians have found that the note issue never reached the level of outstanding bonds, so in principle it would have been possible for banks to issue more notes as needed. The point is that they didn't, so inelasticity was the problem.

pays out the deposits, the bank has the right to pay using bank notes that are its own liability. The bank would of course prefer to back its notes with even higher yielding bonds, but the government took care of that by imposing a tax on all note issues that were backed by anything other than these special 2% bonds.

After

Banking System		Private Sector		Government	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
2% Bond	-Deposits +Bank Notes	-war goods +bank notes		-Deposits +war goods	2% Bond

The balance sheets show the end result. This third strategy for government purchase of war goods is financed by an expansion of <u>private</u> money (the bank notes), which are promises to pay the new standard money (legal tenders) not gold.

I want to now focus on the unintended consequence of this operation, once the war was over and legal tenders were retired, which was a *very inelastic money issue* because the collection of bonds that could be used to back notes was not increased except by special measure. There was also *very inelastic reserves*, indeed perversely elastic because of the reserve requirements, 25% for central city all in cash, 25% for middling cities partly in central city reserves, and 15% for country partly in central city reserves. Already in 1873 the country experienced the first of a series of financial crises, all of which followed a similar pattern.[5]

In slack times the farm banks would find themselves with excess funds for which they could find no local outlet. They might use them to buy a security (bond) but they had always to keep in mind that they would need the funds come fall. So they tended to deposit the funds in New York where they could earn interest. New York banks would therefore find themselves with excess funds, which they also knew were only seasonal, so they wanted a short term investment. They would buy liquid securities or make short term loans. Of particular interest is the phenomenon of the call loan made to stock market speculators. Thus in slack periods (late winter) we might find something like what Young shows (p. 302), where country banks have excess reserves. He mentions the number 50 million as the withdrawal at harvest time, which is pretty close to the excess 2% reserves. At harvest time there is a cash drain from the system, and that means a cash drain from New York, which New York seeks to remedy by calling in loans and raising reserves from abroad.

Thus the cash drain spread into the stock market, causing selling by those who were using call loans to finance their speculative positions. And it spread to the international money market, pulling in gold from London. The consequence was a very definite <u>seasonal pattern in interest rates</u>, as the harvest expansion of credit took place on a fixed reserve basis. The result was not only a seasonal interest rate but also periodic financial crises, caused whenever banks had to make cash payments but lacked the cash to do so. Young makes the correct point that the problem was the inelasticity of reserves. If somehow reserves could be reduced in slack times

[5] The classic reference is OMW Sprague <u>History of Crises under the National Banking System</u> (1910).

and expanded in tight times, the problem could be solved. How to make reserves elastic? The answer was to make reserves a form of credit.

Act 3: Federal Reserve System. What the banking system needed in times of tight money was a temporary increase in reserves (for the farm banks) and a temporary increase in cash (for circulation purposes). Under the National Banking System, typically some kind of stopgap measure would be worked out, involving an expansion of quasi-cash in the form of clearinghouse certificates issued by a consortium of New York banks. If banks agree to accept these certificates among themselves, then for practical purposes these certificates are reserves. The effect was therefore to free up cash for general circulation. After the crisis of 1907, the Aldrich Vreeland Act of 1908 created the legal basis for these ad hoc measures, but more was needed.

The Federal Reserve system routinized the private solution to both problems. Under the national banking system, the quantity of cash was fixed. Under the Federal Reserve System it became elastic. The way it was supposed to work was that banks could take certain kinds of loans (so-called real bills) to their local Federal Reserve bank and "discount" them for Federal Reserve bank notes that could be used in circulation. The various reserve banks could also discount at the Fed itself to get Federal Reserve notes which were legal as well as economic reserve. Shifts in demand as between deposits and cash could thus raise no problems; also cash drains into circulation would raise no problems; and total reserves could also fluctuate elastically.

Member Bank		Federal Reserve Bank		Federal Reserve	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+loan -loan (disc) +FR deposit	+deposit	2% bonds +loan (discount) -loan (redisc) +FR notes	FR bank notes +FR deposit	gold +loan (rediscount)	 +FR notes

The key to all this was supposed to be the discount mechanism, and hence also the discount rate. The idea was that member banks could always get whatever reserves they needed simply by presenting one of their eligible loans to the local FRB for discount. In turn the FRB could itself always rediscount for Federal Reserve notes. If the amount of new discounts exceeded the repayment of existing discounts, then total reserves would expand. If repayment exceeded new discounts, then reserves could shrink. No longer would the system spend half the time with too much reserve (fuelling speculation) and half the time with too little (risking financial crisis). Reserves would always be just right.

But WWI intervened. Instead of real bills, the central bank got stuffed with government paper. In fact, the Fed served as prime dealer for distributing government bonds, taking onto its own balance sheet whatever the private banking system and private bondholders would not absorb. After the war, instead of the planned elasticity driven by demand for discount of private credit,

the Fed used “open market operations”, outright purchase and sale of government securities, to adjust money supply to money demand.

Conclusion. The monetary system is a hybrid system, comprised of both private money (bank deposits) and public money (currency). The central bank is also a hybrid entity, both bankers’ bank and government bank. It is possible in principle to build a monetary theory around either one of these dimensions, focusing on either the private dimension or the state dimension, and treating the other as subsidiary. But any such theory is only partial since it misses the true hybrid character of the system.

The story of American monetary history is a story of getting the public/private balance right. It took a while, and indeed is still ongoing. The story of the development of monetary thought is similarly a story of getting the balance between these two dimensions right in order to capture the true dynamics of the system at any point in time. Sometimes the state dimension is dominant (war time) and sometimes the private dimension is dominant, but all the time both dimensions are present.

4. The Money View, Micro and Macro

http://www.federalreserve.gov/releases/z1/Current/z1.pdf (see full matrix at beginning) Notable features—household deleveraging, switching from credit to money, instrument discrepancy is repo, sectoral discrepancies

Last time we saw how the US banking system was born from the strains of war finance and financial crisis, and we also saw how understanding balance sheet relationships can help us to understand the underlying processes. Today we focus more specifically on the balance sheet approach that will be used throughout the course, and to aid that focus we confine our discussion to the most placid of events, namely the use of the banking system to facilitate ordinary daily exchange.

Payment Systems: Money and Credit

Suppose you and I do regular business with each other. You produce a good that I want, and I produce a good that you want, but for some reason supply and demand for the two goods are not precisely coordinated over time. One way of organizing our interaction is with the help of money. I buy your goods by giving you money, and you buy my goods by giving me money. Over time, my money balances fluctuate and so do yours, while total money stays the same.

Me		You	
Assets	Liabilities	Assets	Liabilities
+goods -ΔM		-goods +ΔM	

Note however that another way of organizing our interaction is with credit or promises to pay. I buy your goods by giving you an IOU and you buy my goods by giving me an IOU of your own, or giving me back one of my own, so there is only net indebtedness between us. Note that the promise to pay is never actually paid, only offset by other promises, so there is no real need for money as such. We could organize the whole thing by promising to pay some abstract unit of account with no physical existence. In some ways this **pure credit** payment system is more flexible than the pure money payment system, since we are not limited by the total money supply, only by mutually agreed credit limits. Observe that in this system, unlike the money system, the quantity of outstanding IOUs fluctuates over time.

Me		You	
Assets	Liabilities	Assets	Liabilities
+goods	+ΔIOU	-goods +ΔIOU	

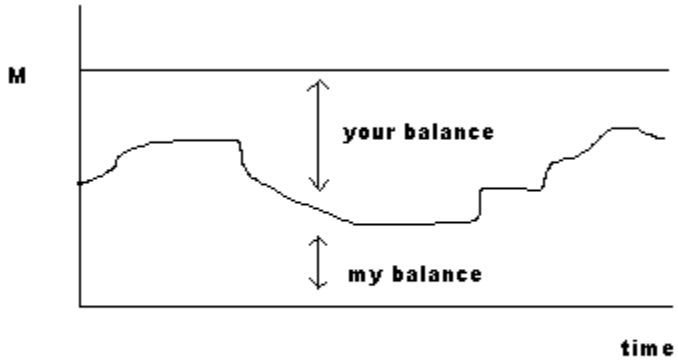
Now consider yet a third way we might organize our interaction. Suppose neither of us trusts one another sufficiently to extend bilateral credit, so the bilateral credit limit is zero. But we both trust some third party, and that third party also trusts each one of us. In this case we can organize our exchange by issuing IOUs to and accepting IOUs from the third party. It seems reasonable to call these third party IOUs "money", and to call the third party a "bank". Then the relevant balance sheet entries are as follows:

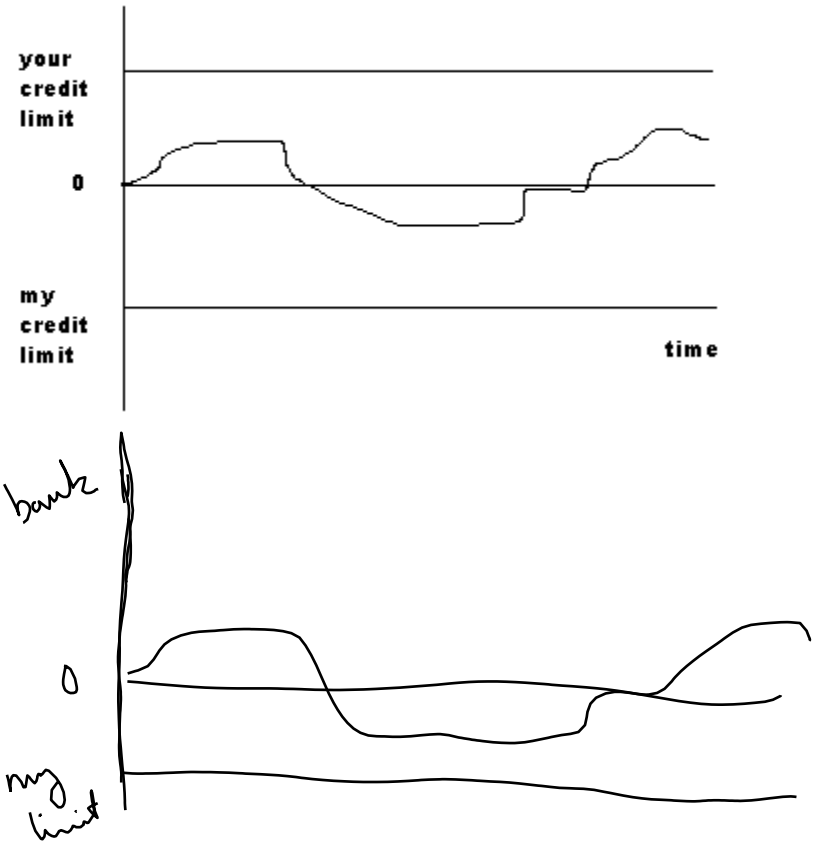
Me		Bank		You	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+goods	+ΔIOU	+ΔIOU	+ΔM	-goods +ΔM	

Note how the quantity of bank money fluctuates over time, as bank credit expands and contracts in order to facilitate the time pattern of trade. It is of course this third system that most resembles the institutions of the modern developed economy.

Discipline and Elasticity

I have emphasized before that at every moment in time there is a balance between discipline (which comes from the scarcity of money) and elasticity (which comes from the availability of credit). In our first example, the discipline came from the limited quantity of money—when either side ran out of money, they could no longer buy and trade stopped. In the second example, the discipline comes from the bilateral credit limit. In the third example the discipline comes from the credit limit and terms imposed by the bank on each borrower, and the elasticity comes from the willingness of the bank to swap its own IOU (which is money) for IOUs farther down the hierarchy (which are credit). We can understand these three examples as representing different balance between discipline and elasticity.





Money and The Real World: Micro

Cash Flow is the most basic concept in this course. Here we follow the lead of Hyman Minsky who writes:

> To analyze how financial commitments affect the economy it is necessary to look at economic units in terms of their cash flows. The cash-flow approach looks at all units—be they households, corporations, state and municipal governments, or even national governments—as if they were banks. (Minsky 1986, p. 198)

What this means is that we view every economic agent as an entity experiencing a certain inflow of cash (receipts of various kinds) and outflow of cash (expenditures of various kinds) over time. The most basic **survival constraint** (or "reserve constraint", also Minsky's terminology) facing the agent is that the inflow must be at least as big as the outflow. Receipts and expenditures on commodities fluctuate over time relative to one another. If at a moment in time expenditures are greater than receipts, then cash flows out from hoards. If hoards are exhausted they may be replenished by borrowing, but that

just puts off the day of reckoning, so it only works if there is a date when receipts are expected to be greater than expenditures.

$$\text{Cash Flow} - \text{Cash Commitment} \geq 0$$

Sources and Uses accounts can help us to understand all this in more detail. Every transaction can be captured as a simultaneous 4-part entry (at least) in this system of accounts.

	Uses	Sources	
Goods and Services	Expenditures	Receipts	
Financial Assets	Accumulation	Decumulation	Credit
Financial Debts	Repayment	Borrowing	Debit
Money	Hoarding	Dishoarding	Money

This set of accounts differs from the T-accounts of last lecture in that all entries are flows, not stocks.

You can see the hierarchy of money below the line, and real expenditures above the line. (Note that I have a separate entry for Financial Assets and Financial Debts, in order to capture the feature that individuals are responsible for their gross debts, not just their net debts.) We could subdivide even more finely to take account of the different kinds of money and the different kinds of credit. Just so, if this were a bank, we would want "money" to refer to reserves. If this is an individual, we might want money also to refer to deposit accounts.

Two rules structure the accounts:

Rule 1: For each agent, every use has a corresponding source, and vice versa.
Rule 2: Each agent's use is some other agent's source, and vice versa.

Example 1: Money. Buy a cup of coffee from Oren's Daily Roast for $2 cash

Me		Oren's Daily Roast	
Use	Source	Use	Source
Expenditure, coffee			Receipt, coffee
	Dishoarding, $2	Hoarding, $2	

Example 2a: Credit. Buy dinner at Vareli for $20 using credit card

Me		Vareli		Mastercard	
Use	Source	Use	Source	Use	Source
Expenditure			Receipt		
		Accumulation (MC)		Accumulation (Me)	
	Borrowing (MC)				Borrowing (Vareli)

Example 2b: settle MC account (Vareli)--daily

Me		Vareli		Mastercard	
Use	Source	Use	Source	Use	Source
			Decumulate		
				Repay	
		Hoarding			Dishoarding

Example 2c: settle MC account (Me)--monthly

Me		Vareli		Mastercard	
Use	Source	Use	Source	Use	Source
					Decumulate
Repay					
	Dishoarding			Hoarding	

This example shows the intimate connection between cash and credit in the payments system. Observe the enormous amount of action "below the line" that is necessary in order to make goods and services transactions "above the line" possible. Observe further the way that credit provides elasticity, expanding to facilitate transactions, while money provides discipline, the requirement to contract credit back down again after a certain time.

Money and the Real World: Macro

The Flow of Funds accounts are built on Sources and Uses methodology. There is however a lot of aggregation and netting.

> Aggregation into sectors: Households, businesses, financial institutions, etc…
> Some netting within sectors: so missing some household to household flows
> Considerable netting over time: quarterly statements only

Despite all this netting, rules 1 and 2 still apply, and the consequence is a disciplined matrix structure of accounts. For any agent, total uses and total sources must balance (column sums). And for any instrument, total uses and total sources must balance (row sums).

http://www.federalreserve.gov/releases/z1/Current/z1r-3.pdf

This set of accounts was originally promoted by Morris Copeland as a possible framework for macroeconomic analysis, in his book A Study of Moneyflows in the United States (1952). He explicitly proposed it as an alternative to the national income and product accounts, which had been the basis of the Keynesian revolution:

$C + I + G + X-M = Y$

In NIPA accounts, the emphasis is on value added and employment, so we focus on final production. But used goods are also exchanged, and also financial assets. These exchanges are shunted off to one side by NIPA but are at the same level of analysis in FoF. Indeed, in FoF the sale of goods and the sale of assets are equivalent ways of achieving a source of funds.

The accounts were also intended as an alternative to the traditional quantity theory:

$MV = PT$

In this accounting system, transactions include everything, not just final income transactions, but all transactions are treated as if they are made with money. In a way the Keynesian framework grows from the quantity theory, with C+I+G+X-M serving as a kind of disaggregation of MV, and Y serving as a specification of a subset of PT. Copeland wanted to go even farther but he did not win out. Actual macroeconomic debate was between Keynesians and monetarists, and FoF remained a specialty interest for those who wanted to track developments in the financial world (below the line).

Money and Time: Liquidity and Solvency

The central concern from a banking perspective is not solvency but **liquidity,** i.e. the survival constraint. Are current cash inflows sufficient to cover current cash outflow commitments? If yes, then we satisfy the survival constraint. If no, then we have to raise additional cash flow in some way and logically there are only three ways to do it:

(1) Spend down hoards of money
(2) Liquidate accumulations of financial assets
(3) Borrow, i.e. build up additional stocks of financial liabilities.

Notice that both (2) and (3) depend on finding someone else to take the other side of your trade, and that might be impossible or extremely expensive in times of crisis. In times of crisis only (1) is dependable. That is why reserves are important.

The survival constraint must be met not only today but also at every moment in the future. Thus, generally, **the problem of satisfying the survival constraint is a problem of matching up the time pattern of cash flows with the time pattern of cash commitments**. The central question is whether at any moment in time actual cash flows are **validating** promised cash commitments. Problems on that score show up in the money market where people unable to make payments from their existing cash flow face the problem of raising cash, either by issuing a financial asset (borrowing) or selling one (liquidation).

To see the point concretely, consider the following examples. In all three cases cash inflows are much greater than cash commitments, so there is no problem with solvency. But in the first case the survival constraint is met in every single period. In the second case the survival constraint is met in future periods but not the current period, so there is a current problem. And in the third case, the survival constraint is met in present periods but not future periods, so there is a potential problem.

Liquid Capital Structure

	T	t+1	t+2	t+3
Cash Flow	10	10	10	10
Cash Commitment	5	5	5	5

Illiquid Capital Structure (current mismatch)

Cash Flow	10	10	10	10
Cash Commitment	20	0	0	0

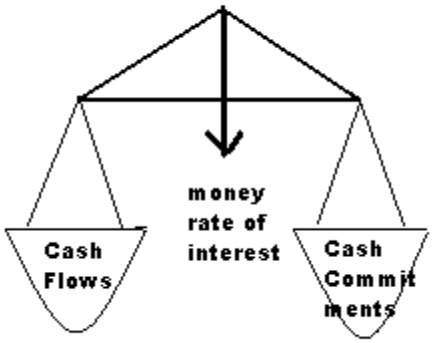
Illiquid Capital Structure (future mismatch)

Cash Flow	10	10	10	10
Cash Commitment	0	20	0	0

The point to emphasize is that at any moment in time there are agents in the economy who fit into each of these cases, and all of them meet in the money market. The ones with current mismatch (like Case 2) are necessitous borrowers—they have to borrow no matter what it costs. The ones with future mismatch (like Case 3)may enter the money market today in order to avoid future problems, or they may decide to wait. The ones with current liquidity (like Case 1) may decide to help the others out, or they may decide to do something else. The result of all this pushing and pulling is the money rate of interest.

The economy thus comes to appear as a system of interlocking balance sheets in which individuals depend on one another's promises to pay (financial assets), and build these promises into their own projections of future cash flows. In the economy as a whole there is a pattern of cash flows emerging from the "real" side, production and consumption and trade. And there is a pattern of cash commitments more or less explicit in the financial structure. At any moment there is a balance between the two, which shows up as a price (interest rate). If there is mismatch, then someone in the economy must be persuaded to give up current cash for a mere promise of future cash, and the relative price of these two is the rate of interest. It is for this reason that **problems of mismatch between cash flows and cash commitments show up as upward pressure on the short term money market rate of interest**.

Flow Balance



At the most basic level, the mismatch between cash commitments and cash flows is something that we see in the current pattern of payments. However that current pattern reflects also a forward-looking view since people can and do anticipate problems they may face in satisfying the survival constraint in the future. If they can move a future

problem into the present where it is easier to handle, they do so. This rearrangement can put current pressure on the money market today even though there is no problem making current payments. In this sense, the state of the money market is an indication of how well or how poorly the entire time-pattern of financial commitments matches the entire time-pattern of cash flows, not just today but also looking forward.

From Flows to Stocks

These fund flows link up to our balance sheet T accounts as changes in the outstanding stock of each particular asset or liability. The stock of outstanding IOUs of any particular description is just the sum of all the past flow accumulations of that particular asset. Just so, we can cumulate all the flows and write the balance sheet for any agent as

Assets	Liabilities
Money Financial Assets	Financial Liabilities Net Worth

Every agent has a balance sheet like this, so we can conceptualize **the economy as a set of interlocking balance sheets**, in which the financial liabilities of one agent are the financial assets of another. These liabilities are promises to pay, usually promises to make a specific series of payments at specific times in the future.

In standard accounting practice, there is a lot of attention paid to the matter of **solvency**:

Value of Total Assets-Value of Total Liabilities = net worth > 0

To make this assessment, we need to value the assets and liabilities. Standard finance theory approaches this problem of valuation as a matter of finding the present value of future cash flows. We want to include not just financial assets but also real assets, since they presumably also involve future cash flows. From a financial point of view, a factory is nothing more than a particular kind of bond.[1]

For example, a factory that is expected to produce an annual net cash flow of $10 in perpetuity has a present value of $200 if the interest rate is 5% and expected to remain so.

Net Cash Flow = {10, 10, 10, 10, …..}
Present Value CF = $\{10/(1+R), 10/(1+R)^2, 10/(1+R)^3, \ldots.\}$
Capital Value CF = $\Sigma\delta^t CF_t = 10/R = \$200$.

Some such calculation gives us the value of financial assets, and the value of financial liabilities. The value of money is of course much easier because the face value is the present value.

[1] Minsky, Stabilizing an Unstable Economy, p.

For our purposes the question of solvency is interesting mainly as an outer bound on **the credit limit** facing each agent. Intuitively it makes sense that that credit limit will be somehow related to the net worth. Solvent agents have unused borrowing power on their balance sheets which they can potentially mobilize to make payments. Thus we can see how asset price fluctuations can cause fluctuations in borrowing power, which might have consequences for immediate liquidity. Solvency problems can easily become liquidity problems.

And liquidity problems can also easily become solvency problems. When money market interest rates rise in a scramble to put off to tomorrow what is due today, the consequence is to discount future cash flows relative to the present, hence fall in asset prices. Any agent holding those assets will have an unrealized capital loss, perhaps large enough to eliminate net worth.

5. The Central Bank as a Clearinghouse

Today we start to get a little more specific about one of the two central themes of the course, banking as a clearing system. We begin with the idea of payment as a balance sheet operation within the banking system, and the observation that, if the banking system were one big bank, there would be no need for reserve flows. **One way to understand the evolution of the banking system is as an attempt to make the system operate as if it were one big bank.**

One Big Bank

Suppose there was only one bank, and everyone in the economy had an account. So the balance sheet might look like this:

Bank A

Assets	Liabilities
Reserves	Deposit account α
	Deposit account β

In this world it is natural to think of the "quantity of money" as the quantity of deposit accounts at the bank. Payments made by any depositor in A to another depositor in A simply subtract and add the same number on the liability side of the bank balance sheet; reserves never move. The only drawback of the Bank A system is that you have to have a positive account in order to make a payment. It's a money payment system, not a credit payment system.

But there is really no reason to insist on that. A more general form of one-bank banking would allow overdrafts. Negative deposit accounts show up on the asset side of Bank B.

Bank B

Assets	Liabilities
Overdraft γ	Deposit account α
Overdraft λ	Deposit account β

This innovation has a somewhat startling implication. Note that if γ makes a payment to α, that involves an expansion of both sides of the bank's balance sheet, in effect an expansion of credit. Contrariwise, if β makes a payment to λ, that involves a contraction of both sides of the bank's balance sheet, in effect a contraction of credit. The use of bank credit as a means of payment thus involves a certain elasticity in the quantity of money.[1]

[1] Note in passing that this way of thinking about the payments system raises deep questions about how properly to measure the money supply. Consider three possibilities:

(1) Σdeposits, but this measures only those with positive balances
(2) Σ(deposits – overdrafts)
(3) Σ(deposits-credit limits)

Let us tally up the effect of different payment flows on the size of the balance sheet of Bank B.

	α	β	γ	λ
α	0	0	-	-
β	0	0	-	-
γ	+	+	0	0
λ	+	+	0	0

Multiple Banks

But of course there isn't just one bank, and decentralized banking causes problems for settling payments. To see the problem, think of a world with two banks.

Bank A		Bank B	
Assets	Liabilities	Assets	Liabilities
Reserves	Deposit account α	Reserves	Deposit account γ
	Deposit account β		Deposit account λ

The problem occurs when someone with an account in Bank A wants to make a payment to someone in Bank B. One way for this to happen is for Bank A to transfer reserves to Bank B. The effect is to shrink the balance sheet of A and expand the balance sheet of B. But if this is how payment actually takes place, we lose almost all the advantages of banking. It's almost as if people are making payments with the gold reserve. There is a better way.

Instead of transferring reserves for each order of payment, suppose that each day each bank collects "due tos" and "due froms" with respect to every other bank.

Bank A		
Assets	Liabilities	
Due from A	Due to A	zero
Due from B	Due to B	Net to B
Due from C	Due to C	Net to C
Due from D	Due to D	Net to D

At the end of the day, each bank nets the payments to each other bank and pays only the net in gold. Call this "bilateral intraday netting". This clearly involves less gold transfer than the first system, but there is a better way yet.

Suppose, for example, that over time net payments to B tend to average out. Then it makes no sense to move gold every day that is just going to move back. Instead it is more

---

The ambiguity about how to measure money in such a world led Fischer Black to propose that we not try. See Black "Banking in a World without money" (1971)

efficient to transfer virtual gold, by transferring "correspondent balances". Thus, if A owes B, A can pay B by drawing on balances held at B, or by increasing the balance held at A to B's credit.

In general we can think of correspondent balances as beginning with a simple swap of IOUs between Bank A and Bank B:

Bank A		Bank B	
Assets	Liabilities	Assets	Liabilities
+deposit at B	+deposit from B	+deposit at A	+deposit from A

Then we use debits and credits to these accounts as a way of making the interbank transfers involved in payment between customers of Bank A and Bank B. Logically there are two possibilities:

Case 1: Correspondent Balances held at B are debited

Bank A		Bank B	
Assets	Liabilities	Assets	Liabilities
Reserves	-Retail Deposits	Reserves	+Retail Deposits
-Deposit at B			-Deposit from A

Case 2: Correspondent Balances Held at A are credited

Bank A		Bank B	
Assets	Liabilities	Assets	Liabilities
Reserves	-Retail Deposits	Reserves	+Retail Deposits
	+Deposit from B	+Deposit at A	

In the first case total deposits (including interbank balances) fall, while in the latter case total deposits (including interbank balances) rise.

In practice, the choice between Case 1 and Case 2 is always resolved so that the more central bank accepts deposits from the less central. In the US case it was country banks holding deposits with city banks. These reserve cities survive as the cities with one of the twelve Federal Reserve Banks.

The invention of correspondent banking amounts to moving from a money payment system to a credit payment system. Note that, since the correspondent system is a credit system, we are not constrained by the quantity of gold, only by the various bi-lateral credit limits. One can imagine an entire banking system using these book entries to clear bi-lateral net payments at the end of the day. But there is an even better way.

Clearinghouse

In a system of bilateral bank payments, any individual bank may at the same time be making payments to one bank while receiving payments from another. Obviously it would advantageous

to devise a system where the bank only had to pay the net across all its correspondents. One simple way that kind of system develops is when all banks hold correspondent balances in only one bank, and use those balances to clear. You can see how that goes some way toward creating a one-bank payment system. The problem arises when any member bank runs his balance to zero. Elasticity of the payments system requires the correspondent to extend an overdraft, but he might not want to, especially if the demand for overdrafts comes from more than one bank at once. (As we will see next time, the correspondent balance system was replaced by the Fed Funds market in which deficit banks could borrow from surplus banks anywhere in the system, not just their correspondents.)

A better system is the clearinghouse, in which banks of approximately equal stature in the hierarchy come together to form a mutual organization, owned by their members in proportion to capital subscribed. Just so, the New York Clearinghouse Association (now CHIPS). Capital is subscribed in gold or other legal reserve, and determines initial holdings of clearinghouse certificates. Subsequently, members treat all sums due to or due from other members as due to or due from the clearinghouse. During the day all members net payments multilaterally with all other members, building up intraday credit or debit balances against the clearinghouse. At the end of the day each member makes or receives only a net payment, depending on whether the net is negative or positive. Settlement involves transfer of clearinghouse certificates, which may require prior bi-lateral borrowing between deficit banks and surplus banks.

NYCA

Assets	Liabilities
Gold (Security)	Clearinghouse Certificates (owned by A, B)
Intraday net due from A	Intraday net due to B

The advantage of the clearinghouse is that all members guarantee all payments. The entire capital is available to make good on the negative clearing of a member in case the member itself cannot. If A cannot pay, then all the members pay in proportion to their capital subscription. Of course no one wants this. The clearinghouse policy committee therefore tends to be given wide powers to help individual banks, lending from the CH or coordinating accommodation through other banks.

Of particular interest is what happens in a financial crisis, when all member banks find themselves short of gold, because of outflows into circulation or abroad. Then there is no chance of solving the problem by arranging for weak banks to borrow temporarily from strong banks. Instead, weak banks borrow from the clearinghouse, which creates additional reserves from thin air simply expanding both sides of the balance sheet.

Assets	Liabilities
Gold (Security)	Clearinghouse Certificates
Member loans (6%)	CH **Loan** Certificates (6%)

In effect, what happens is that intraday deficits and surpluses are not paid but rather put off to

another day. Typically these member loans are highly overcollateralized, and pay interest of 6%.Similarly the CH loan certificates pay interest of 6%. The loan certificates are not gold, and so are not strictly speaking legal reserves, but in a crisis they come to be treated as similar to the clearinghouse certificates which are gold and are legal reserves. After the crisis they are liquidated.

Central Banking

Central banking can be understood as nothing more than one step beyond the clearinghouse, a kind of regularization and strengthening of the clearinghouse system that goes the extra step of obliterating the difference between clearinghouse certificates and clearinghouse loan certificates.

Assets	Liabilities
Gold	Deposits
+Discounts	+Currency

Clearinghouses can be broken by any individual member who insists on redeeming clearinghouse notes for gold. Central banks by contrast cannot be broken by internal forces because their own liabilities are the member banks' reserves. In effect, the central bank makes the system of banks operate as though it were a single bank because they all have clearing accounts at a single bank. All payments from someone within the system to someone else within the system net out. There is however still the problem of managing payments to and from agents outside the system and this problem is the origin of the "Art of Central Banking"

Two kinds of problem arise. It may be that the bank customers demand payment, so reserves flow out of the banking system into circulation as customers demand currency. This is called an internal drain and central bank balance sheet expansion can satisfy it. A more difficult problem is an external drain, where foreigners demand payment in gold. For this, a higher bank rate of interest can satisfy demand by discouraging discounts. (All this will be discussed in more detail in Lecture 9, p. 65).

6. Federal Funds: Final Settlement

Stigum (p. 507) says "The primary job of the manager of a bank's fed funds desk is to ensure (1) that the bank settles with the Fed and (2) that in doing so, it hold no more excess reserves than the amount, if any, that it can carry into the next week." From the perspective of this course, the manager's job can be understood as (1) satisfying the daily survival constraint and (2) satisfying the legal reserve requirement (special Wednesday afternoon settlement). For our purposes, the former is the more generally interesting task, especially today when all banks have massive excess reserves so that the legal reserve requirement does not bind.

The Survival Constraint: Balancing Discipline and Elasticity

The survival constraint says that cash inflows must be at least as large as cash outflows. The precise institutional mechanism that embodies that constraint for banks is the requirement to end the day with a non-negative balance in their reserve account at the Fed. To appreciate the importance of this constraint at the end of the day, it is useful to appreciate the way that banks are **allowed to relax the survival constraint during the day**. Indeed that violation is essential for the smooth working of the payments system because it allows banks to be the "first mover", to make payments before they receive payments. The institutional form that violation takes is the "daylight overdraft".

If Bank A sends over Fedwire an order to pay $10 to Bank B, but has insufficient funds to cover that payment, the Fed does not "bounce the check" but rather covers it temporarily. The effect is a temporary expansion of reserves in the system as a whole, thus:

Bank A		Fed		Bank B	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	+overdraft	+due from A	+reserves B	+reserves	

This expansion of reserves is an expansion of credit on the balance sheet of the Fed, an expansion that at its maximum typically exceeds the size of the rather small required reserves.[1]
Indeed, it is not stretching the truth much to say that these daylight temporary reserves, not the overnight required reserves, are the actual ultimate means of payment. The payment system is much closer to an ideal credit payment system than it is to an ideal money payment system.

The credit element in the payment system is not just on the (public) balance sheet of the Fed but also on the (private) balance sheet of CHIPS, the Clearinghouse Interbank Payments System. Every morning, member banks make payments to the Clearinghouse to open their

[1] Coleman, Stacy. "The Evolution of the Federal Reserve's Intraday Credit Policies." Federal Reserve Bulletin (February 2002): 67-84.

settlement accounts, but during the day they can and do run negative balances. The survival constraint kicks in at the end of the day, when they must bring their accounts back to zero, i.e. repay any daylight overdrafts. Here is the balance sheet relationship in the middle of the day:

Bank A		CHIPS		Bank B	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
-collateral +due from CH	 +due to CH	+collateral +due from A	+due to A +due to B +due to B	-collateral +due from CH +due from CH	

Both the Fed and CHIPS are concerned to make sure that the daylight loans are paid back at the end of the day. To ensure this there are, in both institutions, controls over how much you can borrow (credit limits), and requirements to post collateral as guarantee. At the Fed there is the additional requirement to pay interest on your average daylight overdraft, but if you look at the details (Figure 12.6 p. 527) that requirement is not very onerous, and in fact the interest actually paid is quite small. Notably, interest is paid on the average daylight overdraft over the whole day, so peak overdrafts for an hour are diluted in the calculation. Also, banks are allowed to borrow a certain amount, proportional to their capital, without paying interest.

Once we appreciate the credit character of the payments system, it is clear that every dollar owed as an overdraft by one bank is an extra dollar in the reserve account of some other bank. To get the credit off the balance sheet of the Fed at the end of the day thus requires cancellation of both the negative accounts and the extra positive accounts. The simplest way to do that is for the banks with extra positive accounts to lend their extra to the banks with negative accounts, which then use the loan proceeds to pay off their overdraft. In fact, the bank with extra reserves should want to make the loan since reserves don't pay interest whereas the loan will. The Fed Funds market is where that interbank borrowing and lending takes place for domestic banks (see below). The equivalent interbank market for CHIPS is the Eurodollar market which we will talk about next week.

Suppose this doesn't work, that deficit banks are unable to find surplus banks to lend to them. If a bank is unable to pay back its overdraft at the clearinghouse, then all the members of the clearinghouse are jointly responsible for the loan. If a bank is unable to pay back its overdraft at the Fed, the Fed is willing to extend the loan overnight by lending at the discount window, at the penalty rate of 100 basis points over the Fed Funds target. (Bernanke has lowered the discount rate to 50 points over, but probably that is just temporary.) So banks generally prefer to borrow in the Fed Funds market. And they don't wait until the end of the day; they spend a lot of energy tracking their payment flows and trading in the Fed Funds market during the day.

What Are Fed Funds?

The ultimate means of payment in the US banking system are deposits at the Fed. Call these deposits money. Fed Funds are (usually overnight) loans from one bank to another, payable in

deposits at the Fed. Call these loans credit. Until recently, the underlying deposits paid no interest, but the interbank loan does, and that interest rate is the Fed Funds rate.

To see exactly what is going on consider the following transaction between Bank A (the borrower or buyer of money) and Bank B (the lender or seller of money). Compare the example of Stigum (Table 12.2, p. 495).

Bank A		The Fed		Bank B	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+10 reserves	+10 Fed Funds		+10 Deposit A -10 Deposit B	-10 reserves +10 Fed Funds	

You can see how the transaction amounts to a loan between A and B, which means that the lender has to worry about whether the borrower will be able to repay.[2] After lending Fed Funds to Bank A, Bank B has gotten rid of his excess reserve position held at the Fed. After borrowing Fed Funds from Bank B, Bank A has the means to get rid of his deficit reserve position at the Fed. Like daylight overdrafts, Fed Funds are a credit element that lends elasticity to the payment system. But Fed Funds are overnight credit not daylight credit, and they are interbank credit not Fed credit.

Stigum makes a point that some banks are natural sellers of funds and others are natural buyers. Put another way, the regular business of some banks causes their daily cash inflow to exceed their daily cash outflow, and for some other banks it is just the reverse. Concretely, it seems that the former are small banks in isolated areas that don't face much demand for loans, while the latter are large city banks that can lend out all their deposits plus more. So the Fed Funds market channels excess funds from the country banks to the city banks. Viewed in this way, we can think of the Fed Funds market as analogous to the older pattern of correspondent banking. This country-city flow was largely intra-regional in the past, and so it remains today. (The regional character of correspondent banking is reflected in the location of the 12 Federal Reserve Banks.)

It is important to emphasize that a Fed Funds transaction involves an immediate flow of reserves between the banks. There is no waiting for the end of the day, no netting of inflows and outflows, like at a clearinghouse. Thus, once A has bought the money, she can go out and sell it. Transactions like this happen throughout the day, and so the quantity of Fed Funds contracts increases throughout the day. They are all of course cancelled when they are repaid the next day, but by the end of the day outstanding Fed Funds credit can be several times as large as the outstanding deposits at the Fed. What is the point of that? Elasticity. The system is not constrained by the quantity of reserves.

Payment versus Funding

At a higher level in the system, money center banks trade Fed Funds among themselves as a way of managing intraday cash flow fluctuations. Just so, suppose that Citibank makes a mortgage loan to me to purchase a house from you. On the day of the closing, Citibank has to make a large payment to the seller's bank. In general, there is no reason to suppose that Citibank

[2] This is what Stigum's discussion about "line problems" is about. Banks have FF credit lines with one another in order to limit their exposure. Brokers need to "fine-tune" quotes because it matters who is selling and who is buying.

has adequate reserves at the moment, so to make the payment it has to raise them in the money market. The transaction then takes place as follows:

Citibank (my bank)		HSBC		Chase (your bank)	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+mortgage	+deposit (me)				
+reserves	+FF loan	-reserves +FF loan			
-reserves	-deposit (me)			+ reserves	+deposit (you)
		+reserves	+FF loan	-reserves +FF loan	

The first line shows the mortgage loan, which involves a swap of IOUs between me and my bank. (I don't show my own balance sheet, but you can imagine what it looks like.) The second line shows the immediate **funding** of the loan by buying money from HSBC in the Fed Funds market. The third line shows the **payment** that finalizes the transaction, which involves transfer of reserves from Citi to Chase, as well as debit of my account and credit of your account.

After all this is done, Citibank has a mortgage loan asset that is funded by overnight money. Clearly this is not ideal funding, and the bank has some more work to do, but we leave that aside for the moment to concentrate on the payment rather than the ultimate funding. (The issue of ultimate funding is centrally addressed in Lecture 17.) Note also that Chase now has a large reserve position that pays no interest, so it also has some more work to do. Maybe it will sell in the Fed Funds market, maybe to HSBC? If so, then the ultimate source of funds that Citibank used to pay Chase is from Chase itself! I show this possibility on the fourth line. Note that HSBC is now back to zero in its reserve position at the Fed.

HSBC has in effect acted as an intermediary for interbank borrowing between Citi and Chase. In fact HSBC has acted as a **dealer** in the wholesale money market. And, even more, you have acted as the ultimate source of funds for Citi's loan to me.

Broker versus Dealer

Some of Fed Funds transactions are direct—the trader at Bank A calls the trader at Bank B—but a substantial number are brokered. Indeed, the real price discovery action (which is to say the determination of the "effective Funds rate") happens at brokers and dealers, and the distinction between them will be critical later on in the course, so I introduce it now.

A **broker** just puts together potential buyers and sellers from his sheet, much in the way that real estate brokers do with their listing sheets and client listings. Brokers keep lists of the prices bid by potential buyers, and offered by potential sellers, and they look for matches. There is an extra twist in the Fed Funds market that comes from the fact that sellers are also financing the sale, so that bids need to be fine-tuned to take account of credit lines.

Dealers, by contrast, take positions themselves. Read the chapter carefully and you'll find a number of references to dealer activities. What do they do? They buy funds at a low price and sell them at a high price. This sometimes means bridging different segments of the market--

small versus large tiering, unsophisticated versus sophisticated, domestic versus offshore (Euro). And it sometimes means bridging across time--intraday positions, cross-day positions. From one point of view, dealers are doing arbitrage, picking up nickels on the street and so creating a single price. From another point of view, they are making markets, and that is one big difference from brokers. This latter dimension we'll be talking a lot about two weeks from now. For now, I just want to note that it is there and move on. For present purposes, the important point is that dealing activity expands the balance sheet of the dealer, while simple brokering does not, thus:

Bank A		FF **Broker**		Bank B	
+reserves	+FF loan			-reserves +FF loan	

Bank A		FF **Dealer**		Bank B	
+reserves	+FF loan	+FF loan	+FF loan	-reserves +FF loan	

The Reserve Requirement

We have been focusing on the survival constraint and daily clearing, but what about this "reserve requirement" that Stigum talks about? Banks are required to hold a certain fraction of deposits in reserve (on average over the reserve maintenance period) and they can voluntarily pledge to hold more than that. (The banks holding more reserves than required are the money center banks, and they do it to facilitate the very large volume of transactions they are involved in.) Nowadays this constraint does not bind, but it is nonetheless of considerable historical interest.

Stigum makes a big point about the (monetarist-inspired) shift to contemporaneous reserve accounting from lagged reserve accounting in 1984, which made the central bank's job more difficult. Lagged reserve accounting meant that banks had to accumulate reserves as a fraction of past deposits. Since the demand for reserves was therefore completely inelastic, the Fed essentially was forced to provide them, either through open market operations which made them available through the Fed Funds market, or through the discount window. Given the choices, it preferred the former. The problem with this, according to monetarist thinking, was that it made the supply of reserves endogenous. They wanted banks to adjust deposits to fixed reserves, rather than vice versa.

The monetarists lost out on that one. Today the Fed supplies whatever reserves are needed to hit the Fed Funds target. This involves the Fed in practically daily interventions in the market (see Fig 12.5, p. 524). For our purposes, the important thing about these daily interventions is that they are adjustments of the size of the Fed's balance sheet. If the Fed is adding reserves, that means that the daylight expansion of credit on Fedwire doesn't have to collapse back to zero at the end of day. The Fed is taking some of that credit onto its own balance sheet overnight, or longer.[3]

[3] D. L. Thornton has generated a debate about the "liquidity effect"(p. 507) which we can understand better using the balance sheet approach.

The Fed does these daily adds usually by lending to security dealers in the repo market. We will be talking about the repo market next time. For now, suffice it to say that repos are loans, so the reserve add can be pictured as follows:

Bank + Security Dealers The Fed

A	L	A	L
+reserves	+RP loan	+RP loan	+reserves

If the Bank lends these extra reserves out, that amounts to an additional supply of credit in the Fed Funds market, which puts downward pressure on the Fed Funds rate. (On p. 47 we do the same example, but distinguish bank from security dealer.)

The Wednesday afternoon settlement to which Stigum refers is about meeting this reserve requirement, not the regular daily settlement of payment balances. Note especially how the Fed intervenes to provide the reserves demanded at the targeted FF rate. As Sternlight says (p. 506) "If you want to hang your hat on something, hang it on the average effective funds rate over the settlement period."

Anecdote

When I was at the Board of Governors in summer 1986, there was much discussion about daylight overdrafts on Fedwire. Fedwire is a Real Time Gross Settlement system where a payment from bank A to bank B is immediately good funds that bank B can use to make a payment of its own. The problem is that banks were making more payments than they had reserves during the day, with the plan that they would be receiving more payments at the end of the day. So there was in effect an intraday expansion of reserves that was not under the control of the Fed. The Fed got concerned about the apparent credit element entering here, partly because it put them in the position of lending without control, partly also because they think of reserves as money and they want the credit element to happen somewhere else in the system.

They clamped down initially by establishing credit limits, tailored to individual bank needs. Then in 1994 they started charging interest on average daily overdrafts (now 36 basis points). This has not been that effective in reducing overdrafts, and not effective at all in reducing peak aggregate overdrafts (which come late in the day and are driven by security transactions). One thought is that maybe the price is not high enough? Maybe so, but be careful. The relevant alternative to overdrafts on Fedwire are overdrafts on other payments systems, which typically cost less, even zero. (Here again we hear echoes of the Strong Rule: use the amount of credit on the Fed balance sheet to monitor and control credit more generally.)

For example, CHIPs is an alternative payment system with an even stronger credit element because payments are not good funds until net settlement at the end of the day. One way to think about what happens is that banks do business on CHIPS, taking great care to keep track of their net position and to accumulate what FF they need to clear that position at the end of the day. Note that Oct 1, 1981 Fed required that CHIPs go to same-day settlement, which in effect tightened up on the credit element here as well. And there has been periodic tightening of the

collateral requirements for overdrafts on CHIPs. The idea has been to push the credit element out of the payments system into the repo market which we'll talk about next time.

7. Repos: Postponing Settlement

A repurchase agreement is a **collateralized overnight loan, legally constructed as two security transactions one day apart**. Keeping our eye on the loan interpretation, we can call the agent who "repos securities" (sells securities for repurchase) the borrower of money, and the agent who "reverses in securities" (buys securities for resale) the lender of money. (To make things confusing, it is the latter person who is said to "do repo," meaning to invest in overnight loans.) On the first leg of the repo, money flows from the lender to the borrower and securities flow the other way. On the second leg of the repo, the securities flow back to the borrower, and somewhat more money flows back to the lender. See diagram on p. 534, and refer back to it as often as you need in order to keep straight how the repo works.

Security Dealers

In practice a security dealer is on the other side of most repo transactions. He is either borrowing money by using his securities as collateral or lending money by taking in securities as collateral. Think of the security dealers as a kind of bank. Stigum says that corporations and pension funds invest in repo (make overnight secured loans to security dealers) as a way of earning interest on their money balances, and banks repo their security holdings because it is often the cheapest way to finance them. Figure 13.9 (p. 562) shows the stylized balance sheet of a security dealer, treating borrowing of money as a liability and lending of money as an asset. This will be our convention in this course; I call it "following the money".[1] Here is a version of that figure that includes the balance sheet of the ultimate borrower (corporation A) and the ultimate lender (corporation B). Concretely we might think of corporation A as a bank that funds its security holdings with overnight repo, and corporation B as a pension fund that uses overnight repo to hold its money balances.

Bank Corporation A		Dealer		Pension Fund Corporation B	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	+reverse	+reverse	+repo loan	+repo loan	

As these transactions take place, collateral flows from left to right, from the bank through the dealer to the pension fund. And money flows from right to left, from the pension fund to the dealer to the bank. Note that I call repo a reverse when it is an asset of the dealer, in line with reporting conventions (see below). But it is important to emphasize that it is the exact same instrument, just the opposite side of the exact same instrument.

[1] Since the person borrowing money is also in effect lending securities, in principle you might think of that person as acquiring an asset, and book the transaction as a current sale of a security combined with a forward purchase. But we don't do that. Our convention is suggested by the hierarchical character of the money- credit system.

Using this balance sheet, compare the repo market to the Fed Funds market we talked about last time. We saw how Fed Funds can be understood as interbank borrowing, and can be used to channel funds from banks that have cash inflow greater than outflow (surplus banks) to banks that have cash inflow less than cash outflow (deficit banks). It should be clear now that the repo market does much the same thing as the Fed Funds market but for a much wider class of economic agents, a class that includes just about anybody that owns eligible collateral. Repo can be thus be understood as inter-corporate borrowing, and thought of as channeling funds from corporations that have cash inflow greater than outflow (Corporation B, the surplus corporation) to corporations that have cash inflow less than outflow (Corporation A, the deficit corporation).

Whereas the Fed Funds market is largely a direct or brokered market, the repo market is largely a dealer market, and the most important dealers are the primary security dealers whose balance sheets are monitored by the Fed. Why are dealers so much more important in the repo market? **By serving as the counterparty for most RP transactions, they serve to make a unified and relatively homogeneous market out of what would otherwise be extremely fragmented and heterogeneous (various counterparties, various collaterals, various maturities.)** Like Fed Funds dealers, RP dealers can be thought of as market makers.

The Nitty Gritty

The legal construction as paired security transactions is important because it means that the transaction is more symmetric than an ordinary collateralized loan would be. Think of a house mortgage for comparison. The house itself is security for the money loan, but the lender of money has no right to deliver the house to anybody else, nor even to sell it when the borrower defaults without a lengthy foreclosure process. For repo, by contrast, the lender of money does often have the right to sell the securities (or something pretty close, such as the right of rehypothecation). The legal niceties of the contract are an attempt to establish that symmetry. (They have to get around existing law that gives bankruptcy protection to borrowers and so prevents lenders from selling collateral to recover except after a lengthy legal proceeding.)

But the transaction is not completely symmetric because the price at which the security is transferred is lower than its market price. **This means that the money lender gets control over more securities than a clean purchase would allow.** Stigum refers to that difference as margin. She talks a bit about the question of margin, and suggests that it is not clear as a matter of principle who should be paying the margin, the borrower or the lender. The problem is that the price of the security could go up or down, so giving incentive for default to either the money lender or the money borrower, respectively. Stigum concludes however, "traditionally on a repo transaction, the lender of money, because it is lending the more liquid asset, receives margin" (p. 535). In other words, the asymmetry is a symptom of the money-credit hierarchy.

The way that margin works is that the price of a security that is used as repo collateral is less than its market price by an amount called the "haircut". In the book there is an example (p. 533, Figure 13.1), in which a 10 year Treasury Bond trading at 100-29/32 plus accrued interest is used as collateral for an overnight repo. The haircut is 2%, so the borrower gets only 99.228 for his collateral. The agreed overnight interest rate is 4.92%, which means that the borrower has to pay 99.2280*(1+.0492/360)=99.2416 the next day in order to repurchase his securities. The overnight interest on an approximately million dollar loan is about $135.

(Note that money market convention treats the year as having only 360 days. Also by convention, interest is not compounded, so a 3 day repo would pay back [1+.0492*(**3**/360)], not [1+.0492/360]^{3}).

Repo as Source of Funding for Securities Dealers

I began this lecture with a motivation that referred back to our discussion of Fed Funds, seeing repo dealers as another way of getting deficit agents and surplus agents together to push settlement off for another day. And I emphasized that the repo market is open to many more agents, hence a kind of generalization of the FF market. But there is more to it than that. The FF market is largely a broker market, while repo is pretty much entirely a dealer market, and that means that we have to take more into account what the dealer thinks he is doing. He does not think he is bringing deficit agents and surplus agents together. Rather, he thinks he is bringing buyers and sellers of securities together, and absorbing any mismatch on his own balance sheet.

Figure 13.3 (p. 539) shows that in general security dealers borrow more than they lend in the overnight market (repos>reverses), but lend more than they borrow in the term market (repos < reverses). Like banks, dealers seem to be in the business of borrowing short term to lend long term, but unlike banks dealers are specialists in the very short end of the money market term structure. (Note that the numbers don't balance exactly because reverses are not the only asset held by security dealers. Stigum makes clear however that reverses are a very desirable asset for dealers, and they do as much of them as they can, for the simple reason that they can repo out whatever they reverse in, and earn something on the spread. That is the "matched book" business in a nutshell.)

The figures in Figure 13.3 all come from the weekly report that primary dealers send to the New York Fed. Stigum shows one part of the report in Table 13.1 (p. 564), the dealers' net outright holdings of securities of various types. This table allows us to see the exposure of dealers to fluctuations in asset prices. In the table dealers are long Agency securities but short Treasuries, so they are picking up the spread but will lose money if that spread narrows.[2] For our purposes at the moment we are more interested in the gross holdings of securities, and how those gross holdings are financed using repo and reverses. See http://www.newyorkfed.org/banking/reportingforms/primarystats/deal.pdf and scroll down to Table IV. The following balance sheet can be constructed:

Security Dealers, Sept 2012 (billions)

Assets		Liabilities	
RP loans, overnight	879	1842	RP loans, overnight
RP loans, term	1317	932	RP loans, term
Net Assets Financed		Net Worth (Capital)	

[2] In fact the net outright holdings can be a very misleading indication of exposure because it does not include any exposure that comes from futures and other derivatives which can be substantial. Perhaps we can assume that dealers are making the same bets in their off-balance sheet portfolio as they are in their on-balance sheet portfolio?

I've added in entries for net assets financed and net worth as a gesture toward the entire balance sheet, but let's focus on the RP entries for now. Security dealers have 2.5 trillion dollars of overnight borrowing, which we can consider analogous to bank deposits. You can't spend an RP, but you can certainly spend the next day proceeds, and corporate treasurers use that feature heavily. Compared to those liabilities, their assets are noticeably longer term, even if we just look at the money market portion of the portfolio. Apparently security dealers borrow short and lend long, providing liquidity to the market in much the same way that banks do.

The Fed in the repo market, with security dealers

As economists, we are used to thinking of the Fed as engaged in open market operations, buying Tbills when it wants to increase the supply of reserves, and selling when it wants to contract. Over the course of a business cycle, it may be doing both. Secularly however it is a buyer since the economy is growing and so the supply of reserves grows. Whether secularly or cyclically, these open market operations are made at most a couple times a year.

On a daily basis, the Fed is involved in ensuring elasticity of reserves by standing ready to soak up temporary excess supply and expand to meet temporary excess demand. "Operating factors" that might require Fed intervention include changes in Treasury balances (since Treasury balances are held at the Fed, an increase in balances drains reserves from the banking system), and changes in cash holdings of the public (increased cash holding is a reserve drain). The Fed intervenes to adjust for operating factors by engaging in repo. It does repo (buys Tbills for resale, lends money) like everyone else when it wants to temporarily increase reserves. But when it want to decrease reserves, it does MSP (matched sale purchase).

MSP is like a reverse, but since a reverse is understood as borrowing money, the Fed wants to make it more symmetric. An MSP is legally two distinct sales. The new edition of Stigum suggests that the Fed is now more willing to do outright reverses, but the evidence is that it doesn't do much. The reverses that are on the Fed's balance sheet are almost all done as part of the Fed's repo pool for foreign central banks. Thus foreign central banks get interest at the repo rate on their balances at the Fed, whereas normal reserve balances pay no interest. [UPDATE: The Fed now calls this same transaction a reverse repo, and has a standing RRP facility.]

Figure 13.10 (p. 569) shows an example of the Fed's overnight repo operations. The Fed tells the primary dealers that it wants to do repo, and asks them to submit collateral and bid for the money. It accepts the best bids and does the repo. The effect is to increase reserves as follows:

Bank		Dealer		Fed	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+reserves	+deposit	+deposit	+repo	+repo	+reserves

Thus by expanding its own balance sheet, the Fed expands the balance sheet of dealers and banks as well.

More recent examples are posted daily on the website of the New York Fed, www.newyorkfed.org/markets/omo/dmm/temp.cfm. The annual Open Market Operations report at www.newyorkfed.org/markets/omo/omo2006.pdf has a wealth of summary information about these operations. One important fact is that the Desk arranged short term RPs on all but 8

business days in 2006. Basically before the crisis the Fed was in the market every day, and the size of its daily RP was large relative to total bank reserves.

All that changed with the crisis. Nowadays the Fed intervenes in this way only rarely, in the last few months mostly to test its ability to do reverse repo with the dealers, as a way of shrinking its balance sheet eventually. Here is the balance sheet describing that:

Bank		Dealer		Fed	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
-reserves	-deposit	+reverse repo -deposit			+reverse repo -reserves

You can see that the reverse works to shift its liabilities from reserves (which are high powered money) to repo (which is less high powered, especially if it is term repo), and to shift the Fed's counterparty from banks to dealers. Whether or not this is such a big deal is something we can talk about; the point right now to emphasize is understanding exactly how it all works.

Pricing

Stigum says that in general the overnight repo rate is a bit lower than the overnight Fed Funds rate, and a bit higher than the three month Treasury bill rate. Why should this be? She suggests two reasons for this, but neither is convincing. First, she observes that repo is <u>secured</u> credit whereas Fed Funds is unsecured, and concludes that 5-10 basis point differential is compensation for the higher risk involved in Fed Funds. I don't buy it. In the Fed Funds market, control of credit lines is the way that banks avoid credit risk, and they set these lines in order to ensure that they face essentially zero risk of default. No one lends 1MM overnight to gain only about $100 interest if they have any concern at all about default. It would be better simply to forego the interest, and it is easy to do that simply by foregoing the loan.

The second reason given is that there are many economic entities that <u>cannot invest</u> in Fed Funds but can invest in repo, and they might tend to push rates on repo below Fed Funds. I don't buy this one either. There are plenty of agents who can borrow at the repo rate and lend at the Fed Funds rate—your typical bank for example—so the question is why this arbitrage does not close the gap.

In my view, we are closer to the institutional facts of the matter if we think of the Fed Funds target as a kind of penalty discount rate that dealers have to pay if they are unable to meet their survival constraint by borrowing at the repo rate.[3] Dealers expand their balance sheets to the extent possible on very thin capitalization while holding essentially no cash reserves, depending instead on the repo market to raise cash as needed. If they run into trouble, (which is to say if they find themselves with <u>insufficient collateral</u> for additional repo borrowing) they rely in the first instance on their clearing banks for a dealer loan, which is priced over Fed Funds since the bank depends on the Fed Funds market to fund the loan. But that's just for the

[3] In this respect, the modern US system is very much like the classic British system of the nineteenth century that has been analyzed by Bagehot and Sayers. See my "Monetary Policy Implementation: A Microstructure Approach" (October 2006) for detailed argument from which the following is summarized.

occasional last minute mistake. More fundamentally, and more routinely, dealers can rely on the Fed itself for a repo loan priced at the Fed Funds rate, since that is the rate that the Fed is trying to establish with its daily intervention.

Think of it this way. At the morning auction, dealers bid for the money. Bids that are below the Fed Funds rate will not be attractive to the Fed. It wants to supply needed reserves, but also to retain discipline in the market by keeping reserves scarce. It does that by accepting bids that are at or above the target Fed Funds rate. For their part, dealers are willing to bid above the market repo rate, even if they think they will be financing most of their needs at the market repo rate, so long as they face any probability of having to ask for a dealer loan from their bank, which charges typically FF+50bp. So long as they get the money from the Fed for less than they could get it at their clearing bank, they are happy. The end result is the pattern we see, that the Fed funds rate tends to be above the repo rate.

Observe how thinking about the balance sheet relationships helps us make sense of the typical price relationships. We observe how the institutional mechanism of the daily auction serves to establish a premium on the best money in the system, and that premium provides an incentive for agents throughout the system to try to meet their obligations at the clearing rather than roll them over to another day. That premium shows up in the slight premium of fed funds over repo, and it also shows up in the typical premium of overnight money over longer term money, such as the three month bill. The answer to the otherwise puzzling pattern of interest rates in the money market is nothing more than the natural hierarchy of money and credit.

Above I argued that a situation of repo<FF could be understood as a particular balance between elasticity and discipline, with the Fed keeping the better money (FF) at a premium in order to establish some discipline. Analogously, it seems we could think of the opposite situation of repo>FF as the opposite balance, with the Fed keeping the better money at a discount in order to establish some elasticity. In effect the market rate of interest is the repo rate, and the official rate is the Fed Funds rate. The Fed is trying to set up incentives for banks to borrow at the Fed Funds rate and lend at the repo rate, so supporting short term credit markets more generally. The fact that the gap stands now at 15 basis points suggests that this strategy is not working very well—that is a large gap in money market terms, even if it seems small to us.

8. Eurodollars: Parallel Settlement

Eurodollars are dollar balances held by banks or bank branches outside the country, which banks hold no reserves at the Fed and consequently have no direct access to settlement through Fedwire. Instead, Eurodollar balances settle on the private CHIPS network. The center of the world Eurodollar market is London, not New York.

Stigum provides an example of how Eurodollars arise when a US company Exxon transfers a deposit from Chase NY to Citibank London as follows (p. 212):

Exxon		Chase NY		Citi NY		Citi London	
A	L						
-10 Chase + 10 Citi London		-10 reserves	-10 Exxon	+10 reserves	+ 10 Citi London	+10 Citi NY	+10 Exxon

Because Citibank London has no account at the NY Fed, it holds instead an account at Citi NY which it uses to make and receive dollar payments. It accepts the deposit from Exxon by crediting its New York account (so expanding its balance sheet), and Citi NY accepts that deposit by expanding its holding of reserves at the New York Fed. When we talk about Eurodollars, we are talking about the balance sheet of Citi London.

Stigum makes a big deal of the point that in the process of creating Eurodollars no dollars leave the country. The balance sheets make clear that she is talking narrowly about reserve balances. From the point of view of the New York Fed, all that has happened is a debit from the account of Chase NY and a credit to the account of Citi NY, which leaves total reserve balances in the system unchanged. From another point of view however, the transaction does give rise to dollar balances held outside the country that used to be held inside. If the bank accepting the deposit was Credit Lyonnais rather than a branch of Citibank, we might see the point more clearly.

Exxon		Chase NY		Citi NY		Credit Lyonnais	
A	L						
-10 Chase + 10 Credit Lyonnais		-10 reserves	-10 Exxon	+10 reserves	+ 10 Credit Lyonnais	+10 Citi NY +loans	+10 Exxon +deposits

If we confine our attention to the US, we see that there has been an expansion of the balance sheet of Citi NY but an exactly equal contraction of the balance sheet of Chase NY. However outside the US there has been an expansion of the balance sheet of Credit Lyonnais and no corresponding contraction. From a global perspective, there

has been an expansion of credit.

And that is just the beginning. Since Credit Lyonnais has reserves in NY, it can now proceed to do dollar denominated banking business, both lending and deposit taking, outside the country.

Why is there a Eurodollar market? It seems to be a combination of two reasons. First, the dollar is an important (the most important) international currency, so people outside the US find themselves needing to make and receive payments in dollars. Since the payment system is a credit system, there is need for an interbank market that links deficit and surplus agents, analogous to the Fed Funds market. In principle they could do it in NY, but that's where the second reason comes in.

Historically there were capital controls that got in the way of free movement of dollars in and out of the country, and there were regulatory controls (such as reserve requirements and FICA premiums for deposit insurance) on dollars held inside the country. The Eurodollar market sprang up to provide a necessary service outside the control of US authorities. Most of those barriers have by now been lifted, but the market survives as a separate entity.

But the Eurodollar market is more than a dollar payment system; **it is also the world funding market**. Wherever you may be in the world, if you need to raise money for some domestic project that you are unable to finance domestically, you are going to be borrowing in dollars. The picture to have in mind is these foreign banks taking deposits and making loans, essentially acting as money dealers in the global dollar market.

Borrower Money Dealer Depositor

A	L	A	L	A	L
	Dollar loan	Dollar loan	Dollar deposit	Dollar deposit	

Apparently even now there is some advantage in having the institutions of the <u>international</u> dollar separate from the institutions of the <u>domestic</u> dollar. In fact, the overnight Eurodollar market has recently been larger in volume than the overnight Fed Funds market. Which is the tail and which is the dog? Today, it is still the domestic dollar that dominates, but stay tuned (p. 860).

One issue to flag right away here is the question of monetary policy. The Fed is concerned about employment and inflation within the United States, and monetary policy is an attempt to influence those conditions. To do this, the Fed focuses attention on the domestic money supply and domestic interest rates, leaving largely out of consideration the international money supply in the Eurodollar market. The idea is that these balances are held by foreigners and so may influence their behavior, but not the behavior of domestic consumers and businesses. Increasingly this is a difficult abstraction to defend, since important

entities are global.

Eurodollar as interbank market

The fundamental reason for the Eurodollar market seems to be that foreign banks have customers who wish to hold dollar balances or take out dollar loans from them. This customer-led demand causes some of the banks to have a natural surplus position (more dollar deposits than loans) and other banks to have a natural deficit position (more dollar loans than deposits). They could each resolve the imbalance by doing business with some US bank, but it seems easier all around for them to do business with each other, with the surplus banks lending to the deficit banks. That's what is happening in the interbank market, and the London Interbank Offer Rate, or LIBOR, is the rate of interest charged in that market. Just so:

Deficit agent Citi London			Surplus agent Credit Lyonnais	
Assets	Liabilities		Assets	Liabilities
US$ customer loans	Eurodollar deposit, Credit		Eurodollar deposit, Citi	US$ customer deposits

Both Citi and Credit seek "matched book", which means that their dollar liabilities are the same as their dollar assets, since that protects them from any change in the value of the dollar. Because their natural customer positions offset, they can achieve matched book by trading with each other in the interbank market. This is an unsecured interbank borrowing like Fed Funds, and like in the Fed Funds market banks control counterparty risk by controlling their credit lines to one another. (Tiering of prices is apparently not about credit risk, but more about market organization.)

It is not enough however simply to match total assets and total liabilities. Especially so when we have in mind the special vulnerability of the Eurodollar market to liquidity problems because of its lack of access to the Fed's balance sheet. The liquidity of the Eurodollar market ultimately depends on the ability and willingness of the New York banks to provide liquidity as needed by taking the problems of the Eurodollar market onto their own balance sheets. **Uncertainty about that ability and willingness makes the Eurodollar banks take care to line up the time pattern of cash inflows and cash outflows, in order to minimize their need to use reserves.** (Recall the example in Lecture 4.) That's the reason that Eurodollar deposits are made to specific dates, with no early withdrawal permitted (and no negotiability of the deposit either). The bank wants to know precisely when the cash outflow will happen, in order to prepare to meet it. But what if those dates don't line up?

Balance Sheet Approach to FRA

The most important instrument for that purpose is the Forward Rate Agreement, which I will now proceed to explain. It is an off-balance sheet

instrument, essentially a side bet on the value of LIBOR at some date in the future, but we understand its importance better if we imagine what it would look like as an on-balance sheet instrument. **This mode of analysis, which we'll be using for all derivatives, is presented here for the first time, so it is worth your while to master it.**

Suppose that two months from now Bank X expects to be making a 3-month US$ loan, and two months from now Bank Y expects to be receiving a 3-month US$ deposit. (Perhaps these expectations arise from existing loans and deposits that expire in two months but which the customer is likely to renew, see p. 832). They can help each other line up cash flows in time by swapping IOUs today as follows:

Bank X (forward borrower) Bank Y (forward lender)

2 month deposit, F%	5 month deposit, F%	5 month deposit, F%	2 month deposit, F%

At the moment, suppose these two IOUs have exactly the same value, so no money changes hands. Also, over the next two months the interest rate is the same, so again no money changes hands. But two months from now one deposit matures, which involves a cash flow from Bank Y to Bank X. Five months from now the other deposit matures, which involves a cash flow from Bank X to Bank Y.

Observe that this pattern of cash flows is exactly what the banks need in order to hedge their natural positions. Two months from now Bank Y gets a customer deposit,
and pays it to Bank X, which uses it to make a customer loan. Five months from now the customer repays the loan to Bank X, which sends the money to Bank Y. At the end of
the day, it is the three month deposit that funds the 3 month loan, but the on-balance sheet IOUs allowed the banks to set it up ahead of time.

Bank X Bank Y

3 month loan	3 month deposit, F%	3 month deposit, F%	3 month deposit

We could do all of this with an explicit swap of IOUs on balance sheet, but we don't. Banks are always looking to economize on balance sheet capacity, and there is a more efficient way to get the same job done. For a while banks did something almost like a swap of IOUs, called a forward forward, where Bank Y agreed in advance to place a 3 month deposit in Bank X at rate F% two months from now. Clearly, this forward forward structure just nets out the first two months of our ideal on-balance sheet swap of IOUs.

The Forward Rate Agreement goes even farther by netting out the principal that

is paid from Y to X in two months, and back again the other way in five months. Instead what is paid is the difference between LIBOR and F% two months from now. (See p.
832 for a worked example.) It is because these principal payments have been eliminated that FRA agreements can be off-balance sheet agreements. Bank X plans to fund the loan by borrowing at LIBOR, and Bank Y plans to invest its deposit by lending at LIBOR.

But by engaging in a FRA that pays the difference between LIBOR and F%, they both manage to lock in an interest rate of F% for the funds in question. Bank X doesn't care how high LIBOR might go, and Bank Y doesn't care how low LIBOR might go, <u>just so long as the liquidity is there so they can raise and invest the principal amounts as needed</u>. (I underline this last bit, since it is a key assumption, almost always satisfied but likely to fail when you need it most.)

Bank X			Bank Y
3 month loan	3 mo deposit, LIBOR% 3 mo FRA, (F%-LIBOR%)	3 mo deposit, LIBOR% 3 mo FRA, (F%-LIBOR%)	3 mo deposit

Forward Interest Parity

Turn now to the question of what rate F% is appropriate for this transaction. Whatever rate it is, the effect is to lock in the cost of funding the future loan, and the benefit from accepting the future deposit, so both banks have an interest in getting the number right. Bank X would like a low number, and Bank Y would like a high number. What number will be chosen?

The answer depends on something called Forward Interest Parity. Suppose there is a **market** rate for 2-month deposits (2 month LIBOR) and a **market** rate for 5-month deposits (5 month LIBOR). If you were long one of these and short the other, you would have locked in a 3 month borrowing/lending rate for two months from now, and the rate you lock in can be calculated as follows:

FIP: $[1+R(0,N)][1+F(N,T)] = [1+R(0,T)]$

The notation is that R(0,N) is the rate of interest on a deposit extending from now to N (2 months in our example), R(0,T) is the rate of interest on a deposit extending from now to T (5 months in our example) and F(N,T) is the implied forward rate on a three month deposit two months from now.[1]

[1] I follow the notation of Geoffrey Poitras (2002).

FIP is an **arbitrage condition**. What that means is that if the actual forward rate on a FRA were any different from the implied forward rate, there would be opportunities for riskless profit by lending at the higher rate and borrowing at the lower rate. Because of this, in practice the market forward rate tends to be very close to the implied forward rate. This is the rate that will be chosen for our interbank transaction.

Eurocurrencies and FX swaps

In addition to the Eurodollar market that trades at US$ LIBOR, there are also Euro markets in other major currencies such as the euro, the pound sterling, and the yen. The existence of these other markets provides yet another degree of freedom for bankers looking to line up cash inflows and outflows. (If the desired timing is not available one currency, perhaps it is available in another currency.) However, if we go that route, then have a different problem of mismatch in our currency exposure. That is the problem that gives rise to the foreign exchange swap, which permits banks to trade out of any currency mismatch they may have.

The book talks about a Euroyen swap. Let us think about what that swap would look like as an on-balance sheet operation, and use that to help us understand what is going on. Suppose a bank commits to lend 6-month Euroyen and now needs to fund that loan (p. 855). Suppose that funds are available in 6-month Eurodollars, so now the bank has a problem of currency mismatch. Six months from now it will be receiving yen (from the loan) and paying out dollars (on the deposit). How can it fix that mismatch?

It can solve that problem with the following balance sheet operation:

Bank		Counterparty	
Assets	Liabilities	Assets	Liabilities
6 month loan, Euroyen	6 month deposit Eurodollars		
6 month deposit, R Eurodollars	6 month deposit, R* Euroyen	6 month deposit, R* Euroyen	6 month deposit, R Eurodollars

Observe that this balance sheet operation leaves our bank with completely matched book, both over time and across currencies. Six months from now it will be receiving yen (from the loan) and paying out yen (on the swap), and receiving dollars (on the swap) and paying out dollars (on the deposit).

Let us now suppose that this swap of IOUs has zero value at inception and investigate the details that would make it so. First step is to open accounts that have the same current value according to the current exchange rate, say 115 yen for every dollar, S(0). (Note well that I am defining the exchange rate as yen/dollar, so a rise in S is an appreciation of the dollar and a depreciation of the yen.) Six months from now we'll close the accounts and there will be a different exchange rate, but we can fix a notional rate F(T) today that makes

the present value of the two accounts exactly the same. Here's how.

Today there is some prevailing interest rate for Euroyen, R*, and some prevailing interest rate for Eurodollars, R. So if we open a Euroyen account with 115 yen it will grow by the rate R*, and our $1 Eurodollar account will grow by the rate R. These accounts would have the same value six months from now if the exchange rate changed by exactly the right amount to counter the interest rate differential. The future exchange rate that makes these two accounts have the same value today is the exchange rate F(T) that satisfies Covered Interest Parity

CIP: $[1+R^*(0,T)]S(0) = [1+R(0,T)]F(T)$

The left hand side is the yen value of the yen account at maturity, and the right hand side is the yen value of the dollar account at maturity. Like FIP, CIP is an arbitrage condition, so we expect the actual market forward exchange rate to be very close to the forward exchange rate that is implied by this formula, and it usually is. [UPDATE: Not so since the financial crisis, but we postpone discussion until Lecture 16 on foreign exchange.]

The cash flows involved in our notional swap of IOUs are as follows:

Time 0: $1 from Bank to counterparty, S(0) yen from counterparty to bank Time T: [1+R*]S(0) yen from bank to counterparty, [1+R] dollars from counterparty to bank

If the actual exchange rate realized at time T $S(T) = F(T)$ $[=(1+R^*)S(0)/(1+R)]$ then the cash flows at time T are exactly offsetting in value terms. But if $S(T) > F(T)$ then a dollar buys more yen, so the cash inflow to the bank (from the swap) is worth more than the cash outflow (from the swap). And if $S(T) < F(T)$ then a dollar buys less yen, so the cash inflow to the bank is worth less than the cash outflow.

Now we can see how to do the same thing more efficiently with a swap. A swap contract is nothing more than an off balance sheet way of achieving the exact same net cash flows as the on-balance sheet swap of IOUs. At time 0, the two parties swap yen for dollars at S(0) and agree to swap back again at F(T). (In the case of the FX swap, these principal payments cannot be netted because they are in different currencies, and that is different from the forward interest swap.)

The effect of the FX swap is to shield the bank from any fluctuation in the yen exchange rate. If the yen falls in value, he loses on his Euroyen loan but gains on the offsetting swap. If the yen rises in value, he gains on his Euroyen loan but loses on the offsetting swap. The bank does not care what happens to the yen exchange rate.

EH and UIP—Two relationships you might think should be true but aren't

The Expectations Hypothesis of the term structure suggests that the forward rate should be equal to the expected future spot rate, and this sounds like a reasonable theory. In practice however there tends to be an inequality.

EH: $F(N,T) = ER(N,T)$

Violation of EH: $F(N,T) > ER(N,T)$

What this means is that Bank X typically would have done better by not engaging in the FRA and simply borrowing funds as needed in the spot market. That Bank X is willing to lose this money tells us that it is paying for insurance. Note that it is the borrower of money who pays, not the lender. This violation of EH is another symptom of the hierarchy of money and credit.

Similarly, since the forward exchange rate can be locked in today, we might expect it to bear some relationship to the expected future exchange rate. Uncovered Interest Parity suggests that they should be equal, and this sounds eminently reasonable.

UIP: $F(T) = ES(T)$, or, using the CIP relationship

$$(1+R^*)/(1+R) = ES(T)/S(0)$$

Suppose $R^*<R$, so the yen is a relatively low yielding currency. UIP says that we must expect the dollar to depreciate against the yen (yen to appreciate against the dollar) by just enough to make both currencies have the same yield, at least ex ante.

Nice theory, but unfortunately it seems not to hold up very well in the real world. Study after study finds

Violation of UIP: If $F(T) < S(0)$, [i.e. $R>R^*$], then $F(T) < ES(T)$

That is to say, usually the low yielding currency does not appreciate by enough to give it the same yield as the high yielding currency. (In fact, typically the low yielding currency depreciates, contrary to UIP prediction that it will appreciate.)

9. The World that Bagehot Knew

Review

Up to now, we have been looking at the money market as a mechanism for making a decentralized payments system operate as close as possible to the ideal one-big-bank system.
We started with the observation that, in principle, all payments could be made as book entries on the balance sheet of a single, central bank. This is an ideal, and we can view the institutional arrangements of our actual payments system as an attempt to get as close as possible to that ideal.

From this point of view, in a decentralized payments system, the central bank emerges as a kind of clearinghouse for facilitating interbank payments by netting offsetting payments. The motivation for such an institution is to promote elasticity in the payments system, so that any two people who want to make a trade can in fact do so, while at the same time imposing discipline through, for example, periodic net settlement.

Farther down the hierarchy, we can think of the money market as a system for facilitating interbank payments by enabling surplus and deficit units to find each other easily. Focusing narrowly on banks, we saw how the Fed Funds market enables banks experiencing temporary net cash outflows to borrow needed reserve funds from the banks receiving those outflows as inflows. Thinking about security dealers, we saw how the repo market serves much the same function of allowing dealers with cash deficit to borrow from dealers with cash surplus.

We saw also how, in both cases, there is a well-developed backstop in case surplus and deficit units fail to find one another. The Fed's discount window (and hence the Fed balance sheet) operates as a backstop for the Fed Funds market, and lines of credit at the dealers' clearing banks operate as a backstop for the repo market.

Outside the U.S., the same principles apply in the Eurodollar market. Because foreign banks do not have access to the Fed discount window, they have evolved different ways of making sure that cash inflows match cash outflows, mainly by matching the two in advance using Forward Rate Agreements, and then relying on the overnight Eurodollar market to handle any remaining mismatch. In a similar way, foreign banks have developed mechanisms for matching up inflows and outflows in every currency in which they deal, mainly by using forward currency markets and currency swaps.

By focusing on the payments system, we were able to see the "survival constraint" in action. From this point of view, the significant conceptual distinction was between **money** (=final settlement, good funds) and **credit** (=postpone settlement, promise to pay). If you cannot meet your obligations at the final settlement (clearing) then you must postpone (borrow) or else fail (not survive). Whatever level in the system we are looking at, a payments perspective suggests that the difference between money and credit is all about **time**.

A key feature of the system is its hierarchical character, insofar as what counts as money and what counts as credit depends on what level of the hierarchical system we are considering. That is why the next higher level can operate as a backstop for credit markets at every level. The availability of that backstop is the place where the balance between elasticity and discipline is managed. Smooth operation of the payments system requires elasticity, so that shortage of means of payment does not

get in the way of mutually beneficial trade. But it also requires discipline, so final settlement is only postponed, not put off forever.

We see the tension between these two principles in the simultaneous elasticity of credit and inelasticity of money. We see the tension also in institutions like continuous real time gross settlement and control over daylight overdrafts on Fedwire, limits on overdrafts and daily net settlement on CHIPS, combined with looser limits in the repo and Eurodollar markets. **The effect is to protect the core of the payments system from potential credit problems (excess elasticity), and to protect the core of the credit system from potential money shortage problems (excess discipline).**

In terms of our four prices of money, we have so far been emphasizing only one, par, which is the price of deposit money in terms of currency (or reserve money). We have referred to the interest rate as the price of delaying settlement, but we haven't talked very much about what determines that rate. In our discussion, the clearinghouse was presumed to receive no interest on its assets (gold) and pay none on its standing liabilities (CH certificates). In times of crisis, the clearinghouse was presumed to charge an exceptionally high rate of interest (6%) on loans to members, and to pass that high rate of interest on to members who were willing to hold clearinghouse loan certificates. We see here a primitive kind of lender of last resort function emerging organically from business practice. But such a rate of interest bears no close relation to the rate of interest in normal times.

Preview

The next section of the course approaches the money market from a new point of view. Instead of thinking of credit instruments as forms of delayed payment (different times), we think of them as securities that have a current **price** quoted in terms of money (different prices). The price we will focus most of our attention on is the rate of interest, which is the price of money today in terms of money tomorrow. Furthermore, instead of focusing on the various credit instruments themselves (Fed Funds, repo, Eurodollars), we focus now on the institutions that make "liquid" markets in these instruments. The most significant idea will be that **banks are a kind of security dealer, a market maker in money**. We won't be able to understand that idea fully until next lecture, since first we have to understand what is meant by a security dealer, but it will help you to understand where we are going with this.

In a way, we are reversing the emphasis of Stigum. She says: "the dealer takes in securities on one side at one rate and hangs them out on the other side at a slightly more favorable (lower) rate; or to put it the other way around, the dealer borrows money from his repo customers at one rate and lends it to his reverse customers as a slightly higher rate. In doing so, the dealer is acting like a bank, and dealers know this well" (432). She emphasizes that dealers are like banks, meaning they are a kind of financial intermediary between ultimate borrowers and lenders. I want to emphasize that banks are like dealers, meaning a kind of market maker. And not only are individual banks like dealers, but also the central bank is like a dealer.

These hierarchical layers of market-making knit together the layers of the hierarchy of money. The ordinary market participant's impression that the difference between instruments is a matter of quantity (price) not quality (moneyness) is created by this system of market makers. The hierarchy is still there behind the scenes, and occasionally springs out onto the stage during crises, but most of the time we don't see it.

In this next section of the course, we will be focusing on the money rate of interest as the price of money. This new focus means that we now approach the emergence of central banking also from another angle. We think of banks as directly managing the relationship between payment and promises to pay, which is to say the relationship between money today and money tomorrow, which is to say the rate of interest. Just as in the last few lectures we pursued an analogy between a central bank and a clearinghouse, now we pursue a different analogy between **a central bank and a security dealer**. We begin with 19th century British banking where that analogy can be seen in a simpler form than in modern banking.

The Discount Mechanism

Let us suppose that firm A buys goods wholesale from Firm B. Instead of paying immediately, he promises to pay in 30 days. The tangible evidence of such a promise is a "bill of exchange". Let us suppose further that firm B does not wish to hold the bill, because he needs to make payments of his own (perhaps to his workers, or to repay maturing bills). So he takes the bill to a bank which "discounts" it, paying out notes less than the full face value of the bill.[1] The difference is interest which the bank will earn by holding the bill to maturity. Let us suppose further that the bank buys the bill with notes that it was holding in reserve, so that its reserves fall. Then we have the following:

Firm A		Firm B		Bank	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+goods	+bill of exch.	-goods +bill of exch. -bill of exch. +notes		+bill of exch.	

Note what has happened. The balance sheet of Firm A has expanded on both sides, while Firm B and the Bank have merely changed the form in which they hold their assets. If all goes well, Firm A manages to sell the goods retail, receiving notes in return which he then uses to pay back what he owes to the bank, so shrinking his balance sheet again.

Firm A		Firm B		Bank	
-goods +notes -notes	-bill of exch.			-bill of exch.	

[1] "discount" means sale at a price less than face value. Historically banks also "accepted" bills without discounting them, which means they guaranteed payment. This acceptance is a kind of credit default swap, which we will talk about in a later lecture.

So the balance sheet of Firm A shrinks back down again, and the note reserve of the bank is restored. (For simplicity in the balance sheets, I abstract from any profit firm A might have earned, as well as from the interest the bank has earned by holding the bill to maturity. In fact the notes paid in are more than the notes paid out, by the rate of interest.) Firm A financed its purchase by borrowing first from Firm B, who passed on the promise to the Bank. The result was that Firm B got paid for the goods before Firm A made the payment.

The way we did the balance sheets, that payment came from notes that the Bank held in reserve. I did it that way in order to make clear the cost to the Bank of helping Firm B receive a payment before Firm A makes a payment, namely a reduction in reserves relative to total assets, and hence increased vulnerability to demands for payment. The bank opens itself to that vulnerability in order to earn interest on the bill it discounts. This vulnerability is what we mean by liquidity risk. Concretely, with fewer notes on hand, the bank is more vulnerable to an unexpected withdrawal of its deposit liabilities.

Alternatively, the bank could discount by expanding its deposits. This too increases liquidity vulnerability since the new deposits may well be withdrawn, but probably less so than paying out notes right away.

Firm A		Firm B		Bank	
A	L	A	L	A	L
		-bill +deposit		+bill	+deposit

One other possibility is also worth noting, which is that the firm does not discount the bill but rather only "accepts" it, which means committing to pay the bill at maturity in case Firm A does not pay. Acceptance is thus a kind of guarantee, a contingent liability of the Bank and a contingent asset of Firm B. Firm B pays for this guarantee, but in return gets an interest bearing asset that he can more readily liquidate if it turns out he needs cash before maturity.

Firm A		Firm B		Bank	
A	L	A	L	A	L
		+acceptance			+acceptance

Both of these mechanisms operate to economize on scarce notes (money). In the first, the bank finds a way to discount without paying out notes. In the second, Firm B finds a way to ensure access to future cash without demanding any present cash.

The Discount Rate

Now suppose that the bank is doing business not just with Firm A and B but with lots of other firms, and suppose there are lots of other banks too. On any given day it is both extending new

discounts, and receiving payment from maturing discounts, so notes are flowing in and out. The whole secret of successful banking of this type is to match these inflows and outflows while keeping (non-earning) note balances as low as possible.

In order to do this, the one thing the bank can control is the discount rate. If it finds that requests for discount are exceeding its available note capacity, it can lower the price it is willing to pay (which means raising the discount interest rate). The effect should be to discourage new discounts relative to repayment of old ones, and hence a shift toward net inflow of notes.

r↑ discourages discount, contraction of credit at given bank
r↓ encourages discount, expansion of credit at given bank

Every bank is doing this. Each one is moving its own discount rate up and down in order to achieve the desired net flow of paying assets. If any bank finds itself with too many notes, it may lower the discount rate in order to attract new discounts. If it finds itself with too few notes, it may raise the discount rate in order to discourage new discounts. In the system as a whole multiple rates encourage holders of bills to seek out one bank or another.

Also, banks themselves may borrow and lend notes to and from one another by using "rediscount". A bank finding itself with too many discounts and too few notes can rediscount some of the excess to another bank in the opposite position. The significant point about this, for our purposes, is that the **discount banks are acting in effect as dealers, quoting a two-way market for bills of exchange and meeting the resulting flow of demand and supply by altering their own inventories and balance sheets**. Individual dealers set their own prices, with a view to managing their own inventories, maximizing profits and controlling exposure to liquidity risk. But competition and arbitrage ensures that the multiple rates are coordinated around a single rate we'll call the market rate of interest.

Given what we've said about how banks adjust their discount rates, it follows that the marketwide rate of interest fluctuates depending on the marketwide balance between cash inflows from maturing bills and cash outflows from new discounts. When discount pressure is high, or when firms are having difficulty repaying, the market rate of interest rises.

Bank Rate

Now let us think about the role of the central bank. Ever since Peel's Act, the Bank of England was divided into two departments:

Bank of England

Assets	Liabilities
Issue Department Gold Government Debt	Notes
Banking Department Notes Discounts Advances	Deposits

The effect of this division was to fix the note issue in aggregate, which is a source of discipline. The elasticity in the system comes from the willingness of the Bank of England to hold more note reserves than it really needs, and to make them available to other banks through the discount mechanism. A further source is the elasticity of deposits, which other banks are able to use as reserves when notes are particularly scarce. (Advances are loans, generally against good security as collateral.)

So far we have been implicitly thinking of the banking problem as merely involving allocation of a fixed quantity of notes. Fluctuating market rates create a profit incentive for the banking system to expand and contract its overall balance sheet, by holding fewer notes when credit demand is higher. But it also creates a profit incentive for whoever issues the notes to change the quantity. So a profit-seeking central bank would respond to a period of general credit expansion as follows:

Banking System		Central Bank	
Assets	Liabilities	Assets	Liabilities
-bills +notes +deposits		+bills -notes	 +deposits

Just as the individual banks were encouraged to take on liquidity risk by the prospect of profit, so too would be a profit-seeking central bank. As in the system with a fixed quantity of notes, there will be a tendency for all discount rates to converge on a single rate.

But notice that whenever there is scarcity of note reserve (i.e. "money is tight"), the central bank is in a position to affect the interest rate as a matter of its own policy. It may take the view that there are sufficient notes outstanding already, and set its rate sufficiently high that net demand for rediscount is minimal. Or it may set a lower rate, and allow its balance sheet to expand. In fact it may use the rediscount rate to control the rate of that expansion. Contrariwise, it may take the view that there are excessive notes outstanding, and set its rate so high that net demand for rediscount is negative, so contracting its balance sheet, and it may use its discount rate to control the rate of contraction.

Bank rate thus moves as a policy variable, while the market rate moves as a market price.

The Art of Central Banking

Here we have the origin of monetary policy. The question of course is how the central bank decides what is the right level of bank rate. The beginning of wisdom, in this respect, is the realization that narrow profit-maximization is not obviously the right policy for the central bank. Rather, the ability to relax the survival constraint for banks lower down in the hierarchy brings with it responsibility for managing the balance between discipline and elasticity, choosing when to relax the constraint and also when to tighten it. The point to hold on to here is that a central bank is at a level in the hierarchy above other banks.

In times of crisis, bills are not repaid on time so banks face a reduction in cash inflow. Unless they can replace that cash inflow by discounting at the central bank, they will be forced (the survival constraint) to curtail their own discounts, and the result will be a general credit contraction and perhaps a cascade of payment failures. Against that possibility, the central bank holds excess reserves for the system as a whole, so that in times of crisis it can "lend freely but at a high rate". This is the Bagehot Principle established by Walter Bagehot in his Lombard Street (1873). (Bagehot called this "lender of last resort", but I will call it "dealer of last resort" in order to fix ideas that will be more useful when we extend our analysis to modern conditions.)

In the face of an **internal drain**, the job of the central bank is to lend out its own note reserves "to the last farthing", and even possibly to create additional deposits as a substitute for notes. The reason for the "high rate" is to make sure that banks asking for accommodation really need it, and to provide incentive for them to pay back as soon as conditions return to normal. See the balance between elasticity (lending freely) and discipline (high rate)?

The central bank can help with an internal drain because it stands above other banks, but the same is not true of an **external drain**. The central bank is not above other central banks. Suppose the demand for discounts at the central bank is coming from people who really want gold, the international money, so they convert the notes they receive into gold and take gold out of the country. Then the central bank has a problem that it cannot fix by paying out freely, much less by expanding its own balance sheet. People don't want notes, and they don't want deposits, they want gold, the international reserve.

Young is at great pains to describe the banking department of the Bank of England as essentially a credit operation built on an extremely small foundation of gold reserves, so that any incipient gold outflow calls for immediate action. Because of the enormous volume in the bill market, there are always bills maturing and new bills being offered. All one has to do to forestall a gold drain is to make new borrowing slightly less attractive by raising the rate of interest, which is the same thing as lowering the price of bills. Then maturing bills are not replaced but rather paid and the resulting gold inflow stems the tide.

Here we see the central bank operating according to the very same principles we laid out for banks lower in the system:

r↑ discourages discount, contraction of credit at given bank
r↓ encourages discount, expansion of credit at given bank

This is the economics of Walter Bagehot, focusing on the discount rate as defense of the rate of exchange under the gold standard. When the Bank of England tightens, it contracts not only its own balance sheet but also the balance sheets of all other banks below it, which tightens credit in the entire country. Thus, to some extent, it could be said that domestic credit policy was subordinated to the dictates of the gold standard, at least whenever those dictates were binding.

This distinction between external drain and internal drain is thus an early attempt to find the appropriate balance between discipline (external) and elasticity (internal). The distinction makes a certain practical sense in a gold standard world since if the central bank does not defend the gold parity it is not clear who else would. But from a larger analytical perspective, the distinction is a lot muddier than it appears on first sight. If we take the perspective of the world as a whole, then gold flows between countries have no more effect on the aggregate world gold stock than internal flows have on the domestic gold stock. If we accommodate the latter, then why don't we accommodate the former?

Practically of course the reason is that such accommodation would require cooperation between central banks—the French central bank would have to lend back the gold that was flowing out of England. But that is only a practical problem, and it is conceivable that evolution of central banking would provide an institutional solution. There is no necessity for the scarcity of gold to constrain the system as a whole. The system does however require some overall disciplining factor.

10. Dealers: Liquid Security Markets

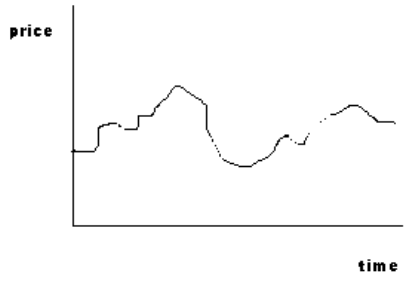
I said last time that the focus of the next section of the course will be on how different financial institutions make liquid markets that resolve the differences between qualitatively different kinds of credit into mere quantitative differences of price.

Liquid Markets and Inventories

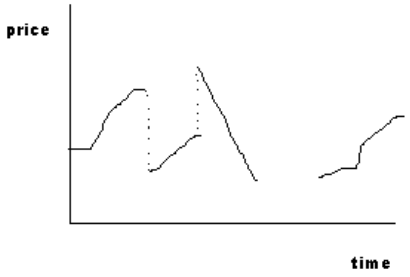
A **liquid market** is "one in which an individual transaction does not disrupt the continuity of the market". More specifically, it is a market in which you can buy and sell (1) quickly, (2) in volume, (3) without moving the price much. This feature of markets is absolutely crucial for the smooth operation of our economy, so crucial that it tends to be taken for granted. All of microeconomics revolves around the idea that suppliers and demanders are trying to find the optimal supplies and demands **given** the market price. They never consider whether they will actually be able to complete desired trades at that price.

One way of understanding the concept of liquid market is therefore as continuity in the time path of market price. There are no jumps or gaps as price evolves over time. The question is, what is the institutional basis of a liquid market?

Continuous Liquid Market



Discontinuous Illiquid market



Consider an example (inspired by Hicks, Market Theory of Money). When I pass the Westside market in the morning there is almost no one in the store, but the shelves are all stocked. When I pass the Westside market in the evening the whole world seems to be in there, and the shelves are in some places seriously depleted. And yet, despite these intraday fluctuations in the flow of demand, the price of the various goods remains the same, and we would be seriously surprised if it did not. The prices of goods fluctuate over time, across days and weeks and years, but not within the day. The Westside market is a liquid market.

My concern at the moment is not so much with why merchants find it useful and profit-maximizing to make liquid markets, as with how merchants manage to do it. A moment's thought reveals that the secret is inventories. Merchants take delivery of a large quantity of a particular good, set the price so as to ensure a profit on the load, and then proceed to supply individual demanders from the inventory until it runs sufficiently low that they make another order.

The answer is simple, but it is also deep, because when you think a little more you realize that Westside chooses its inventories on the assumption that it can get resupply and that resupply depends on inventories somewhere else in the system, and so forth all the way back to the factory or the original producer. The **continuity** of the market price for a particular good thus seems to depend on a hierarchical structure of inventories. Even so-called "just-in-time" production is best viewed as a very careful adjustment of the flow of ultimate supply to ultimate demand in order to minimize inventories. We are amazed at the Japanese just-in-time inventory system because this is really very difficult to do, but it does not eliminate inventories. Rather it pushes the ultimate inventory back before production into inventories of the parts and components held by suppliers.

Security Dealers and Market Makers

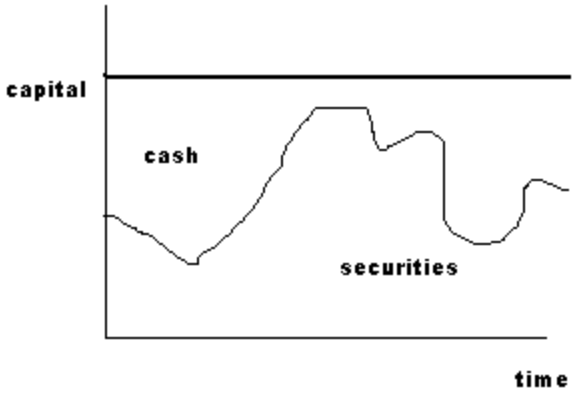
A security dealer is somewhat like Westside market, more so than might first appear. A security dealer who acts as an agent in a primary issue of corporate securities buys the securities from the corporate clients at one price and then sells them at a higher price to its retail clients. That's exactly what Westside market does. More generally, however, we are interested in two-sided dealers, who offer to sell retail but also to buy retail. Security dealers quote two prices--a lower **bid** (to buy) and a higher **ask** or offer (to sell). (In money markets prices are quoted as yields, so the bid is a higher number than the ask.) It is as if the Westside market was willing to buy mangoes from its retail customers, as well as sell them.

In the securities market both bid and ask prices are retail prices, and wholesale (interdealer) prices are within the bid-ask spread. The one sided dealer (like Westside) only sells, so it only needs an inventory of the good that it sells. The two sided dealer also buys, and so it also needs an inventory of cash. We begin therefore by thinking of the dealer holding inventories of both cash and securities in order to be prepared to fulfill when bids are "hit" and offers are "lifted".

Hypothetical Dealer Balance Sheet (showing inventories)

Assets	Liabilities
securities	capital
cash	

The consequence from the point of view of retail trade is a liquid market, meaning the ability to buy or sell without moving the price. My hypothetical dealer has inventories of cash and securities, and uses those inventories to absorb fluctuations in demand and supply. Any increase in one inventory is matched by a decrease in another. The result is that imbalances in the time pattern of demand and supply show up as balance sheet changes, not price changes.

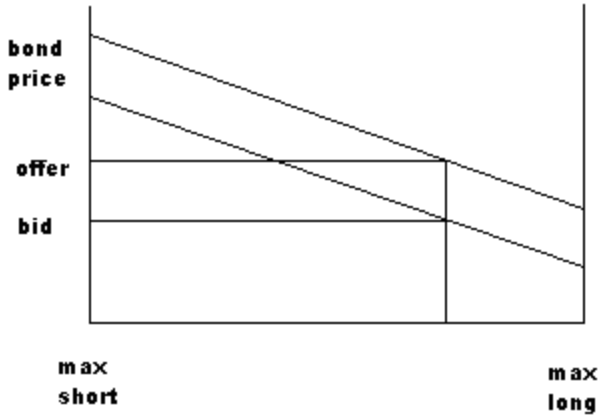


One way to appreciate what dealers do is to think about what would happen without them. Imbalances in supply and demand would cause prices to jump around, or force ultimate suppliers and demanders to wait for the opposite side. The security market would be like the housing market in which individual buyers have to find individual sellers and negotiate individual deals.

Instead, in the securities market, buyers and sellers only have to find a dealer, and the dealer takes the opposite side of their trade. The dealer makes liquid markets by buying when there is excess selling pressure, and selling when there is excess buying pressure. So he is buying at a price higher than would otherwise be (higher than if excess supply had to drive prices down far enough to attract demand), and selling at a price lower than would otherwise be (lower than if excess demand had to drive prices up far enough to attract supply). He has to pay attention to these prices if he is going to survive, and make sure that he is not paying too much or selling for too little. Dealers thus operate to smooth prices as well as to make continuous markets out of discontinuous order flow.

Economics of the Dealer Function

Now, the fundamental value of securities can and does change over time, and this is the major source of risk in the dealing business: "inventory risk". Also there is the risk that a customer might know more than the dealer, so the dealer will be buying at too high a price and selling at too low a price: "adverse selection risk". The Bid-ask spread compensates somewhat for this risk, and in times of high volatility (and for securities that have especially high volatility) we see dealers widen the bid-ask spread. Equally important, however, is the dealer's ability to **change the price as inventories change**. Treynor offers the following model of the dealer function (my picture is a stripped down version of his, so refer to his for the full model):



Here the offer (or ask, or selling price) is above the bid (buying price), and the spread means that the dealer is always buying low and selling high. The level of both prices also changes with the inventory. By lowering the price at which he is willing to add to a large long position, he protects himself from the risk that price may fall. By raising the price at which he is willing to add to a large short position, he protects himself from the risk that price may rise.

Three key pieces make up this model. First, the position limits of the dealer. He has limited capital and limited credit (from his clearing bank) and perhaps also limited taste for risk. We capture all these in the position limits, max long and max short. Second, the "outside spread", which is the price at which value based traders (the ultimate suppliers and demanders) are willing to buy and sell. This is a wide spread, maybe 20% below and 20% above fundamental value as estimated by the value based trader (VBT). When the dealer hits his position limits, the VBT becomes market maker of last resort. Third, the volatility of price and the prevalence of adverse selection risk, which influences the inside spread.

We can think of the dealer as supplying liquidity because he is offering the option to trade. If you want to trade he is willing to trade at the quoted prices, either way, but if you don't want to trade that's okay too. It's up to you, not the dealer. He supplies liquidity and you demand it, whether you want to sell or buy.[1]

[1] In fact, anyone who offers to trade, by placing a "limit order" is supplying liquidity to the market. Anyone can be a dealer in this respect, although it takes a lot of work and attention to make money at it. (Amateur day traders often supply liquidity until they run out of money.) This way of conceptualizing liquidity is developed in more detail in Larry Harris, Trading and Exchanges: Market Microstructure for Practitioners (2003).

Leverage

Actual dealers differ from my hypothetical dealer in one important respect. They actually hold almost no inventories of either cash or securities. (The balance sheet I showed you is actually more like the balance sheet of a value based trader who makes the outside spread.) Competition among dealers forces them to offer a very tight "inside spread", so the profit on any one trade is very small. They compensate for this with leverage. One step toward greater leverage is to reduce cash inventories to a bare minimum and rely instead on borrowing, so the dealers balance sheet looks like this:

Assets	Liabilities
Securities	Loans

A dealer that looked like this would experience fluctuation in the size of his balance sheet as he bought more securities to absorb selling pressure, and then sold them off to absorb buying pressure. But a balance sheet like this is always net long securities and net short cash, and that involves risk exposure (to falling security prices) that is not necessary to the market making business. Ideally, the dealers would like to reduce inventories to zero, so they eliminate inventory risk, and they can do this only if they hold as many short positions as long positions. Typically real world dealers wind up net long, but they move as close as they can to the following ideal balance sheet:

Assets	Liabilities
Reverse	Repo

Outside the dealer balance sheet there are other people holding inventories of securities and cash. The dealer accesses the inventory of securities using reverse and the inventory of cash using repo. We can think of the repo as borrowing money to finance the dealer's long security positions, and reverse as lending money (borrowing collateral) to finance the dealer's short positions.[2]

In effect, the dealers have very good access to cash (repo market) and to securities (reverse market) when they need them, so they can behave as though they do have inventories even though the inventories are actually out in the market some place. In effect, dealers operate a just-in-time inventory system.

In a crisis, as we will see later, it matters where the inventories are. Ultimately access to cash comes from higher up in the hierarchy, from banks, and access to securities comes from lower down in the hierarchy, from security holders. So the dealer is in effect straddling layers of

[2] I remind you that I construct these balance sheets by "following the money". Thus, what I have on the liability side is a money debt, a promise to pay money at a later date that is secured by security collateral. And what I have on the asset side is a money loan. This makes sense of the accounts, but you have to see the security flows as well in order to see what is happening as knitting together the layers of the hierarchy. The money debt is a security sold and repurchased later, and the money asset is a security purchased and resold later. See Stigum Table 10.1 (p. 434) for actual data on the balance sheets of security dealers.

the monetary hierarchy. Sometimes they get into trouble when they have to come up with securities they have reversed in and then sold. The more significant troubles come when the dealer has to come up with money they have repoed in and then spent.[3] They are dependent on banks for refinance. Meanwhile, in normal times, the dealers make the market between cash and securities, standing ready to take the opposite side of trades that others may wish to make.

Arbitrage

The business of making markets is sufficiently competitive that profits are hard to come by. The real profit in the business comes from trading on the information one gets from knowing the state of the market better than anyone else, and from having privileged access to both money and securities in the repo and reverse markets which means the ability to put on a position more cheaply than anyone else.

Dealers take "positions", which means they speculate on how prices will change in the future. They deliberately mismatch their book in the direction they think will be profitable. They may be net long securities if they are bullish on security prices, but mostly the trades are more subtle than that. They may be net long some class of securities and short another class, betting on the relative price between them. Sometimes this means betting on what the Fed will do, since the Fed influences the price of money in Fed Funds and all other markets key off that. Sometimes this means betting what the market will do, since spreads can change.

Example: yield spread arbitrage, in which the dealer identifies apparent mispricing at one segment of the yield curve, takes a position but hedges overall interest rate exposure by taking an opposite position at another segment of the yield curve. The initial distortion in asset prices that the dealer identified as offering an arbitrage opportunity can be understood as a (temporary) fluctuation in demand or supply. Observe that, by taking advantage of the distortion, the dealer in effect spreads its impact into other markets, so reducing its impact, indeed counteracting the distortion. (Dealers engaged in arbitrage are acting as "porters" of liquidity from one market to another.)

This kind of arbitrage is important. The consequence is that markets for individual securities are in fact not separate, each with its own flow of supply and demand that causes price to fluctuate. Speculators are joining the separate markets into a single market, and in doing so they bring about a result that is no part of their intention, namely liquidity. Arbitrage and liquidity are in this sense two sides of the same coin.[4]

---

[3] In this recent crisis, dealers have had problems of both kinds. So-called "fails", meaning failure to receive back collateral that had been used to secure a loan, became so widespread that the Fed intervened using its Term Securities Lending Facility, lending its own inventory of Treasury securities to dealers. Also, access to cash became a problem as the repo market collapsed, so the Fed opened its Primary Dealer Credit Facility, essentially a lender of last resort facility for dealers.

[4] A philosophical question: What is the relation between liquidity and asset prices? The modern theory of finance is built on the assumption of perfect liquidity, so that prices can be fully efficient. It is supposed to be arbitrage that creates perfect liquidity by entering to take advantage of even the smallest deviation of price from fundamental value. To say that liquidity is perfect is to say that liquidity is a free good.

But in a fully efficient market, arbitrage would not be profitable since all bets would be fair bets. So

In financial theory, it is common practice to assume perfect arbitrage, and hence also complete liquidity. Assets are assumed to trade at their fundamental value since any other price would create an arbitrage profit opportunity. In effect, the world that the finance theorists imagine is a world in which the VBT outside spread is very very narrow, so there is no room and no need for dealers. In the real world, the outside spread is quite wide, dealers offer prices inside that spread but the prices can deviate very far from fundamental value. That is the world Fischer Black was talking about in his infamous presidential address to the American Finance Association when he said that he thought markets were efficient, meaning price was usually within a factor of two of true value.

We can understand what Black is saying by referring back to the Treynor model. Suppose that fundamental value is the price that dealers would quote if their inventories were exactly zero, so they are not exposed to any price risk. The Treynor model then shows how market making by dealers pushes price away from fundamental value, on one side or another, by more or less depending on the size of the outside spread and the dealer's maximum long and short position limits. Standard asset price theory abstracts from this effect, in effect treating the outside spread as collapsed around fundamental value, so there is no need for dealers. Some markets are close approximations to this, but others are not; some times are close approximations to this, but others are not.

---

position takers would not make money. And if position takers do not make money, they will not compete so much for the market-making business. This seems to imply that markets would be less liquid--wider bid-ask and more volatile prices. Thus, in practice it seems that we must expect liquidity to enter into asset prices.

11. Banks: The Market for Liquidity

Last time we introduced the concept of a continuous or liquid market, and argued that the institution of a dealer is key to achieving a liquid market. The dealer supplies market liquidity by quoting prices to buy and sell, and absorbing the resulting order flow onto his balance sheet. The consequent positions, long and/or short, have to be financed somehow, which means that the dealer is a demander of funding liquidity.

Today we ascend the hierarchy and consider banks as dealers. Way back at the beginning of the course, I said we would be thinking of the banking system as making markets by being prepared to trade currency for deposits and vice versa, at a fixed price par. When we try to apply the Treynor model directly to banks in this sense, we immediately confront a puzzle. From the perspective of the Treynor model, we are looking at a limiting case, in which the spread is zero and the price is fixed, and from what we know about dealers this kind of restriction makes such dealing both impossible and unprofitable. And yet apparently it happens; that's the puzzle.

Instead of confronting this puzzle head-on, let's start from what we already understand. Let's start by thinking of security dealers as banks.

***

The Treynor model focuses attention on the net position of the dealer, because he is mostly concerned about price risk. However, if we want to understand banking, we need to bring into the picture the gross position as well, which is an order of magnitude larger. Here is a stylized balance sheet, dividing the dealer's balance sheet into two pieces.

	Assets	Liabilities
Matched book dealer	Securities in, 100	Securities Out, 100
Treynor "speculative" dealer	Net financing, 10	Loans, 10

The matched book part is comprised, in principle, of equal and opposite long and short positions, so that fluctuation in security prices have exactly offsetting effects. Only the speculative book is exposed to price movement, and it is an order of magnitude smaller than the matched book. I am showing the dealer as net long, and funding this net long position with loans. The data for primary dealers in the US shows their outright position, but not necessarily their actual exposure since it does not include derivative positions, so we cannot exactly translate between the conceptual distinction and the data.

Security Dealers as Money Dealers, Term Rates

The data does however show in some detail how all these positions are financed, and it is instructive to look at closely. The data below is from October 3, 2012, Table 4 in the dealer statistics: http://www.newyorkfed.org/banking/reportingforms/primarystats/deal.pdf

	Assets	Liabilities
	854 overnight reverse 1253 term reverse 515 net financing	1796 overnight repo 826 term repo

(Might be useful to refresh your understanding of repo, lecture 6, p. 46.) Note that I have added net financing as a balancing item, since borrowing is greater than lending. Net financing is not the same as the outright positions shown in Table 3 (which add to about 270), and one reason is the way that mortgage backed security positions are treated. We don't have to worry about that, since we are trying to understand concepts.

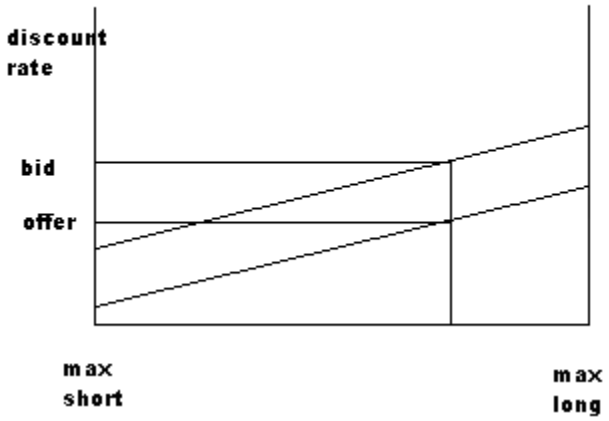
I have made the point before that security dealers operate somewhat like banks, insofar as they are borrowing short term and lending long term, which you can see in the data. Now I want to go farther using our idea of matched book. Suppose that the term of reverse lending and reverse borrowing are exactly the same. Then we can rewrite the dealer balance sheet as follows

	Assets	Liabilities
Matched book dealer	854 overnight reverse 826 term reverse	854 overnight repo 826 term reverse
Treynor "speculative" dealer	427 term reverse 515 net financing	942 overnight repo

In fact we can go farther still, by distinguishing two different kinds of price risk on the speculative dealer side. Just add 515 term repo on the liability side and 515 term reverse on the asset side

	Assets	Liabilities
Matched book dealer	854 overnight reverse 826 term reverse	854 overnight repo 826 term reverse
Treynor "speculative" dealer	942 term reverse 515 net financing	942 overnight repo 515 term repo

So the dealer has 515 exposure to security price risk, and is funding that exposure using term repo. That's the risk we talked about last time. But he also has 942 exposure to liquidity risk, borrowing short and lending long in the money market. That is new, but we can adapt the Treynor model to handle it.



Liquidity Risk

Since here we are talking about the money market, prices are quoted as yields, so the bid is higher than the offer and the dealer quote curves slope up rather than down. But the same idea applies, that the dealer is willing to take on more risk, in this case liquidity risk, only if compensated by higher expected return, in this case the difference between the term rate and the overnight rate. (I label the vertical axis as discount rate to remind us of our discussion of the world that Bagehot knew, where the discount rate was a term rate. For now think of the overnight rate as fixed by the Fed's official Fed Funds target. In effect, the vertical axis is the spread of 3 month over overnight)

The point to emphasize here is that the security dealer chooses not only how much price risk to take (and chooses price quotes to achieve that) but also how much liquidity risk to take (and chooses yield quotes to achieve that). The balance sheets show clearly that, if the dealer wanted to, he could take either less or more liquidity risk than price risk—the two numbers do not have to be the same.

Now let's translate all this into banking language. Overnight repo is analogous to a demand deposit account, and term repo is analogous to a short term loan. So our Treynor diagram is not just about the determination of term repo rates, but also about bank term rates as well. Banks make money, in part, by issuing money as their liability and investing the funds they receiving in interest bearing securities. That is exactly what the dealers are doing as well.

Digression: The Evolution of Banking

It will be helpful when reading Stigum to have in mind the traditional picture of banking that she would have learned when she was in school. In this picture, a bank is an institution that makes loans to corporate customers and takes deposits from households retail. These are essentially passive activities. Where the action comes is in managing the difference between these two. If lending exceeds deposits (as is typical for money center banks) then they must raise

additional funds in the money market. If deposits exceed lending, then they must find profitable outlets for excess funds, typically some money market asset or longer term security.[1]

This traditional view sees banks as intermediaries between savers (household depositors) and investors (corporations). All the money market does is to move those deposits from banks with excess to banks with deficit, so in the aggregate what is happening is that deposits are funding loans. Banks are intermediaries that facilitate that movement, and also intermediaries in the sense that borrowers and lenders both face the bank, not each other. In this traditional view, banks are important mainly because of their role in fostering capital accumulation.

Corporations		**Banking System**		**Households**	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	Loans	loans	deposits	deposits	
		securities	borrowing		
		reserves	capital		

This is a pretty good picture of what things looked like in the fifties. One thing it leaves out is that government securities were a large portion of bank portfolios, so that much expansion of lending was replacing government securities with private loans. (As you read you will note a lot of ways in which government securities are given privileged legal status. The government in time of war uses the banking system to support the market in its debt.)

Things have however changed. Stigum talks about the "Death of loans"--top corporate customers have access to open market credit, particularly commercial paper, where they can borrow more cheaply than banks. Banks have adjusted to the loss of this business by instead providing backup lines of credit to commercial paper issues.

The second big change I would call the "Death of deposits". Reg Q worked to provide banks with low cost funds by limiting interest payment on deposits, but this eroded over time. Rise of money market mutual funds as competitors. So-called "disintermediation" as ultimate borrowers and lenders began to take interest rate risk, without the bank standing between them.

Finance Companies		**MMMFs**	
Assets	Liabilities	Assets	Liabilities
Loans	Commercial Paper	Commercial Paper	MMMF shares

See data at http://www.federalreserve.gov/releases/z1/Current/
MMMFs are table L.121, Finance Companies are table L.127

[1] There is also the matter of managing interest rate risk and liquidity risk. See Stigum and Branch Managing Bank Assets and Liabilities; Strategies for risk control and profit (1983).

As Stigum was writing, banks were losing their core business, but at the same time being prevented from going into any others. The Securities Industry Association (SIA) resists bank attempts to get into underwriting non-exempt securities. The Fed imposes capital adequacy regulations on banks that are not imposed on others. This accelerates the tendency to strip the balance sheet, to go for off-balance sheet exposure, substituting FRA for deposits, futures for actuals, etc....

This apparent discrimination against the banks was also in part motivated by concerns about safety and soundness of the payment system. Also the fact that banks, unlike their competitors, have privileged access to the Fed, and the Fed wants to keep use of that access to a minimum.

In the last ten years or so we have seen a further development of this parallel banking system, into what now people call the shadow banking system. Essentially we seem to be moving from a system of bank credit to a system of capital market credit. Corporations pioneered in this shift, but the latest development has notoriously been more about household credit, namely mortgage credit.

Households		Shadow Bank		MMMF	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	mortgage	RMBS	MM funding ABCP RP	 ABCP RP	Shares

The shadow banking system faces the same problems of liquidity and solvency risk that the traditional banking system faced, but without the government backstops (mainly Fed LOLR and FDIC deposit insurance). Instead the shadow banking system relies on the market for both, the wholesale money market and the CDS market mainly. [2] We will focus on liquidity risk and hence the wholesale money market.

Overnight Rates—Fed Funds

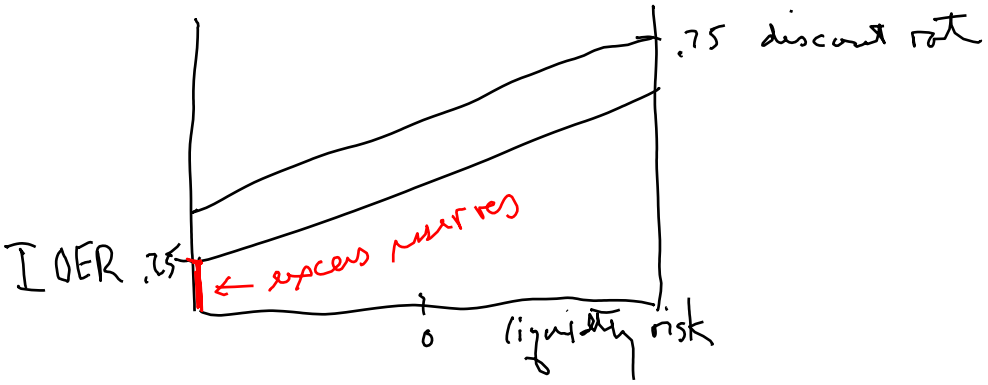
We turn now to the question of the overnight rate, which we have so far treated as though it were pegged by the Fed. The full story is more interesting. Essentially the Fed does two different things—it fixes the outside spread, and it trades in the repo market to influence the supply of reserves. (Remember that Fed Funds are overnight promises to pay reserves.)

Outside spread. Interest on reserves is currently .25, and the discount rate is currently .75. These are essentially an outside spread. It is a pretty narrow outside spread (only 50 basis points), and the Fed's communication makes clear that it anticipates widening that spread in the future. The

[2] The lender of last resort to the shadow banking system is the traditional banking system, operating through various lines of credit and liquidity support commitment.

interest on reserves, in particular, it imagines will be 75 basis points below the target. And before the crisis, discount rate was 100 basis points above the target. At the moment the effective funds rate is actually below IOER rate (16 bp vs. 25 bp), but that is because there are some lenders in the market who are not eligible to receive interest on reserves, and there are lots of excess reserves, so they have to take what they can get. In a way, we should think of the effective FF rate as 25, not 16, and that helps us a bit to understand why repo rates are higher than FF. A lot of distortion in the system right now.

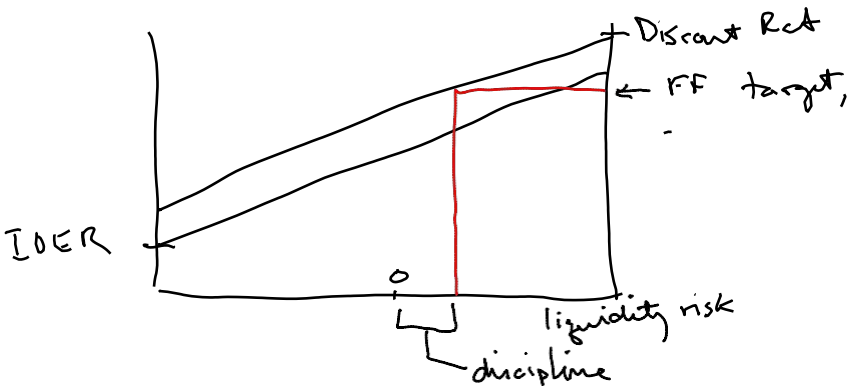
That said, we can use the Treynor diagram to understand what is happening:



In this situation clearly the Fed can just establish the FF rate as IOER. In the future however, it intends just to set the outside spread and then perhaps influence the level by trading. Just so, recall how temporary open market operations work:

Bank		Dealer		Fed	
A	L	A	L	A	L
+reserves -loan			+repo -loan	+repo	+reserves

Here I am showing the dealer using the Fed's loan to pay off bank loan. (Before I showed it used to accumulate a deposit.) The result is that the Banking system as a whole now has more reserves, which should ease some of the need to borrow reserves and hence lower the effective funds rate. The point is to keep the FF rate near the target by intervening daily. The outside spread is intended only for anomalous situations for individual banks.



So now we can return to the puzzle we started with—how do banks manage to make markets in currency and deposits at a zero bid-ask spread and a price that is fixed at par, and how do they make profit doing so? The answer is that they are also in a complementary business, the business of bearing liquidity risk by issuing demand liabilities and investing the funds at term, and this business is highly profitable. They cannot change the price of deposits in terms of currency, but they can expand and contract the quantity of deposits because deposits are their own liability, and they can expand and contract the quantity of currency because of their access to the discount window at the fed. Security dealers are stuck with the quantity of securities out there, and stuck also with the quantity of cash out there. Banks are not stuck with the quantity of deposits or currency, so although they have less flexibility on price, they have more flexibility on quantity.

Term Rates, Redux

In normal times, the Fed targets overnight rates, leaving term rates to be determined in the market. But during the crisis the Fed did a lot of intervention in term markets, which broke down, mainly lending. And now the Fed is experimenting with intervention on the other side of the market, by issuing term deposits of its own, first time Sept 10, 2012.

http://www.frbservices.org/files/centralbank/pdf/termdepositfacility_offeringresults_A26_091012.pdf
We can see the consequence for the Fed's balance sheet:

http://www.federalreserve.gov/releases/h41/Current/

What is the Fed trying to undo? Topic for next lecture.

http://www.newyorkfed.org/markets/Forms_of_Fed_Lending.pdf

12. Lender/Dealer of Last Resort

http://www.newyorkfed.org/markets/Forms_of_Fed_Lending.pdf

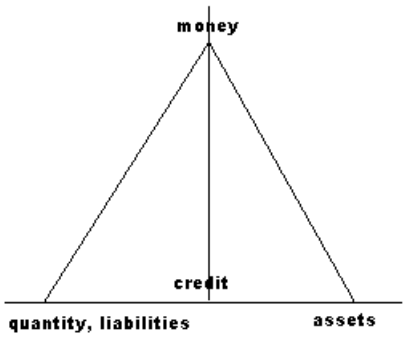
Now that we have built up a picture of the entire (domestic) monetary system, we can turn our sights to the problem of managing that system. As Walter Bagehot tells us, "Money does not manage itself, and Lombard Street has a great deal of money to manage." Today we treat crisis management, which was the origin of management more generally, and indeed the focus of Bagehot's famous Lombard Street (1873). Once the central bank took responsibility for crisis management, it was inevitable that it would take responsibility also for trying to prevent crisis by intervening well before any actual trouble. That is to say, lender of last resort is the logical as well as historical origin of modern money management.

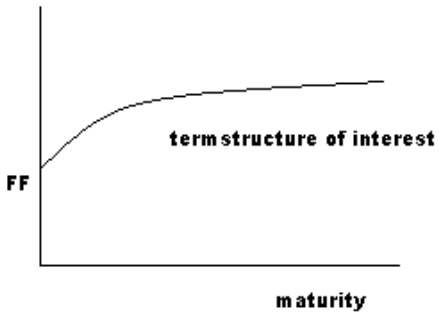
Today we reverse the order, taking monetary management first, and crisis management second.

Evolution of Money Management

As you read Stigum Ch. 9, keep in mind the question whether the Fed controls the Fed Funds rate or some measure of reserves, i.e. price or quantity. We can think of Fed interventions in both ways. First, the Fed is using its control of money (at the top of the hierarchy) to try to influence the expansion of credit (at the bottom of the hierarchy). Second, the Fed is using its control of the money rate of interest (at the short end of the term structure maturity spectrum) to try to influence the bond rate of interest (at the long end).

In both cases, it is actually the private credit market that will determine quantities and prices, but the Fed is an important player in that market and can influence it. The point of this lecture is to understand exactly what it is that gives the Fed leverage.





In the earlier 3rd edition, Stigum made clear that standard economic theory has no very good answer to the question of how the Fed gets leverage over the real economy. Thus: "from long experience, Fed technicians knew that the Fed could not control money supply with the precision envisioned in textbooks" (395); Goodhart's Law (400). Her general thrust is non-monetarist: "Much of the macro theory that links money supply to the price level depends on concepts that

have little relevance to the workings of a modern financial system, but at least such a theory exists and has even earned a Nobel prize. No comparable macroeconomic theory for credit aggregates yet exists." (412) She is talking here about the quantity theory of money:

MV=PQ

which suggests that control of the quantity of money (left side) gives control over the price level and aggregate output (right side).

Leaving theory aside, and focusing on practice, Stigum further suggests that the talk about reserves (free reserves, borrowed reserves, non-borrowed reserves) is a cover for what the Fed is actually doing, which is targeting interest rates. Post 1951 it targeted free reserves, then explicitly interest rates. Volcker's 1979 Saturday Night Special switched to non borrowed reserves, then in 1983 to borrowed reserves. Since 1987 there has been a more or less explicit targeting of the fed funds rate.

Date	Target	Quantity or Price?
1951	Free Reserves Interest	Quantity?
	Rates Nonborrowed	Price
1979	Reserves Borrowed	Quantity
1983	Reserves	Quantity?
1987	Fed Funds Rate	Price

Here are some definitions to help make sense of it all:

Total Reserves = Required Reserves + Excess Reserves

Total Reserves = Non-borrowed Reserves + Borrowed Reserves

0 = (RR-NBR) + free reserves (ER-BR)

Throughout all of these regimes, **one constant is the Fed's idea that its job is to attend to the balance of elasticity and discipline in the monetary system as a way of controlling the flow of credit**. Too rapid credit growth requires more discipline; too slow credit growth requires more elasticity. Operational practice changes, but the underlying goal is the same.

From these equations, you can see that targeting excess reserves amounts to accommodation of changes in required reserves, and targeting borrowed reserves amounts to accommodation of changes in non-borrowed reserves. So these "quantity" targets are not the kind of strict quantity control favored by the Currency School and modern monetarists. They involve accommodation in terms of quantities, but maybe not in terms of price. To maintain constant excess reserves, for example, the Fed may have to allow interest rates to change. In fact the idea of these policies was to use the quantity target to exert pressure on the market, since banks don't like to be borrowing from the Fed, but to allow the market to determine the rate of interest.

Even the remaining quantity regime, the Volcker regime of nonborrowed reserves, might be viewed as a stealth interest rate target. Crandall is good on this.

Modern Money Management[1]

Today, we are no longer on the gold standard, and central banks are not required to control the exchange rate. Instead they tend to focus on internal stabilization. They do not want to accommodate domestic credit expansions and contractions, but rather to counteract them in order to smooth business fluctuations. Modern money management looks not at gold flows but rather at movements of the price level as an indication of incipient imbalance between the pattern of cash flows and cash commitments.

Central banks typically follow some version of the so-called Taylor Rule, which I'll write as follows:

$$R = \rho + \pi^e + \alpha(\pi^e - \pi^*) + \beta(Y - Y^F)$$

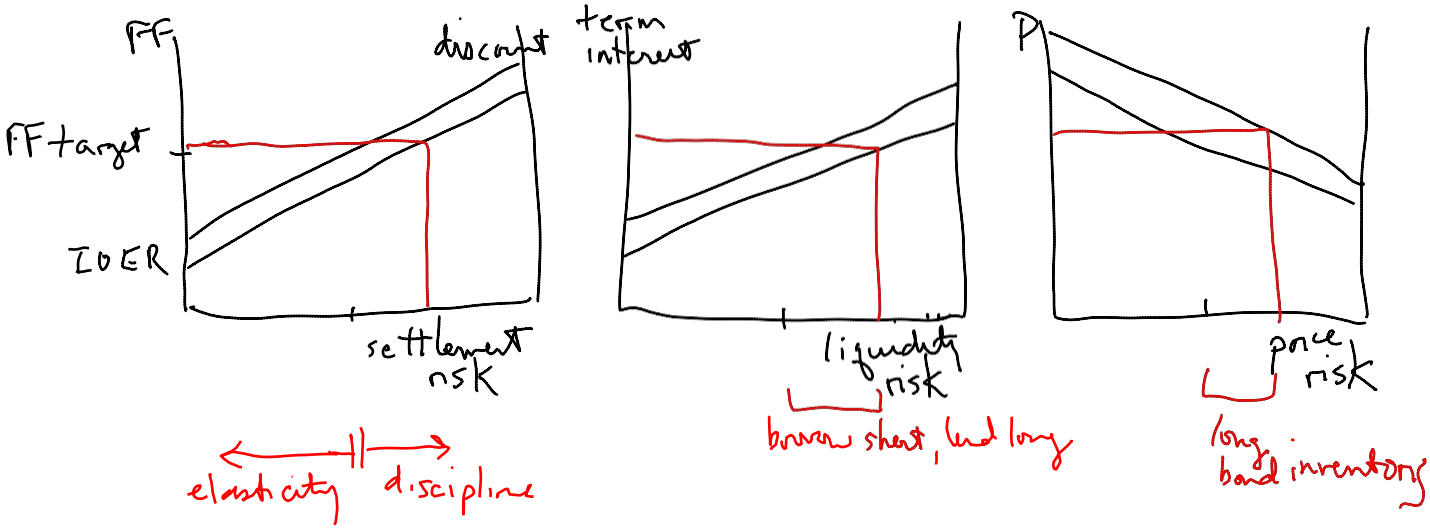
The first three terms of this equation express the idea of the Fisher Effect, which says that the private negotiations between borrowers and lenders tend to produce a nominal rate of interest that takes account of the expected rate of inflation (so $R = \rho + \pi^e$ where ρ is the real rate of interest and π is the rate of inflation). This is what the private markets tend to do on their own. Fisher himself thought that private markets tend to have "money illusion", so they systematically underestimate future inflation, and hence mistake merely nominal interest rate changes for real interest rate fluctuations. In his view, monetary policy was mainly about correcting for these systematic mistakes.

In modern treatments, we tend to assume "rational expectations" which means that the private forecast is as good as can be achieved from available information. In modern treatments, the problem is not that people make mistakes, but that their adjustments of the nominal rate to expected inflation do nothing to bring inflation under control—they merely accommodate it. The role of the central bank is to do more than that, by raising interest rates even more when inflation exceeds the target, and lowering even more when inflation falls below the target. (That means that $\alpha>1$.) Translated into our language, the central bank imposes discipline when prices are rising too fast, and tries to instill some additional elasticity when prices are not rising fast enough.

In the current crisis, this "inflation targeting" framework has come under some question. Many people think that the central bank played a role in causing the credit bubble by keeping interest rates too low for too long, and that it did so because it was watching inflation not asset prices (housing prices). So expect to see some evolution in thinking about money management, but that evolution will be about finding a more appropriate way to achieve balance between

[1] Unlike previous years, I skip over all the developments between Bagehot and Taylor. Thus there is nothing about the Strong Rule of the 1920s, or Monetary Walrasianism of 1960s and 1970s (both Keynesian and monetarist), which represented stages of the intellectual evolution from one to the other. All these stages can be interpreted as different ways of understanding the source of imbalance between cash flows and cash commitments, and hence different approaches to managing the balance of elasticity and discipline.

elasticity and discipline. Maybe we will be seeing a return to Hawtrey, with his emphasis on the inherent instability of credit, operating through the effect of lending on collateral valuation.



These are the three diagrams from last time. We can think of the leftmost as describing how the Fed controls the FF rate; the rightmost is about how the market determines the price of long term bonds. And the middle diagram shows how both of these two influences come together to determine the term rate of interest.

If the Fed wants to tighten, it raises FF target. That immediately lowers the profitability of the liquidity spread in the middle diagram unless the term interest rate rises by the full amount (and perhaps a bit more to compensate for anticipated future tightening as well). This is funding cost for bond dealers, so they are willing to hold existing bond inventories only at a lower price, hence higher expected profit.

This chain of events I propose to you as a sketch of the monetary transmission mechanism in the modern economy. Note it all goes through the price of money and the price of securities. This chain of events is just for normal times. What about crisis?

Anatomy of a "Normal Crisis"

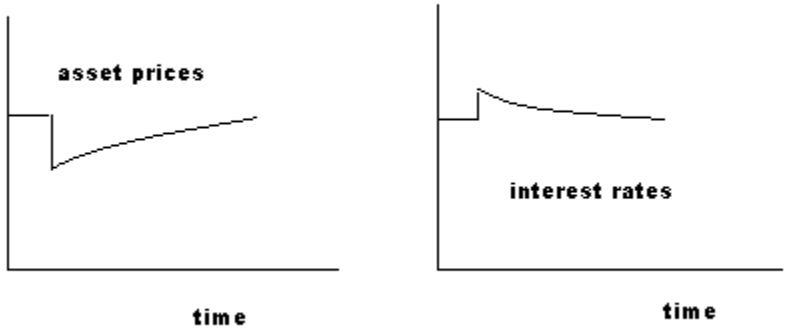
Suppose there is a sudden shift in preferences in favor of money and against securities. (I will come to the question of the origin of this shift by the end of lecture, but for now take it as given.) One way to put it is that people aren't prepared to delay settlement anymore; they don't want promises to pay, they want money.

If it's a tiny and temporary shift, it might be able to be accommodated by the dealer system as follows.

Households		Dealers		Banks	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
-securities +deposits		+securities	+RP	+RP	+deposits

Thus dealers swap IOUs with the bank, and the effect is to increase the means of payment; this increase is what the dealers use to pay the households. The households get the trade they want because the dealer system is prepared to take the opposite side of the trade.

Why are they willing to do so? In order for dealers and banks to participate in this kind of action, they have to have an expectation of making money. Thus, dealers buy securities below equilibrium price, banks make loans above equilibrium price, and the distance from equilibrium is determined by the amount of risk they think they are taking. (You can think of this price change as determined by their inventory change, following the dealer model.) They will not be willing to do any of this unless they see a light at the end of the tunnel in terms of eventual return to equilibrium prices.



From this point of view, two important things happen during crises. First, the dealer system **buys time** by letting the private sector think that accounts are settling rather than being delayed; the private sector gets cash today instead of a promise of cash tomorrow. Of course accounts are not actually settling--the cash is new cash created for the purpose—but it is the same as old cash from the perspective of the private sector. Second, during that extended delay, prices are pushed away from equilibrium, and the disequilibrium prices **put pressure** on the system to actually settle rather than delay; just so, higher interest rates raise the cost of delaying payment. So although actual settlement is delayed, incentives are put in place to encourage more rapid settlement in the future.

In the course of normal economic interaction, there are many times when the marketmakers face trouble of this kind and turn to the banks for refinance, and the banks are able to meet the need without trouble, harvesting a nice return for doing so. That is to say, there are many cases where liquidity crises for market makers do not turn into banking crises. These are cases of "normal crisis" where the liquidity constraint is doing its job of ensuring continued coherence in the

economy.

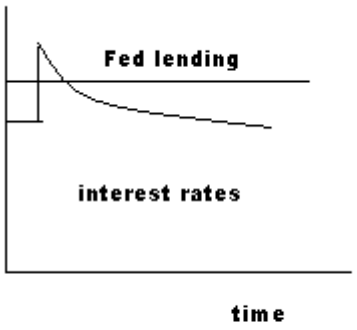
Anatomy of a Serious Crisis

Sometimes however the problem gets out of hand, and banks themselves run into trouble. Then they turn to the Fed.

Dealers		Banks		Fed	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+securities	+RP	+RP +reserves	+deposits +discount loan	+discount loan	+reserves

Just as the dealer system could take the problem off the hands (the balance sheet) of the households, so too can the Fed take the problem off the hands (the balance sheet) of the dealer system. It does so by expanding its own balance sheet.

In such a circumstance, Fed lending can put bounds on the fluctuation of interest rates and so prevent the problem from spiraling out of control. If the Fed is willing to lend at 10%, then interest rates will not go higher than 10%, and that provides a lower bound to asset prices as well.



[This diagram is not exactly right. The Fed's discount rate, shown as "Fed lending", is the interest rate paid by banks, usually 100 basis points over the Fed Funds target. The interest rate charged by banks will be a markup over their cost of funds, not the same as the Fed lending rate. But this is a refinement. The important point is that the Fed's willingness to lend to the banks as a rate lower than they would lend to each other makes it possible for the banks to lend to the dealers at a rate lower than they would otherwise charge. Putting a ceiling on the money rate of interest thus indirectly puts a floor on asset prices. In the recent crisis this transmission broke down and the Fed moved to put a floor on asset prices directly by buying them, not just financing their purchase by dealers.]

Why **can** the Fed help? When the Fed helps the banks, what it does is to expand reserves. Hence the money supply expands. We have seen that the market makers are long securities and short cash. What the Fed does is to backstop those short positions by shorting cash itself. The advantage the Fed has is that it prints the stuff, and so there can be no short squeeze on the Fed. (Actually, this is strictly true only domestically. Internationally the Fed itself faces a reserve constraint. This international constraint becomes an issue only if the problem is so large that it cannot be solved within one country.)

Why **does** the Fed help banks in a liquidity crisis? Partly because that is its job. But that job assignment is backed up by a very real threat. If the Fed does not help the banks, what happens?

Why intervene (why elasticity)?

(1) Government Bank. Consider this. Your book makes the point that banks meet the demands of dealers, first by expanding up to their reserve constraint, and then by replacing security holdings with loans. In effect a liquidity crisis in any market tends eventually to get funneled into a demand for bank lending, and if that demand is high enough banks will eventually try to meet it by dumping their holdings of securities. Note well that the securities banks hold are overwhelmingly government securities. What this means is that, if the Fed does not help, there is always danger that liquidity problems in any market get translated into liquidity problems in the government securities market--and that cannot be tolerated. Government securities trade at a premium not only because they are default free, but also because of the liquidity of the markets. A Fed that failed to maintain the market for government debt would be a Fed soon under direct Congressional control. A Fed that ensures liquid markets for government debt in effect ensures liquid markets for other debts as well.

(2) Bankers Bank. Actually this problem of potential dumping of government securities is just one dimension of a larger potential problem. Sharp moves of security prices and interest rates can disrupt balance sheets, especially leveraged balance sheets, all over the economy and force emergency liquidation. If your security holdings fall in value, you can't borrow as much using the securities as collateral, and if you can't finance your holdings then you might be forced to sell. As more securities get thrown upon the market, the initial dislocation (consumers want money rather than securities) gets larger rather than smaller. Bank liquidation of security holdings is only one example of this more general phenomenon. The point is that such liquidation makes the problem worse not better, and drives security prices and interest rates even farther from their equilibrium levels. From time immemorial, the lead banker in the system has stepped in to play a stabilizing role in a crisis, even from self-interest.

(3) Stabilization Policy. Allowing liquidation to take its course can have even more profound effects through macroeconomic channels. One way for borrowers to raise needed cash is simply by reducing their spending. But their spending is the income for other agents in the economy, and reduction in their income receipts may thus spread the debt repayment problem throughout the economy. Something like this is what happened during the debt deflation of the Great Depression. In fact, the attempt to liquidate holdings of commodities also drove down the price level, which increased the burden of nominal debt and so made it harder to repay. Once again the

important point is that the natural mechanisms of the economy may make the problem worse rather than better.

So far so good. It looks like the Fed can solve any liquidity crisis simply by providing the needed liquidity. So why doesn't it just go ahead and do so? **Because the crisis is typically a symptom of some underlying deeper problem**. What the Fed wants to do is to buy time for adjustment, but not so much time that people lose the incentive to make an adjustment. Then the problem of the Fed is to ensure enough price movement to put pressure on the system but not so much as to make the problem worse.

Why not intervene (why discipline)?

Minsky's conception of financial crisis is helpful here. He emphasizes that at any moment in time the history of past borrowing and lending (based on expectations of future cash flows) has given rise to a certain pattern of cash commitments, some due today and some out into the future. At the same time, there is a pattern of cash flows emerging from the real economy, today and expected out into the future. If these two patterns are lined up, then debt is serviced and we have no problems.

When they begin to get out of kilter, cash commitments cannot be met and have to be delayed. If the people who are receiving the cash don't need it, again there is no problem because they'll just lend it, and the consequence is an expansion of credit. This is what happens in a boom. A boom is a credit expansion, which is to say a deliberate delay of settlement, on the basis of expectations about the future. The problem comes on the reverse side of this expansion, when there are demands to delay settlement and no supply, or demands to make payment and no matching cash flow.

Such a disjuncture between the pattern of cash commitments and the pattern of cash flows shows up as pressure in the money market, because it involves/requires an expansion of inside credit. (Such disjuncture also drives the kind of portfolio preference shift out of securities into money that I used as a characterization of financial crisis earlier.) What makes the problem worse (at least potentially) is that "the cash commitments of each unit depend on the cash commitments of every other unit. The whole web of interlocking commitments is like a bridge we spin collectively out into the unknown future toward shores not yet visible." That means that price movements in the money market may not be sufficient by themselves to bring the two patterns back into line with one another. Now the Fed could, as we have seen, solve the immediate problem by providing liquidity. The fear is, however, that by preventing crisis it prevents the kind of adjustment that will ultimately resolve the underlying problem.

One way to put this is to say that once the Fed expands the money supply to accommodate immediate needs (elasticity), it faces the problem of contracting again sometime in the future (discipline). But if the Fed always meets the demand for liquidity, then it allows people to delay adjusting their balance sheets to their changed circumstances, so the discipline never comes and the imbalance continues. If the imbalance is prevented from showing up in a liquidity crisis, it

will show up somewhere else, as for example in inflation (or exchange crisis). A summary way of saying this is that if the Fed does not succeed in imposing **discipline**, then its offer of **elasticity** means that the problem shows up somewhere else. This is a crude statement of a deeper truth--that masking the symptoms of the underlying problem does not resolve the problem.

Thus the Fed faces a dilemma. If it proves too ready to provide liquidity, it prevents necessary adjustment. If it does not prove ready enough, then adjustment may prove fatal to the system, or at least be more costly than need be. It is a fine line. From this point of view we can understand the job of the Fed better. It can paper over problems in times of crisis, but needs to be careful that in doing so it contributes to their solution not to a delay of their solution.

Dealer of Last Resort (Money Dealer)

Let us apply this idea to the shadow banking system, and the stresses it experienced in recent years. One of the first things that happened is that MMMFs refused to roll over ABCP—essentially they wanted to convert their money market assets into cash. In the first stage of the crisis (Fall 2007), this was handled by replacing ABCP with RP, still secured funding but now shorter term. In the second stage of the crisis (Bear, March 2008), RP became suspect as well, haircuts were increased, so forcing shift to unsecured Eurodollar and Financial CP funding. In the third stage of the crisis (Lehman), these also became suspect. In each crisis trouble shows up as upward pressure on money market funding.

We need to understand these pressures as coming ultimately from the need to maintain par. Banks that had backup lines of liquidity to the shadow banking system had themselves to step up to the plate in many instances, like so:

Shadow Bank		Citibank/Lehman		MMMF	
Assets	Liabilities	Assets	Liabilities		
	-ABCP + loan	+loan	+RP/ Fin CP/ Eurodollar	-ABCP +RP/ Fin CP/ Eurodollar	

This set of balance sheets shows more or less the anatomy of a normal crisis, and that's what everyone thought this was when it started. The Fed did not involve its own balance sheet, but it did try to encourage banks to involve their own balance sheets, by cutting the Fed Funds rate from 5% to 2%. But it wasn't enough.

The banks and more generally the dealer system, having taken on responsibility for financing the shadow banks, now began to run into problems themselves, and that required more serious Fed intervention. The cornerstone of the domestic intervention was various forms of term lending to banks and dealers. The cornerstone of international intervention was liquidity swaps with

foreign central banks that then made term loans to their own banking system. Here is the domestic story:

US Banks/Dealers | Fed | MMMF

A	L	A	L	A	L
	+term loan	+term loan -Treasury bill		+Treasury bill	

Here is the international story:

Foreign banks | Foreign central banks | Fed | Treasury | MMMF

A	L	A	L	A	L	A	L	A	L
	+term loan	+term loan	+liquidity swap	+liquidity swap	+Treasury deposit	+Treasury deposit	+Tbill	+Tbill	

The domestic story caused the Fed to liquidate about half of its holdings of Treasury bills. The international story caused the Fed to expand both sides of its balance sheet.

You might ask why the Fed got involved in supporting foreign banks. One reason is that they wanted to stop foreign banks from liquidating their holdings of RMBS, hence driving down the price and adding to the liquidity spiral. Another reason was the pressure that foreign banks were putting on domestic banks through the correspondent banking system. That pressure showed up in the spread between the Eurodollar rate and the term Fed Funds rate that we talked about back in Lecture 8.

Here is the correspondent pressure story:

In practice the first line of liquidity defense for the Eurodollar system flows through domestic banks, call it “lender of first resort”. Remember that the Eurodollar system uses correspondent balances in New York as its reserves. So even before the British Bank goes to the Bank of England, it will go to its New York correspondent. That New York correspondent can then go to the Fed Funds market. If the Fed is committed to a particular rate in the Fed Funds market, it has no choice but to supply the needed funds. In this case the lender of last resort operation works as follows:

British Bank		New York correspondent		Federal Reserve	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
-reserves +reserves	-eurodollar deposit +interbank loan	+interbank loan	+daylight overdraft -overdraft +repo	+daylight overdraft -overdraft +repo	+reserves

In the first step, the New York correspondent lends at the Eurodollar rate to the British bank, using daylight overdraft at the Fed as its source of funds. It then enters the Fed Funds market to look for reserves needed to meet end of day clearing, and that tends to push up the Fed Funds rate. The Fed, committed to keeping Fed Funds at target, intervenes in the market to provide the funds itself.

The spread between Eurodollar and OIS is the incentive for New York correspondents to do this business, so allowing foreign banks indirectly to tap the Fed. The liquidity swap channel took off this pressure by channeling funds instead through official central bank channels.

13. Chartalism, Metallism, and Key Currencies

In terms of our hierarchy of money and credit, we have so far been paying most attention to currency and everything below it, so our attention has been on two of the four prices of money, namely par and the interest rate. Today we begin a section of the course that looks into forms of money that lie above currency in the hierarchy, and hence at a third price of money, the rate of exchange.

Metallism

Under a gold standard, the extension of our analysis would be straightforward. Gold is the ultimate international money, an asset that is no one's liability. Under a gold standard, each currency has its own mint par, and the exchange rate is determined by the ratio of mint pars. In this view of the world, the multiple national (state) systems relate to one another not directly (money to money) but only indirectly (credit to credit) through the international (private) system. Each national currency has an exchange rate with the international money and it is that pattern of exchange rates that sets up a pattern of exchange rates between national currencies.

Dollar = x ounces of gold
Pound = y ounces of gold
Dollar = x/y Pounds [S(1/x)=(1/y)]

Exchange Rate in a Metallic Standard World

	Gold	
	X oz. Y oz.	
Dollar	S=X/Y	Pound
Deposits		Deposits
Securities		Securities

From this point of view, the central bank is a **banker's bank**, holding international reserves that keep the national payment system in more or less connection with the international system.

This mint par ratio idea does not work exactly in practice because of the cost of shipping gold around; sometimes the currency trades higher and sometimes lower. This idea can be rationalized within the metallist frame by using the covered Interest Parity relationship.

CIP: $[1+R^*(0,T)]S(0) = [1+R(0,T)]F(T)$

where we think of the cost of shipping gold as putting lower and upper bounds on S:

$X/Y - \delta < S(0) < X/Y + \delta$

Spot rates, and even moreso forward rates, can move away from mint par, and so can foreign and domestic interest rates diverge a bit, all within the limits provided by the gold points. You probably feel an application of the Treynor model behind the scenes here, and that is exactly what we'll do but not yet. The prior question is whether this line of analysis is useful at all in the modern world, since we are not on a gold standard. Let's look at the world, and the facts that any adequate theory would need to explain.

The Facts

Figure 1 maps current arrangements, in which the dollar serves as world reserve currency. According to the most recent BIS numbers, 51 percent of the volume of foreign exchange trading involves only a few major currencies—the dollar, euro, yen, and sterling—and fully 84.9 percent of trading volume has the dollar as one leg of the trade. This latter institutional fact has led one participant-observer to opine that "the foreign exchange market is largely the price of the dollar" (DeRosa 2011, p. 4). But it is important to appreciate that the dollar in question is substantially the international private dollar, which is to say bank money not state money.

Figure 1

	Asia	US	Europe
Reserve Currency		International Dollar	
Key Currency	Yen	Domestic Dollar	Pound Sterling, Euro
National Currency/ National Deposits/ National Securities	Australian Dollar	Canadian Dollar	Swiss Franc

Further, the hierarchical character of the FX market is more than the special role of the dollar relative to everything else. The finer texture of the hierarchy is reflected in the language of

trading, which distinguishes "majors" from "minors". The majors are high volume, liquid markets, with tight bid-ask spreads, and all majors have the dollar as one leg: EUR/USD, GBP/USD, AUD/USD, USD/JPY, USD/CAD, USD/CHF. So-called "cross-currency" pairs have no dollar leg, but "euro-crosses" have a euro leg.[1] The minors trade as cross-currency pairs with some major as the other leg. With only a few exceptions, minor cross-currency pairs do not trade.

The hierarchical organization of spot FX markets carries over also into FX derivatives markets: forwards, futures, and options. Most derivative trades have the dollar as one leg, and most also involve other majors, not minors. By volume, the derivative market is larger than the spot market. Of the $4 trillion a day of FX trades, $1.5 trillion are spot transactions while $1.8 trillion are FX swaps, and the remainder are outright forwards ($.5 trillion), options and exotics ($.2 trillion), and currency swaps ($43 billion). The overwhelming majority of the market is short term; the FX market is fundamentally a money market, not a capital market.

These are the facts, but how shall we understand them? In particular, how do we understand the value of a currency that is not convertible into any metal, and how do we understand the relative price of two such currencies. The mint-par anchor of the gold standard system is nowhere to be seen, and we find ourselves intellectually adrift.

Chartalism

I have stated the problem as if it were a new one, arising from the breakdown of the gold standard, but in fact it is ancient. Joseph A. Schumpeter, in his magisterial History of Economic Analysis (1954), identifies two traditions in monetary thought that he calls chartalism and metallism. **Chartalism** proceeds analytically by thinking of money as a creation of the state, or the king. The quintessential form of money is fiat currency, a piece of paper with the king's image on it that costs essentially nothing to produce. By contrast, **Metallism** proceeds analytically by thinking of money as a creation of private business. In this tradition, the quintessential form of money is some precious metal, not the liability of anyone, and all lesser monies are promises to pay that ultimate money.

Why are there these two traditions? Both trace their origins to the distant past when in fact there were two parallel monetary systems operating at the same time.[2] Within nations, the hand-to-hand currency used by the common people was the king's money (and its credit derivatives); the chartalist tradition grew up to explain how that money worked. Between nations, the money used to facilitate international wholesale trade between businesses was a metallic currency (and its credit derivatives); the metallist tradition grew up to explain how that money worked.

[1] Before the current Euro crisis, some people thought that the Euro might emerge as a challenger to the dollar, and as a consequence certain euro-crosses were for a while considered majors.

[2] See Fernand Braudel, Wheels of Commerce, and especially Boyer-Xambeau, Deleplace, and Gillard Private Money and Public Currencies: The 16th Century Challenge (1994).

	Public/Retail	Private/Wholesale
Money	King's Money (Base)	International Money (Gold)
Credit	Domestic credit	International credit

The exchange rate was fundamentally a relative price between domestic and international currency, both of which circulated side by side. There was no mint par to anchor the value of domestic currency, and yet it had value. Why so? Further, the value of the king's money seemed secularly to fall relative to the international money. Why?

Historically, domestic coinage was the king's prerogative and, given the undeveloped tax system, kings frequently used that prerogative as a source of funds to finance state projects. (Similarly, in more recent history, domestic coinage has been the prerogative of the state, and we have seen in the Young reading how the Civil War was financed in part by greenback issue.) From a chartalist standpoint, the central bank is just a creature of government finance; it is a **government bank**. By monopolizing the issue of currency, the government monopolizes the cheapest source of finance in the nation.

Treasury		Government Bank	
Assets	Liabilities	Assets	Liabilities
Authority, especially taxing authority	Tbills 5%	Tbills 5%	Currency 0%

The government also gains access to a further possible source of finance, which arises from depreciation of the value of the currency, depreciation which diminishes the value of both currency and any debt denominated in that currency that might be outstanding. The central bank is just a way of issuing zero-interest debt, debt that can furthermore be repudiated over time as need be by the simple stratagem of inflation.

For drawing out the full implications of this point of view, it is helpful to take advantage of a certain affinity between the chartalist view of the money standard and the quantity theory of money.

MV=PT Quantity Theory of Money

The quantity theory does not question that the government has the power to assert what is money, but it points out certain limits to that power. If the government issues more money than people can use, any excess just depreciates the value of money. Just so, if we think of the volume of transactions T as determined by the patterns of real trade, and velocity of money V as determined by monetary institutions, then an overissue of money on the left hand side can only show up as an increase in the price level on the right hand side. From this point of view,

inflation ($\Delta P/P$) is caused by and is a symptom of the government's irresponsible expansion of money, an overreaching of state power.[3]

We can expand this point of view to consider international exchange between states by thinking of each country's price level as determined by the quantity of money issued in that country. Then the exchange rate between currencies seems like a relative price linking two essentially valueless currencies. Define S as the number of pounds that exchange for a dollar (so when S goes up that means the dollar appreciates, or becomes more valuable). Then we have the idea that

$SP=P^*$ **Purchasing Power Parity**

or **$\Delta S/S + \Delta P/P = \Delta P^*/P^*$**

In this way of thinking, exchange rates reflect differential scarcity of two fiat monies, and changes in exchange rates reflect differential inflation rates, which are understood as arising from differential money growth rates. This is one way of looking at the problem of exchange rates, and it makes a certain amount of sense in a world of fiat currency.

Exchange rate in a Fiat Money World
Parallel National Hierarchies

Dollar	$S=P^*/P$	Pound
Deposits		Deposits
Securities		Securities

This is a coherent theoretical structure, to be sure, but it doesn't seem to describe real economies very well, at least in the short run. There is a lot of slippage between M and P in the quantitity equation, and even more in the purchasing power parity equation. One way to think about the problem is to observe that price levels move very slowly but exchange rates and monetary quantities fluctuate rapidly. Advocates of the chartalist theoretical tradition therefore argue that the theory is about the long run tendency of the system. But in the long run doubts arise about the direction of causation. Money and prices move together, to be sure, but maybe prices drive money rather than vice versa, especially under modern conditions where most money is actually credit? And maybe exchange rate movements drive domestic prices, rather than vice versa?

I mention these doubts not so much as decisive criticisms, but merely to suggest that we keep our minds open about alternative theoretical possibilities. As further motivation, I want to observe that **PPP thinks of the exchange rate as the relative price of goods, whereas CIP**

[3] The assumption that V and T are independent of monetary expansion leaves out reaction of the money holding public to hold less money (in real terms) if they are expecting inflation. It also leaves out the reaction of issuers of private credit promises to pay the king's money, which promises may serve a close substitutes for money.

thinks of the exchange rate as the relative price of assets. This is the economics view versus the finance view. But neither thinks of the exchange rate as the relative price of money, which is the money view. To get that into the picture, we need to adopt a payments and marketmaking perspective.

A Money View

We start with what Minsky called the "survival constraint", which for our purposes might better be called the "reserve constraint" since it focuses attention on the end-of-day clearing in a multilateral payments system. Every day payments go in and out, but at the end of the day net payments must be settled. If a country has sold more than it has bought, it is a surplus country; if a country has bought more than it has sold, it is a deficit country. The survival constraint is the requirement that deficit countries find a way to settle with surplus countries.

In present conditions, the world reserve currency is generally the dollar so, without loss of generality, we can say that the deficit country needs to acquire dollars. Possibly it has a small reserve holding but more generally it will need to acquire dollars in world foreign exchange markets. Either way, the point to emphasize is how this need to acquire dollars disciplines the behavior of the deficit country. If it cannot acquire the necessary dollars, it will be unable to complete its purchases, and some of its transactions will have to be reversed.

Figure 1 shows a stylized example of how a deficit country might acquire dollars by relying entirely on the **private FX dealing system**. This is only the starting point; the role of central banks will be introduced later.

Figure 1

Surplus Country		FX Dealers		Deficit Country	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
$10 due from					$10 due to
-$10 due from +$10 spot		+$10S FX spot +$10 term	+$10 spot +$10S FX term	+$10 spot -$10S FX spot -$10 spot	-$10 due to
		+$10S FX term	+$10 term		

The first row shows the net positions of the two countries before settlement.

The second row shows how the FX dealer system facilitates settlement by creating credit, specifically a spot dollar liability which we suppose the deficit country buys from the dealer at the spot exchange rate using local currency, and then transfers to the surplus country. The consequence is expansion of the dealer's balance sheet on both sides, expansion that exposes the dealer to exchange risk, namely the risk that the dollar price of its new FX asset might fall.

As a hedge against this price risk, the second row further shows the dealer entering into an offsetting forward exchange contract by borrowing term FX and lending term dollars, taking its cue from the Covered Interest Parity condition. (Taking our own cue from CIP, we adopt the convention of booking forward transactions as a pair of term credits, lending in one currency and borrowing in another.) At the end of the day our FX dealer has "matched book"—if the dollar price of its new FX spot asset falls, then so also will the dollar value of its new FX term liability. It does however still face liquidity risk since maintaining the hedge requires rolling over its spot dollar liability position until maturity of its term dollar asset position.

The third row shows the position of a second "speculative" dealer who provides the forward hedge to the first dealer. This second dealer does not have matched book and so faces exposure to exchange risk, but in the forward market not the spot market. (In practice he might himself hedge with a futures position, or an FX options position, but that doesn't eliminate the risk, only shift it to someone else.) In effect this second speculative dealer is engaged in a "carry trade", paying the dollar interest rate and receiving the FX interest rate. If the realized spot rate is different from the forward rate, this speculation will make a profit or a loss.

I have been treating these two dealers as private agents, but in fact they might be central banks, and in general both private dealers and central banks are involved. In major currencies, the action is mostly with private dealers. In minor currencies it is mostly with central banks. We will expand on this point in future lectures.

Private and Public Money: From Parallel Systems to Integrated Systems

I have presented two traditions in monetary thinking. Metallism arose from attempts to understand the private international money system; Chartalism arose from attempts to understand the public domestic money system. Let me now observe that neither of these traditions seems to be very well lined up with the current facts, as I have described them. **The modern system is not a parallel money system but rather a hybrid or integrated system.** How did this come to pass?

Historically, wars have provided the opportunity for kings to discover the usefulness of the private money system. During wars, governments face a dire need for funds, and typically the kind of funds they need are metallic not fiat, in order to pay for crucial war materials. The private metallic system can provide gold since it runs on metallic reserves.

Suppose, as was usually the case, there is a banker's bank separate from the government bank, and suppose such a bank issues deposits against a 100% gold reserve. (Think of a clearinghouse, which we have already seen is a kind of primitive central bank.) I am showing

here parallel money systems, although it could be said that the private gold money is better, i.e. higher in the hierarchy.

Prewar Parallel Banking Systems

Government Bank		Private Banker's Bank	
Assets	Liabilities	Assets	Liabilities
Treasury bills	Domestic Currency	Gold Private Credit	Deposits

Now consider what happens when the government borrows from the private bank and takes payment in gold.

Wartime Hybrid Banking System

Treasury		Government Bank		Private Bank	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+Gold	+Loan	Treasury bills	Domestic Currency	-Gold +State Loan	Deposits

Now consider what happens when the government repays the loan in currency and bills (not repayment really, just refinancing), holding onto gold as the international reserves but requiring banks to use its own currency as domestic reserves.

Postwar Hybrid Banking System

Treasury		Central Bank		Private Bank	
Assets	Liabilities	Assets	Liabilities		Liabilities
-Gold	-Loan +TBills	Gold +Gold +TBbills	Currency +Currency	+Currency +Treasury	Deposits

The effect is to put the government bank at the top of the hierarchy, while lower down the private bankers gain the right to issue forms of money that compete with the king's money. It is a win-win. The hybrid structure brings more credit to the king for war, and then more credit to the economy for peacetime expansion.

We have seen in Lecture 2 how Civil War finance created the national banking system, but not yet any national central bank. Indeed what passed for central banking took place in the

New York Clearinghouse, a private banker's bank. So the hierarchy remained with the private money on top and the public money below. The establishment of the Federal Reserve system in 1913 switched that hierarchy, putting public money on top in the domestic banking system. That proved enormously useful for the war finance that followed, but it stored up problems once the war was over because the Federal Reserve System left private money on top in the international gold standard banking system. This was a contradiction that would take another hundred years to work out, and it is not completely worked out even today.

14. Money and the State, International

Mundell puts forward a fascinating and provocative way of understanding the 20th century as one long excursion away from the gold standard and back again, or at least moving in that direction. For our purposes, the interesting thing about Mundell is the way his perspective expands the money view in the direction of international money.

Most important, the discipline that is imposed by the gold standard is a payments discipline; countries settle payments ultimately in gold, although there are myriad opportunities (using money markets) to delay net settlement for a while in anticipation that some offsetting payment will arrive meanwhile. This discipline not only polices the behavior of individual nations, but also knits the entire collection of nations into a more or less unified, integrated, and stable moneyflow system.

The problem is that individual states, and their central banks, may not be willing to submit to the discipline. If they could coordinate with other states, either to revalue gold (increase its price) or to create accepted substitutes for gold (such as the SDR), they could relax the discipline somewhat (more gold in circulation relative to a given quantity of credit). But in practice they have not been able to coordinate because individual states, and their central banks, are also generally not satisfied with the place assigned to them in the system, and so always looking to improve their position.

As a consequence, the 20th century monetary experience has had not much to do with gold and much more to do with the behavior of the dominant nation (the US) and its central bank (the Fed.) Mundell blames monetary mistakes for most of the travails of the twentieth century, not only deflation and inflation but also wars, social unrest, communism and Naziism.

I proceed by recasting Mundell's account of the 20th century in the analytical framework of this course, namely the natural hierarchy of money.

Act 1 (1900-1933): "Confrontation of the FRS with the Gold Standard"

The Federal Reserve System was established in 1913 on the eve of WWI. Only the US remained on gold during the war, but that was no great feat because throughout the war gold flowed in to pay for war material. At the end of the war, most of the world's gold was in the US, and the international monetary system was comprised of national currencies, none of which was properly convertible into a higher level international money. Nevertheless, the decision was made to move in the direction of a return to the international gold standard which had been the system before the war.

The problem was that almost no one realized, and certainly no one prepared for, the deflationary consequences of doing so without revaluing gold. Over the next decade, the value of gold would increase each time another country that returned to gold, simply because of the increased monetary demand for gold, and the consequence would be downward pressure on other prices (quoted in gold) throughout the world. (France was a partial exception because France devalued its own currency against gold, and so got a fresh start.)

International Money	Gold			
	|			
National Money	Dollar	Pound	DM	Franc
	|			
Credit	Deposits	Deposits	Deposits	Deposits

Add to this structural deflationary tendency some mismanagement by the Fed, and you got the ingredients for a real disaster. Mundell identifies the key mistake in 1931: "Instead of pumping liquidity into the system, it chose to defend the gold standard" (p. 330) by raising the rediscount rate from 1 ½ to 3 ½ percent. Thus the Fed compounded the secular deflationary tendency with explicitly deflationary policy.

In terms of our hierarchy, we can think of the Fed as defending the position of the dollar relative to the purportedly better international money gold. This was the wrong thing to do on two accounts. First, the whole point of raising the rediscount rate was to attract gold, but that just increased the upward pressure on the price of gold, which means the downward pressure on prices generally, and on other currencies. In the event the pressure was too much for the weaker currencies to stand, and they simply abandoned gold, as did the United States itself eventually, and the end result was the collapse of the international monetary system, and along with it much of international commerce.

Second, by defending the dollar against gold, the Fed was in effect leaving domestic banks on their own to defend the par value of bank deposits against the national dollar. Falling agricultural prices meant widespread loan default, and tight central bank policy added liquidity pressure. The result was widespread domestic bank failure, which added a further impulse to domestic deflation.[1] On both accounts, Mundell suggests that the right policy would have been instead to engineer a worldwide revaluation of gold, which means a worldwide depreciation of currencies against gold. "Had the price of gold been raised in the late 1920s, or, alternatively, had the major central banks pursued policies of price stability instead of adhering to the gold standard, there would have been no Great Depression, no Nazi revolution, and no World War II."

Act 2 (1934-1971): "Contradiction between Keynesian <u>national</u> management and the Bretton Woods <u>fixed rate</u> system"

At the end of WWII, at the famous Bretton Woods conference, gold was once again in the US, and the international system once again involved a collection of national currencies that were not convertible into any higher international money. This time, unlike WWI, there was no thought of trying to return to the gold standard. Instead there were several proposals for fundamental reform, most notably Keynes' bancor plan which would have created an <u>elastic</u> international money.[2]

[1] See Irving Fisher, "The Debt Deflation Theory of great depressions" (1933)

[2] Also noteworthy, the proposal for replacing gold with a basket of commodities. See Benjamin Graham, World Commodities and World Currencies, 1944.

Keynes Bancor Plan

Deficit countries		International Bank		Surplus Countries	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	Bancor loans	Bancor loans	Bancor Deposits	Bancor deposits	

A key feature of this plan was an attempt to create symmetry between the deficit and surplus countries, which means weakening the discipline of the survival constraint which binds on the deficit countries but not the surplus countries. Countries with large and persistent surpluses were to be penalized, and so provided with incentives for spending the surpluses (on goods or capital from the deficit countries) rather than saving them. At the time, the US was obviously going to be the only surplus country, as everyone else rebuilt from wartime damage, so the US did not like this feature and instead put forward a plan that fixed the quantity of international money once and for all, the IMF.

International Monetary Fund		Member Country Central Bank	
Assets	Liabilities	Assets	Liabilities
+Gold +National Currencies	+SDR	-Gold +SDR	+Currency

This rather harsh <u>discipline</u> at the international level might have caused problems (similar to those caused by the inelasticity of money under the National Banking System in the US, see Lec 3) but in fact there was an element of elasticity added elsewhere. The dollar was convertible into gold, and all other currencies were convertible into the dollar, the so-called "anchored dollar" system, but the quantity of dollars was, in principle, elastic because in effect the US (the entire US, not just the Fed) was the world's bank. This point is not so much emphasized by Mundell, but was a constant theme in the work of his teacher, Charles Kindleberger.[3]

United States		Rest of World	
Assets	Liabilities	Assets	Liabilities
Long term bonds Gold reserves	Short term $ deposits	Short term $ deposits	Long term Bonds

In these balance sheets we see an expansion of dollar reserves for the rest of the world that does NOT require the US to be running a trade deficit. Instead they are just the liability counterparts of gross capital outflows to finance the redevelopment of the rest of the world. So long as the rest of the world accepts dollar deposits as if they were gold, the system works fine. But it comes under unsustainable pressure if gold comes to be undervalued, so that holders of dollars ask for the promised convertibility. Over time, in the 1960s, this started to happen with increased frequency.

[3] Depres, Emile, Charles P. Kindleberger, and Walter S. Salant. "The dollar and world liquidity" The Economist (Feburary 5, 1966): 526-29.

Gold Anchor	Gold
International Money	Private Dollar ($35/oz.)
National Money	Dollar Pound DM Franc

One way we might have fixed the problem would have been to revalue gold; Mundell says that Arthur Burns was trying to get Nixon to do that. Another way was to increase the supply of gold, or rather of SDR (Special Drawing Rights) fiduciary issue of "paper gold" by the IMF. This was attempted in 1967 but it was too little too late. Mundell suggests that if more had been issued, so that central banks could have substituted SDRs for gold reserves, that would have reduced worldwide demand for gold, so lowering its price relative to the dollar, and taking the pressure off. Instead, the pressure continued and in 1971 the US simply went off gold. (Recall that one response to the present global financial crisis was to have the IMF issue additional SDRs.)

Act 3 (1972-1999): Flexible exchange, learning from experience

When the US unilaterally broke the connection with gold, even the weak discipline of the gold anchor was lost. The rationale for other currencies to peg to the dollar was therefore lost, and they broke away. Instead of a hierarchical system, we got a system of national currencies and floating exchange rates. A system of national currencies is typical in war time, when commerce is typically severely restricted, but quite anomalous in peace time.

International Money	???
National Money	Dollar Pound DM FF

Some economists (most notably Milton Friedman) persuaded themselves that flexible rates would actually work better, on the grounds that market prices tend to work better than administered prices.[4] Perhaps he thought that price stability domestically would inevitably result in exchange rate stability internationally, through the mechanism of Purchasing Power Parity. The result however did not bear out these hopes.

The result was inflation and stagnation simultaneously until the 1980s when we began to get some discipline into the system again, and ordering around a few key currencies. The discipline came from commitment by central banks to a regime of inflation targeting domestically, even absent any international disciplining device (such as gold convertibility). Such regimes were very successful in producing price stability. But internal price stability did not translate into international exchange stability, not at all. Volatility of major rates—dollar, euro, yen—continued even after reduced inflation in each area.

One important legacy of the period of disorder was the rise of currency futures markets for hedging exchange rate risk, and these markets remain to this day extremely important. We will be talking at length about these markets next week. Notwithstanding these markets, the

[4] It is clear that Mundell does not agree with Friedman on this point. Mundell talks in a number of places about the advantages of fixed exchange rates. Just as par clearing is an important institution for knitting together the farflung states of the United States, so too fixed exchange is an important institution for knitting together the farflung countries of the world.

continuing problem was persistent volatility of exchange rates as well as continuing lack of any truly global currency.

Mundell doesn't talk much about speculative currency markets, maybe because he sees them as eventually being replaced by a fixed exchange system. For him, the great example of success, in this regard, is the emergence of the Euro in 1999, which eliminated exchange rate volatility in a large portion of the system, leaving the system organized around three main currencies: dollar, Euro, yen. Volatility however remained between these currencies.

International Money	???		
Key Currencies	Dollar	Euro	Yen
National Currencies			

Ten years later, we see that his optimism seems misplaced. One of the things he misses is that the three-currency picture he paints is not so symmetric. **The world funding market is basically a dollar funding market, not euro or yen.** The dollar is very much the dominant player, as we saw in the recent global financial crisis.

Act IV: Global Financial Crisis

During the financial crisis, one of the big problems was a shortage of dollar funding for dollar assets held outside the US. When the system was working, there was a kind of unregulated international dollar market that worked like this:

Global Shadow Bank		MMMF	
Assets	Liabilities	Assets	Liabilities
RMBS	MM Funding	MM Funding	"deposits"

When the crisis happened, the Money Market funding all dried up as MMMF refused to roll their loans to the shadow banks, and demanded instead high quality money market assets such as Treasury bills. In the aftermath of Lehman and AIG, the funding problem was fixed temporarily by means of a liquidity swap between central banks: the Fed lent dollars to foreign central banks, which then lent them on to the shadow banks located in their countries. Thus the central bank network substituted for the collapsing money markets.

Here is a stylized picture of the operation in its full glory, using the European Central Bank as my example[2]:

Shadow Bank		ECB		Fed		Treasury		MMMF	
A	L	A	L	A	L	A	L	A	L
MBS	-MM funding	+$ dep.	+€ dep.	+€ dep.	+$ dep.			-MM funding	"dep"
	+$loan	-$dep. +$loan				+$ dep.	+Tbill	+Tbill	

The first line shows the withdrawal of MMMF funding of the shadow bank, and the beginning of a solution in the central bank liquidity swap. The second line shows ECB providing dollar funding to the shadow bank, and the MMMF investing in newly issued Tbills, the proceeds of which the Treasury deposits at the Fed. So you can see how ultimately it is still "deposits" at the MMMF that fund the Shadow Bank holding of Mortgage Backed Securities. But now there is no direct lending from MMMF to the Shadow Bank. Rather the MMMF lends to Treasury, which lends to Fed, which lends to ECB, which lends to the Shadow Bank.

That was how we managed lender of last resort internationally in the heat of the crisis. In the aftermath, all this temporary construction got dismantled, essentially by taking MBS back to the US where it was easier to cobble together dollar funding directly. It is no accident that there is now over a trillion MBS on the balance sheet of the Fed.

Fed		Banks		MMMF	
A	L	A	L	A	L
MBS	Reserves	Reserves	Deposits or other funding	Deposits or other funding	"deposits"

Act V: Global swap network

As of October 2013, the Fed has signed permanent and unlimited swap lines with five major central banks: ECB, Bank of England, Bank of Japan, Bank of Canada, Swiss National Bank. So now this emergency backstop is present all the time. So far it has not been needed, but that's mainly because private markets are more willing to do the job, knowing that the backstop is in place.

[2] For more detail, see contemporaneous explanation in BIS Quarterly Review.

15. Banks and Global Liquidity

Today and next time, we pursue an adaptation of our dealer model to the foreign exchange market under a gold standard. In lecture 10 "The World that Bagehot Knew" we talked about a world of bills of exchange, discounting banks, and the central bank as liquidity backstop. We start our discussion of international money there, by specifying that the firms issuing and accepting bills are both outside England, while the discounting banks and the central bank are inside.

Surplus Firm/Country		City of London Banks		Deficit Firm/Country	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
-goods +bill -bill +notes		 +bill -notes		+goods	+bill
		  -bill +notes		-goods +notes -notes	 -bills

I'm showing here the surplus firm discounting bills for sterling pound notes (closer to reality would be deposits, but intuition is sharper with notes), and by assumption that is not his own domestic currency. Similarly, I am showing the deficit firm paying off the bill with sterling pound notes, and by assumption that is not his domestic currency either. So we've got the surplus firm with sterling it has to exchange for domestic currency, and the deficit firm with domestic currency it has to exchange for sterling. These are two different currencies, and the exchanges happen at two different times.

This is a job for a foreign exchange dealer who makes two-way markets in the various currencies, by quoting prices and absorbing the resulting order flow on his own balance sheet. Here are the two dealers, and I am also showing the Bank of England:

Surplus Dealer/Bank		Bank of England		Deficit Dealer/Bank	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+notes/gold -domestic currency		 +/-gold	 +/-notes	 -notes/gold +domestic currency	

At the inception of the trade, the surplus dealer is creating domestic currency and building up inventories of international reserves. At maturity, the deficit dealer is destroying domestic currency and drawing down inventories of international reserves. And at any point in between, either dealer may decide he'd rather hold reserves in gold, or the other way around, and go to the Bank of England for that purpose. So we have three markets here.

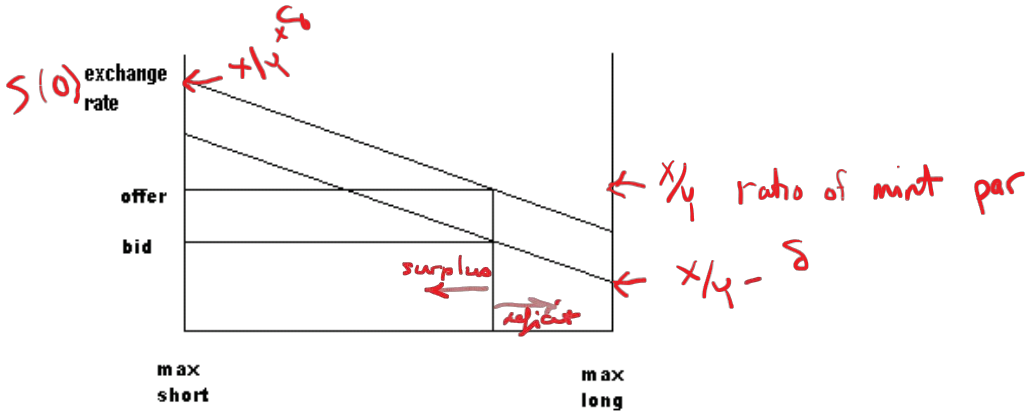
Dealer Model

So long as we are on a gold standard, the analysis is fairly straightforward.
There is a mint par, defined as a quantity of gold that private agents can take to the mint and get pounds. But, because there is some cost of transporting gold abroad, the exchange rate for pounds can move a bit away from the mint par, on either side, without creating incentive to convert pounds into gold. These "gold points" are the outside spread, established by the central bank, within which the private dealers make markets, establishing the inside spread.

I remind you of the notation introduced in Lecture 13:
Dollar = x oz. of gold, mint par
Pound = y oz. of gold, mint par
Exchange rate S = £/$ = x/y, ratio of mint pars, so actual exchange rate fluctuates between gold points:

$x/y - \delta < S < x/y + \delta$

In a world like this, dealers are willing to add to their inventories of foreign exchange if they can get them at a good (cheap) price, but once that price falls to the gold point they are unwilling to add any more. At that point, anyone who wants to sell foreign exchange must sell it not to the private dealers but rather to the central bank who pays gold (mint par) for them.



In the diagram I am showing the economics of a dealer who is taking on liquidity risk by shedding international reserves (pounds) and holding instead dollars. This is a position like our deficit dealer. He is willing to bear a certain amount of liquidity risk, but when it gets too much, so exchange rate falls to the gold points, he stops and the central bank steps in.

Central Bank defense

In a world like this, consider how the US central bank that is committed to maintaining a fixed exchange rate against the pound would actually do it. Upward pressure on the currency can be met by issuing the currency to buy foreign reserves. No problem. But downward pressure (same as upward pressure on the price of gold) must be met by buying currency. Big problem.

There are three basic channels toward this end.

First, the central bank can buy in domestic currency with its holding of international reserves. (Note, this is exactly what the deficit dealer was doing in the example above.)

Assets	Liabilities
-$100 gold reserves +100 currency -100 currency	 -100 currency

The net effect is a contraction of the balance sheet, since the currency purchased is the central bank's own liabilities. But this works only so long as the bank has reserves.

Second, he could sell some assets, presumably the most liquid ones, i.e. the public debt. In this case defending the currency looks like this.

Assets	Liabilities
-$100 domestic credit (Tbills) +100 reserves -100 reserves	 -100 currency

Now we have contraction of the balance sheet, but also possible downward pressure on Tbill price, hence upward pressure on interest rates, perhaps very large upward pressure if speculators fear devaluation. We often hear loose talk about central banks defending their currency by raising interest rates, and now we understand better what really happens; central banks offer interest bearing securities in exchange for currency, in the hope that the interest will prove sufficient incentive to prevent asset holders from demanding payment in international reserves.

Third, if the central bank has no more reserves and is unwilling to push interest rates higher, one further possibility is to rely on lines of credit with other central banks, (or with the IMF which, in

this respect, operates somewhat like an international central bank). These borrowed reserves then are sold for dollars.

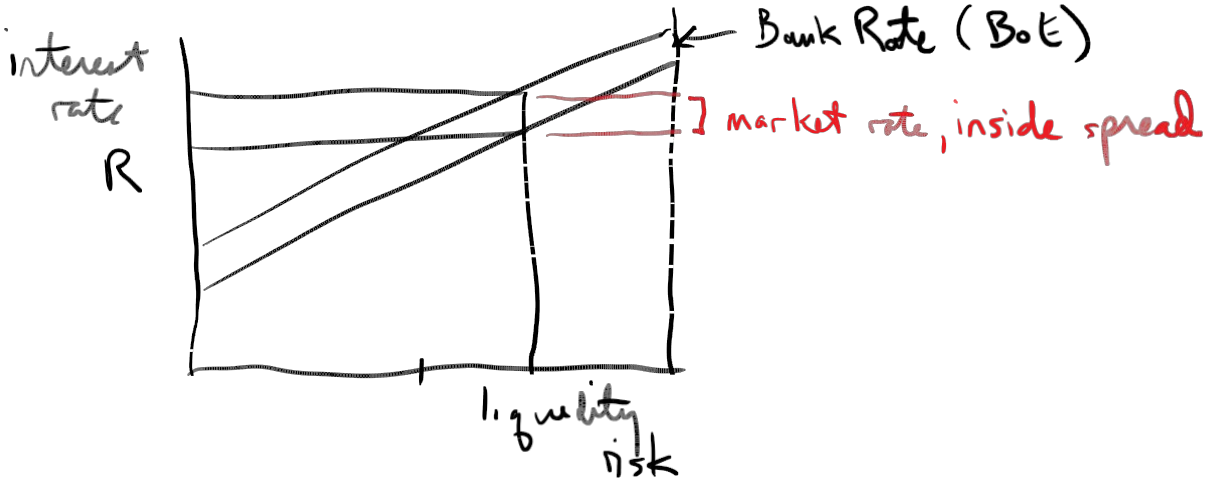
Assets	Liabilities
+100 reserves -100 reserves	+100 borrowed reserves -100 currency

In the end, this third channel amounts to a portfolio adjustment on the liability side. **More generally, the point of these examples is to make clear in what sense contraction of the money supply supports the international value of the dollar.** The most straightforward way of understanding what happens is that the US central bank enters the market as a buyer of dollars, but since dollars are its own liability, its purchase of dollars is in effect a contraction of its balance sheet.

As always it is the deficit entity that is forced to adjust, but the deficit entity is not necessarily below the Bank of England in the hierarchy. In some cases, the deficit entity could be the Bank of England itself, if the pound drops to the gold points (or equivalently, in our dealer diagram, the dollar rises to the point where gold flows from the Bank of England in defense of its mint par). Let's explore that possibility for a while.

City of London and Bank of England Defense

Put yourself in the shoes of the City of London banks. Their global discount business involves daily note outflow, and also note inflow as discounts reach maturity. Each individual bank is watching the balance between outflow and inflow, and adjusting its discount rate in order to bring them into line. Higher rate discourages new discount, so stemming outflow for a while until inflow restores liquidity. All banks are doing this, and competition between them establishes the market rate of interest, which fluctuates over time depending on the aggregate note outflow and inflow between the banking system and the rest of the world.



We have seen in previous lectures how to understand this discount system using the dealer model. Banks are willing to take on additional liquidity risk, by continuing to discount even when note reserves are falling, if they get compensated by a higher interest rate. But at some point, they have enough, market rates rise to bank rate (the discount rate quoted by the Bank of England) and the Bank of England takes over. It may discount (or rediscount) bills and pay out notes, just like the private banks. But it may also discount bills by creating deposits, which are promises to pay notes, insofar as banks are willing to accept deposits as substitutes for notes.

Bank of England: Banking Department

Assets	Liabilities
+bills -notes +bills	 +deposits

We have seen in previous lectures how this kind of lender of last resort activity—lending freely but at a high rate against good security—can stem an "internal drain", meaning a net outflow of notes from the banking system into circulation. The important point to emphasize now is the limits on such activity when the problem is an "external drain", meaning a net outflow of notes to the rest of the world. The rest of the world doesn't want notes, it wants gold, so if it finds itself with excess notes it just takes them to the issue department of the Bank and exchanges them:

Bank of England: Issue Department

Assets	Liabilities
-gold	-notes

You can see from the balance sheets that the Bank of England now faces exactly the same problem that our deficit country central bank faces. It can pay out gold for a while, but reserves inevitably fall short. The next steps are

--raising Bank Rate, so discouraging the world from bringing bills to London for discount, and allowing repayment of maturing bills to restore gold reserves. The problem with this policy is that it inevitably raises domestic interest rates as well, even though domestic credit was not the source of the problem.
--borrowing from other central banks, who may have larger gold reserves
--suspending specie payments.

This latter is the "nuclear option" but the Bank of England historically was forced to resort to it on numerous occasions. Obviously once the Bank of England suspends gold payments, the gold point story about determination of exchange rates no longer operates. So what is the story then? That will be our question for next time. You will see then why I have been putting so much emphasis on Covered Interest Parity

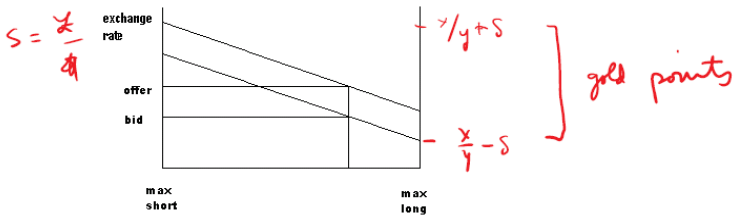
CIP: $[1+R^*(0,T)]S(0) = [1+R(0,T)]F(T)$

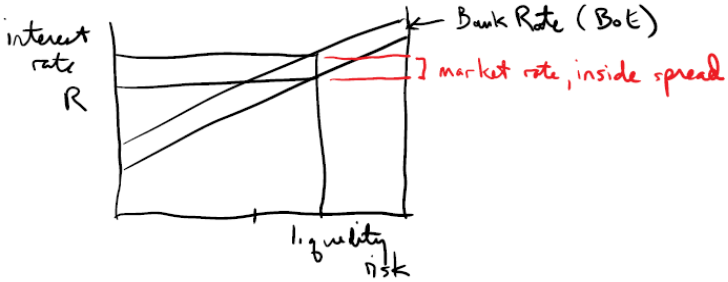
Meanwhile, in summary, observe how important central banks are for the operation of the gold standard. They establish two critical outside spreads that provide bounds on the system within which profit maximizing dealers can operate to make markets. One bound is on the exchange rate, i.e. the gold points. The other is on the interest rate, i.e. Bank Rate, which is a term interest rate. In modern exchange rate systems, neither of these bounds is effective. Nonetheless, as we will see next time, the analysis we have just been through can be adapted to explain how things work.

In closing, I refer you to a passage from Keynes (1924) Tract on Monetary Reform, where he refers to the gold standard as a "barbarous relic", and emphasizes the role of Ministers of Finance in discovering how to operate national currencies without the discipline of gold. He urges them to focus their efforts on stabilizing the domestic price level, hoping that such stability will also (through purchasing power parity) stabilize exchange rates. We know now, from our own experience, that price stability is not enough, as Mundell emphasizes. That means that we need a theory of exchange that does not come from purchasing power parity, but rather from somewhere else, i.e. CIP.

16. Foreign Exchange

Last time we introduced two new Dealer diagrams in order to help us understand our third price of money, the exchange rate, but under the special conditions of the gold standard.





In both cases we are treating the pound as the world reserve currency, and quote exchange rates as pounds per dollar, so lower exchange means depreciation of the dollar against the pound. The bank rate we discussed was the Bank of England's bank rate at the center of the system.

Today we build on that analysis to develop a more general theory of exchange rate determination without a gold standard anchor.

In the first lecture of this series, I contrasted metallist and chartalist explanations, and suggested that the money view involves a third approach that somewhat melds the two, but goes beyond them. Figure 1 shows the stylized example I used then of how a deficit country might acquire dollars by relying entirely on the **private FX dealing system**. This is only the starting point; the role of central banks will be introduced later.

Figure 1

Surplus Country		FX Dealers		Deficit Country	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
$10 due from					$10 due to
-$10 due from +$10 spot		+$10S FX spot +$10 term R	+$10 spot +$10S FX term R*	+$10 spot -$10S FX spot -$10 spot	-$10 due to
		+$10S FX term R*	+$10 term R		

The first row shows the net positions of the two countries before settlement.

The second row shows how the FX dealer system facilitates settlement by creating credit, specifically a spot dollar liability which we suppose the deficit country buys from the dealer at the spot exchange rate using local currency, and then transfers to the surplus country. The consequence is expansion of the dealer's balance sheet on both sides, expansion that exposes the dealer to exchange risk, namely the risk that the dollar price of its new FX asset might fall. Note that we are quoting exchange rates as units of FX per dollar, so that a lower number means dollar depreciation or FX appreciation, and that R* is the interest rate on FX term deposits.

As a hedge against this price risk, the second row further shows the dealer entering into an offsetting forward exchange contract by borrowing term FX and lending term dollars, taking its cue from the Covered Interest Parity condition.

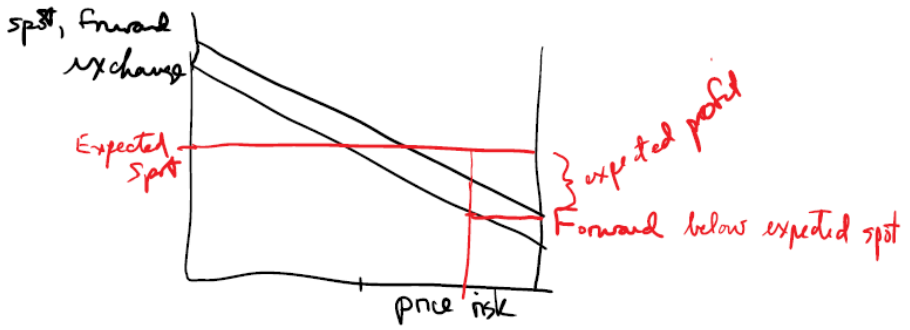
CIP: $[1+R^*(0,T)]S(0) = [1+R(0,T)]F(T)$

Taking our own cue from CIP, we adopt the convention of booking forward transactions as a pair of term credits, lending in one currency and borrowing in another. At the end of the day our FX dealer has "matched book"—if the dollar price of its new FX spot asset falls, then so also will the dollar value of its new FX term liability. It does however still face liquidity risk since maintaining the hedge requires rolling over its spot dollar liability position until maturity of its term dollar asset position.

The third row shows the position of a second "speculative" dealer who provides the forward hedge to the first dealer. This second dealer does not have matched book and so faces exposure to exchange risk, but in the forward market not the spot market. (In practice he might himself hedge with a futures position, or an FX options position, but that doesn't eliminate the risk, only shift it to someone else.) In effect this second speculative dealer is engaged in a "carry trade", paying the dollar interest rate and receiving the FX interest rate. If the realized spot rate is different from the forward rate, this speculation will make a profit or a loss.

Economics of the Dealer Function

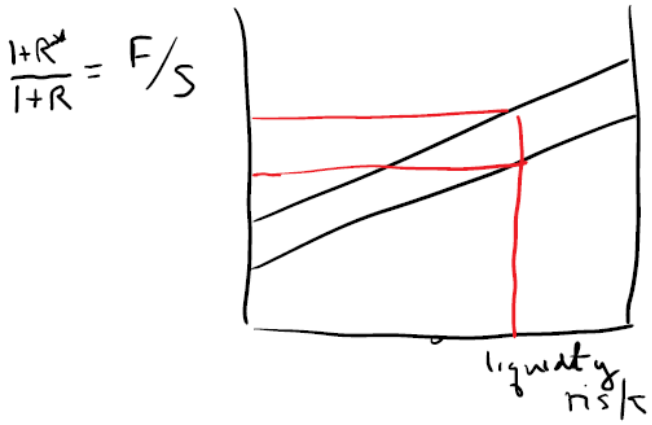
Now think about the economics of the dealer function, and let's start with the speculative dealer. He is concerned about price risk. He will be willing to shoulder more only if the forward rate for FX depreciates relative to expected spot, so he has a larger expected profit. I am showing this depreciation as a downward sloping forward curve (quoting FX forward as 1/F, and FX spot as 1/S)



Notice how the economics of the speculative dealer function require violation of UIP, Uncovered Interest Parity. The expected profit that lures the dealer into a naked forward position is entirely about deviation between forward and expected spot.

Notice also how the expected spot plays the role, in a flexible exchange rate system, that the mint par ratio plays in a gold standard system. Deviations from mint par create opportunities for dealer profit, so long as there is some expectation that the system will at some time return to mint par—mint par is a kind of long run expected spot. In a flexible exchange rate system, there is not so much of a long run anchor for expected spot. Indeed, if people believe UIP then they believe that forward rates are unbiased forecasts of future spot rates, so depreciation of forward rates which arises to encourage dealers to absorb imbalances can easily be interpreted instead as a sign that future spot rates will also involve depreciation. This is a source of potential volatility in the flexible exchange rate system.

The matched book dealer faces dollar liquidity risk; the larger his book, the larger the risk. He is buying FX spot and selling FX forward. He is willing to enlarge his book, and hence his liquidity exposure, but will insist on buying spot relatively more cheaply than forward, so F/S increases with the size of the book. (S goes up, but F goes up more; spot depreciates but forward depreciates more.)



The first thing to notice now, by contrast to the gold standard case, is that the outside spread is not set by the central bank, at least not directly. Modern central banks typically operate on the overnight interest rate, not the term rate. Supposing the US and foreign country maintain constant overnight interest rates, that still leaves scope for payments pressure to move around term rates. But by CIP, F/S=(1+R*)/(1+R), so as positions build up term interest rates have to move apart as well. Assuming that the dollar rate is unaffected, that means FX term rates rise.

Central banks are involved indirectly because of the expectations hypothesis of the term structure, which states that term rates should be the same as the expected result from rolling over a series of short term deposits. EH fails empirically, of course, but that is because the implied arbitrage (borrowing short and lending long) exposes to liquidity risk. Our FX dealer is borrowing short and lending long in dollars, hence exposed to liquidity risk, and will be willing to do more of it only if he is compensated. One way to compensate him is to keep overnight rates low even as term rates rise.

Aside: Note that we have here a story that makes sense of the failure of EH and UIP both. They are both a consequence of the need to offer dealers expected profit in order to make liquid markets.

Central Bank Backstop

But there is a limit to this. If private dealers hit their limit, and the central bank is still not willing to let the overnight rate go, the central bank itself can get into the speculative dealer business, selling the hedges that the matched book dealers want. And if that isn't enough, the central bank can get into the matched book business by doing liquidity swaps with other central banks. **The point to emphasize is the way that central bank's commitment to target overnight rates implies FX intervention whenever the F/S ratio moves to the extreme end of the possible range. Simply put, central banks that target overnight interest rates are inevitably drawn into serving as FX dealers of last resort.**

Figure 4 shows the limiting case in which the deficit central bank borrows reserves from the surplus central bank for a certain term at interest rate R^g, trades those reserves with the private citizen at the spot rate s^g, and then sterilizes the consequent domestic monetary contraction by buying a domestic Treasury bill. The surplus central bank, for its part, lends term reserves by creating a spot deposit to the credit of the deficit central bank, which winds up in the hands of its own private citizen, and then sterilizes the consequent domestic monetary expansion by selling a domestic Treasury bill. (For simplicity, we do not show explicitly the private citizen counterpart to either of the sterilization operations.)

Figure 4

	Surplus Country		Deficit Country	
	Assets	Liabilities	Assets	Liabilities
Private Citizen	\$10 due from -\$10 due from +\$10 spot		-10S^g FX spot +\$10 spot -\$10 spot	\$10 due to -\$10 due to
Central Bank	+\$10 term, R^g -\$10 Treasury bill	+\$10 spot -\$10 spot, CB +\$10 spot, PC -\$10 spot	+\$10 spot -\$10 spot +10S^g FX spot +10S^g FX Tbill	+\$10 term, R^g +10S^g FX spot

The first point to emphasize here is that, at the end of the day, the deficit country central bank is borrowing dollars term and lending FX term, which is essentially a naked forward position analogous to the position our speculative private dealer was induced to take by the expectation of private profit (compare Figure 1). But the central bank need not, and probably does not, expect to profit from its speculation. For one, the interest rate at which it borrowed from the surplus central bank need not, and probably does not, match the term funding rate in private credit markets. And the spot rate at which it

sold dollars to its own citizen need not, and probably does not, match the spot rate in private FX markets. These are both, at least potentially, policy rates and as such can be expected to reflect the non-commercial relationship between one central bank and another internationally, and the non-commercial relationship between the central bank and the needy private citizen domestically.

The second point to emphasize is that, at the end of the day, the surplus country central bank is lending dollars term to a foreign central bank, instead of to its parent government. This is very unlikely to be a move about which the parent government is neutral, hence our characterization of it as a limiting case. But other counterparties, other than the surplus country central bank, would also work. There are sources of reserves other than the Fed, sources such as regional liquidity pooling arrangements or the International Monetary Fund which do not depend on the goodwill of the surplus country's government.

Even more, once it is recognized that deficit country dealer of last resort essentially involves willingness to take on a naked forward position when no one else will, it becomes clear that the whole operation need not involve another central bank as counterparty at all. The deficit country central bank could, if it so chose, instead facilitate private matched-book dealing by serving as the speculative dealer to enable forward hedging of spot exposures. Or it could go even farther, facilitating the term dollar borrowing of its own private citizens by directly offering them forward hedges, so taking their exchange risk onto its own balance sheet. Again, all of this is policy, so pricing of these forward contracts need not be, and probably will not be, on commercial terms. The limiting case along these lines comes when central banks offer forward cover to all comers at the same rate as current spot exchange, in effect fixing the exchange rate as a matter of policy.

The danger, of course, whenever you offer to trade with all comers at non-commercial prices, is that in doing so you offer arbitrage opportunities for speculators.[1] The positive case for doing so must therefore rest on an argument that commercial prices are in some sense wrong. We have seen (in section II) that one consequence of the view that the exchange rate is the relative price of money is to draw attention to the way that order flow can push prices around for reasons that are not fundamental. The potential problem is that private agents take these distorted prices as parametric for their economic decisions, and so make distorted decisions.

There is thus potentially an argument for central bank intervention, certainly in extremis to prevent breakdown of the payments system, but also in less extreme situations where, for one reason or another, private markets are not making markets at all, or only reluctantly doing so at the cost of substantial price distortion. But it is an argument that applies only under specific conditions, not universally. The exchange rate is not a free variable left open for state choice, but neither is it a market price that always fully reflects fundamental valuation.

[1] Just so, DeRosa (2009, p. 80): "…all these crises were preceded by the accumulation of substantial at-risk positions that were short the U.S. dollar and long local currency. When the crises occurred, the entire market, not counting the central bank, had to buy dollars and sell the local currency immediately in order to hedge."

Renminbi

McCauley reports on the state of play in terms of the internationalization of the renminbi. China has very strict controls over most of its money and credit system (track one) with some periphery allowed to price competitively (track three). What is happening now is a third track, basically centered in Hong Kong, both money markets and capital markets. It is clear that McCauley has in mind as an analogy the evolution of the Eurodollar market, which took the US authorities by surprise when it happened, but then they allowed it. Today world funding markets are dollar funding markets.

That's a long way away for the renminbi probably. The first thing will be breaking down the barriers between track three and track two, and then also track one. Central bank control will in the course of this shift from a gold-standard like price setting function to a modern price backstopping function.

The concern in all of this is exchange rate instability. The key player in everything is the speculative dealer who is willing to take a position only if the forward rate is lower than the expected spot. But what anchors expectations about expected spot? Depreciation of forward looks like a forecast of future spot, and if that forecast leaks into expectations of actual future spot, then the whole system can get unhinged.

17. Direct and Indirect Finance

Today we begin our fourth section of the course, which extends the money view to capital markets and asset prices.

It is common practice to treat capital markets in finance courses, and money markets in banking courses. But in the real world the two markets are quite completely integrated, most obviously in the shadow banking system, which I define as money market funding of capital market lending. Thus the intellectual habit to treat them separately is outdated, but also more than a little misleading and even dangerous. It is important however to appreciate that the tight integration of the two has historical and institutional roots.

In Bagehot's day, arguably, the two were much more separate than now. Banking, as we have seen, was about discounting of short term bills that financed goods on their way toward sale. Bank liquidity was assured by creating a portfolio of bills maturing at different times in the future, so that you always have cash inflow to meet possible cash outflow, either from deposit withdrawal or from new discounts. On the other hand, there was a capital market, for government and corporate bonds, and also equities.

In stylized form, Bagehot's world looked like this:

Bagehot's World

Primary borrower Business		Bank		Primary lender Household	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	Bills	Bills	Deposits	Deposits	
	Bonds Equities			Bonds Equities	

We can see in the first line the money market, and in the second line the capital market. Households of course buy bonds and equities by transferring their bank deposits to business, which spends them on investments. But for the most part banks did not buy bonds or equities; they were thought to be inappropriate assets for banks because they were not "self-liquidating".

Economic historians argue whether British excellence in money markets held back the capital development of the nation by emphasizing short term finance at the expense of long term finance. I don't have a developed view on the matter. But I do observe that in the U.S., which was a developing country compared to Britain, the pressing need for long term finance caused the banking system regularly to hold long term bonds, and even some of the purported short term loans were really long term, since they were intended simply to roll over at maturity. As a consequence, banks did not have regular cash inflows to provide liquidity and they had to devise

another means. (They also did not have a central bank to provide rediscount, and so they had to devise other means for that as well.)

Instead of self-liquidating short term assets, they held substantial cash reserves, including correspondent balances, and devised elaborate mechanisms of secured interbank borrowing using their bond holdings as collateral.

The New World

Deficit bank		Surplus Bank	
Assets	Liabilities	Assets	Liabilities
Bonds Loans Cash Banker's balances -Cash Reserves -Banker's balances	Deposits +Interbank borrowing (repo)	Bonds Loans Cash +Cash Reserves +Interbank Borrowing (repo)	Deposits Banker's balances -Banker's balances

By the way, this explains why the US was the first to have bond rating services. A bank in need of funds needed to sell some of its bond holding to another bank, or put it up as collateral (early repo), and no one had time or interest in investigating the fundamental value of the bond. So they relied on rating agencies to tell them, and the rating agencies focused on the question of whether there was likely to be a default in the next year. Default means that the coupon would not be paid, although possibly it was only delayed. You can see how the importance of the timing of cash flows enters the American system, but in a different way.

In the American system, liquidity was all about "shiftability", meaning salability. The word was first used by Harold Moulton (JPE 1918), who was worrying that the new Fed discounting policy might interfere with capital accumulation. Shiftability means market liquidity, the ability to buy and sell. The net effect of this system was to mobilize bank deposits as a source of funding for long term capital finance, by using the repo market, not the real bills discount mechanism, as a source of liquidity. However the founders of the Fed saw this shiftability system as a source of instability and tried to replace it with proper real bills discount banking.

One way of understanding the banking collapse of the 1930s is that it was a run on the shadow banking system of the time, namely this system of private funding liquidity using repo on non-Treasury security. The Fed refused to support it, so it collapsed. The consequence was that the

government had to take on responsibility for long term lending itself through various mechanisms, most famously the Reconstruction Finance Corporation. This mistake was not, thankfully, repeated in 2007-8!

The larger lesson to draw is that, in the United States, securities markets and money markets have always been intertwined, going back even before the Fed. In previous lectures, I have told a story about monetary transmission, using central bank control over funding liquidity to affect market liquidity in securities markets, and hence asset prices. I think that is not recent, but in fact ancient (at least in the United States), although understanding how it works is not so ancient.

The Alchemy of Banking and Development Finance

One person who did understand the importance of banking for development finance was Schumpeter, whose experience was with the Continental banking system, not the British. As early as his PhD thesis he insisted on the importance of bank deposit creation as a mechanism for development finance, a way of giving purchasing power to entrepreneurs who do not have it, without requiring any individual saver to cough up saved funds.

Development Bank

Assets	Liabilities
+capital loan	+deposit

We've got to be careful here. It is very easy to go astray, and most of those you will find posing as monetary alchemists are in fact monetary cranks. The ability to create money from nothing is not the same as the ability to create bread from nothing. So important has it sometimes seemed to banish the unsound reasoning of monetary cranks, that economics has sometimes come close to adopting as a kind of Creed, "there aint no such thing as a free lunch". But there is.

Adam Smith, back in 1776, urged the adoption of a paper money system as a way of economizing on the use of gold. We could trade the gold for real capital assets, which is a trade of a sterile commodity for a productive resource. In a way that is what we have been doing ever since, moving toward a pure credit economy. I mention Adam Smith to establish the bona fides of the search for a monetary free lunch, but in fact the main source lies elsewhere than substituting inside money for outside money.

The key is the use of banking to mobilize unused resources. If entrepreneurs use their new deposits to buy things that otherwise would have been unsold, they don't raise prices, they increase economic activity. Schumpeter emphasized technological change. Others have

emphasized other margins of mobilization—see Lewis model on economic development and the role of banks in mobilizing underemployed labor, moving it from the traditional subsistence sector to modern manufacturing sector.

Payment versus Funding

The problem comes clearest if we think of the primary borrower as a corporation borrowing in order to purchase some physical capital asset such as a machine or a building. This involves the corporation in current expenditures far in excess of current receipts, and it is only over the long lifetime of the capital asset that the corporation expects to reverse that imbalance. Society's problem thus is to find someone to hold that illiquid asset over *its entire productive lifetime*.

More concretely, the issue is the distinction between **payment versus funding.** (Recall Lec 6, p. 39). When we were adopting a banking point of view, we always emphasized the way that expansion of balance sheets by swapping IOUs creates elasticity in the payments system. Just so, a corporation can acquire means of payment simply by swapping its own IOU with that of a bank. And that, more or less, was where we stopped (except for wrinkles about whether the quantity of IOUs is limited by the quantity of reserves, or only the price). Now we go the next step, and think of the corporation as using that means of payment to acquire a physical asset from society. Then the question arises whether society as a whole is happy with the asset portfolio implied by the swap. Is society happy holding money rather than the physical asset?

Initial Payment Point of View (Money Mkt)

Corp as Borrower		Bank		Society as Lender	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
+100 deposit -100 deposit +machine	+100 bank loan	+100 loan	+100 deposit	+100 deposit -machine	

If the answer is yes, then we are done for the moment and the new capital is funded by an increase in outstanding money balances. But this is a tricky matter because money is a promise to pay on demand, while the capital asset cannot be turned into payment except over a long time. So there is significant mismatch between commitments and cash flows, which (as we have seen) can be a cause for crisis. So a safer form of finance will be when the financing matches better the characteristics of the asset being financed. This is the funding problem.

If society is not happy acquiring additional money balances, then the recipients will attempt to use those balances to buy a different financial asset. (Note that we are ruling out by assumption any attempt to spend the balances on goods. We assume that savings equals investment. The only imbalance at issue is a portfolio imbalance.) This will drive up the price of alternatives to money and so provide an incentive to issue such alternatives.

One way this all can work out is the following. Suppose that the bank loan was only bridge financing until the capital asset is up and running. Once its ability to produce revenue is proven, the corporation can pursue permanent financing in the form of bonds. The proceeds of that bond offering are then used to pay off the bank loan, and in effect society swaps the deposits it doesn't want to hold for a bond that it does want to hold.

"Permanent" Funding Point of View (Capital Mkt)

Corp as Borrower		Bank		Society as Lender	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	-100 loan +100 bond	-100 loan	-100 deposit	-100 deposit +100 bond	

Note that the bridge financing by deposit expansion is a kind of indirect finance, while the takeout financing by bond issue is a kind of direct finance. The Gurley and Shaw point of view emphasized the use of intermediaries as sources of indirect finance for the capital development of the nation.

What is an intermediary?

The word itself gives us a clue. It is a financial institution that mediates between the primary borrower and primary lender of funds, holding as assets the liabilities that the borrower issues, and issuing as its own liabilities the assets that the creditor holds. Standard banking texts (such as Mishkin) make a big point about the <u>empirical</u> importance of such **indirect finance** by contrast with **direct finance** in which borrower and lender meet directly, as by the direct issue and holding of stocks and bonds.

Primary Borrower Corporation		Intermediary		Primary Lender Household	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	Direct Finance --stocks --bonds			**Direct** Finance --stocks --bonds	
	Indirect Finance --stocks --bonds --loans	**Indirect** --stocks --bonds --loans	**Indirect** --pension --insurance --money	**Indirect** Finance --pension --insurance --money	

Gurley and Shaw, in their classic <u>Money in a Theory of Finance</u> (1960), made a big point about the <u>theoretical</u> importance of such financial intermediation, along the following lines. They

argued that intermediaries help to facilitate economic growth by bridging any potential mismatch between the kind of liabilities that borrowers want to issue and the kind of assets that creditors want to hold. It is by changes in the quantity of indirect finance that capital markets equilibrate, and channel funds from savers to investors.

They overdid it. Modern finance, by contrast, emphasizes that no risk is eliminated in the process of intermediation, only transferred, and sometimes quite opaquely. Thus it is by changes in price that capital markets equilibrate. Attempts to avoid this by indirect finance only create arbitrage opportunities that over time have transformed the system.

Paradigmatic Intermediaries: Insurance and Pension

In idealized form, we can think of these intermediaries as holding the following balance sheets:

Insurance	
bonds	policies

Pension	
stocks	pension plan

For the more complete balance sheet, see www.federalreserve.gov/releases/z1/current/z1r-4.pdf Tables L.116 and L.117 show insurance companies. Tables L.118 and L.119 show Pension Funds.

There can be no question that these are intermediaries, since their assets are clearly of different kind than their liabilities. Bonds pay regular coupons no matter what happens. Insurance policies pay only if some insurable event occurs. Stocks pay dividends and capital gains. Pension plans pay when owners retire, an amount depending on final wage, inflation, job tenure and the like (for defined benefit plans).

One of the trends in insurance is replacement of the traditional whole life insurance policy, which included a savings component, with term insurance. One of the trends in pensions is replacement of the traditional defined benefit plan with the defined contribution plan, such as a 401(k). Both trends reflect the rise of mutual funds as competitors to traditional indirect finance.

Here is a bond mutual fund, and a stock mutual fund.

Bond Mutual Fund	
bonds	shares

Stock mutual fund	
stocks	shares

(For details see tables L.122 and L.123)

At first glance, it appears that in a mutual fund there is intermediation, but that is mainly just pooling of risk through diversification, and not much transformation of risk. By construction, the

shares have exactly the same risk properties as the underlying pool of bonds or stocks. There is some benefit for the mutual fund shareholder from diversification, also management services. There is also some liquidity benefit perhaps, because open end funds typically promise to buy back shares at NAV. But that just means that mutual funds have to keep cash or lines of credit for the purpose, both of which will lower the return and hence are paid for by shareholders. Thus final glance confirms initial glance.

Basically bond funds have replaced the saving component of whole life insurance policies, and also bank time deposits, as fixed income saving instruments. And stock funds have replaced defined benefit plans as retirement income savings instruments. This change reduces the degree of transformation in financial intermediation. Nowadays, mismatch between the preferences of borrowers and the preferences of lenders is increasingly resolved by price changes rather than by traditional intermediation.

Banks as Intermediaries

So far in this course we have been emphasizing the special role of banks in the liquidity hierarchy on account of the fact that their liabilities (bank deposits) are means of payment. Thus banks are able to turn private debts into purchasing power by accepting them, in effect swapping IOUs.

As it happens, individual households and firms in the economy not only want to make payments (flow), they also want to hold means of payment (stock). This makes room for the banking system as a whole to issue a **permanent short position in cash**. The role of banks as intermediaries comes from their use of this short position to fund long positions in non-cash assets. (The huge literature on money demand is built on this institutional fact.)

You might think that the liquidity of bank liabilities requires liquidity of their assets, but this is true only at the margin. To achieve liquidity on the margin, it is sufficient that the banking system hold some cash reserves, and have the ability to replenish those reserves. For this latter purpose, individual banks depend on access to the FF market and discount loans at the Fed, and on holding of some easily sold "secondary reserve" assets. (Here we see the distinction between liquid asset and liquid market again.)

The important point is that, after liquidity needs are taken care of, there will be some fraction of total deposits left over which can be invested in less liquid assets.

Assets	Liabilities
Cash reserves	Deposits
Secondary reserves	
Loans	Net Worth

Because bank deposits are substitutes for currency, and because currency typically pays no interest, deposits typically pay zero or low interest as well. This makes it profitable for banks to

make loans at relatively attractive rates. Not surprising, there is considerable competition to get access to this cheap money. Historically, the biggest player in this competition is the government. In times of war, we always see bank balance sheets fill up with loans to the government or to government-favored enterprises. Similarly, for countries trying to jumpstart a development process, it is very tempting to begin by trying to mobilize idle balances on the balance sheets of domestic banks.

In the United States, politics has resulted in a division of the spoils of this cheap source of funds during peacetime. Until recently, the central bank has invested almost entirely in government debt. Commercial banks historically specialize in making commercial and industrial loans, though they do other things as well. Savings and loans historically specialize in mortgage lending and they issue "shares" which have come to look more and more like deposits. You can find data at http://www.federalreserve.gov/releases/z1/Current/z1r-4.pdf , L.108 Monetary Authority, L.109 Commercial Banks, L.114 Savings Institutions, L.121 Money Market Mutual Funds.

Central Bank		Commercial Bank		Savings and Loan	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
Treasury secs	High-powered money	C&I loans	deposits	Mortgage loans	Shares

Building on the discussion from last time, it should be clear that even if we grant that a certain portion of deposits are permanent, there are still considerable risks involved in using those deposits to fund C&I loans, and mortgages. Liquidity risk is one of them. But once we turn our attention to intermediation, **solvency risk** becomes the main focus. Main sources of solvency risk are interest rate risk and credit risk. Again the financial revolution has transformed banking, most dramatically in the transformation from Jimmy Stewart banks to shadow banking.

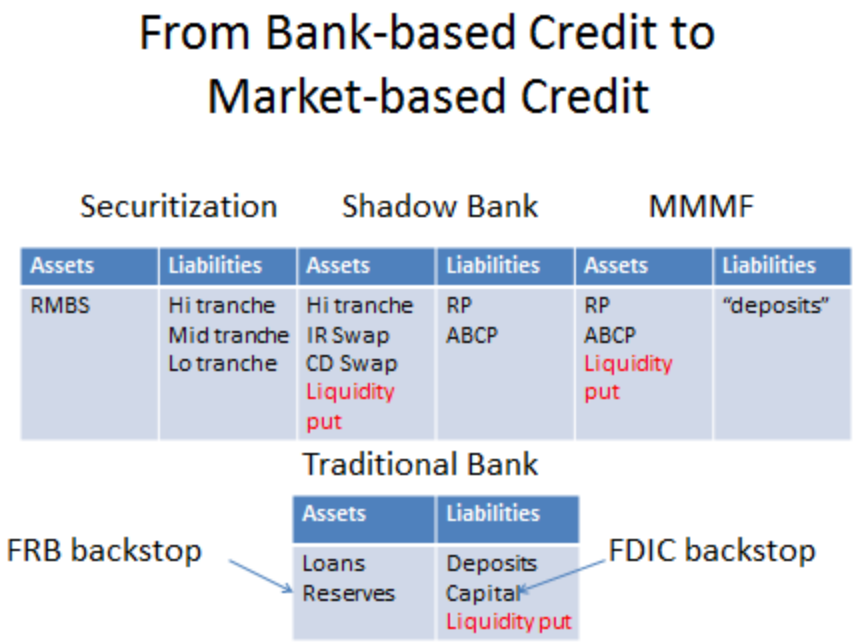
The Shadow Banking System as Intermediation

This development needs to be understood as part of the challenge of finance. From a finance point of view, an intermediary that offers liabilities with different risk characteristics than its assets must itself be bearing the risk of that transformation. But this means that its stockholders (or the government), presumably primary lenders, must be bearing that risk, so the transformation of risk is illusory. Risk is just getting moved around, not eliminated.

From the finance point of view, the illusion that risk is being eliminated comes from the fact that risk is not being (directly) priced. The answer is to strip out every risk exposure and price it separately. From the finance point of view, any mismatch between the kind of liabilities that borrowers want to issue and the kind of assets that creditors want to hold is equilibrated not by

changes in quantity but by changes in price. Creditors bid up the price of assets they like, and bid down the price of assets they don't like, until the price exactly compensates for the risk.

That is one way of understanding what shadow banking does. Thus, a modern course in banking has to understand the various swaps that are used to strip out the risk exposure and price it separately. That's what we will be doing in the next lectures.



18. Forwards and Futures

This is the first of a series of three lectures intended to bring the money view into contact with the finance view of the world. We are going to talk first about interest rate forwards and futures, then interest rate swaps, then credit default swaps. I have been treating forwards as a kind of swap from the beginning, (as all banking is swap of IOUs), but even so it will take a little bit of doing to connect this up with the finance view of swaps. Here is an idea to have in mind to keep you oriented, the idea of "Banking as Advance Clearing".

Remember how we began the course, in lectures 5-9 on "Banking as a Clearing System". In those lectures, we saw how, at any moment, a particular intertemporal pattern of cash flows and cash commitments resolves itself into a particular pattern of clearing and settlement at a moment in time. Deficit agents buy cash (borrow) today to delay settlement, and the elastic availability of loans is the essential source of elasticity in the payments system. By means of **credit**, current imbalances in the pattern of cash inflows and outflows are pushed into the future where, hopefully, they can be offset against a pattern of imbalances going the other way. The cost of pushing those imbalances into the future is the current money rate of interest, which operates therefore as a symptom of the degree of imbalance but also as an incentive to pay up soon; here is the discipline in the system. Financial crisis arises when delaying tactics no longer work; it is an extreme form of discipline.

What I mean now about advance clearing is the way that emerging imbalances in the future show up as cash flow imbalances in the present, again with the money rate of interest serving as a symptom, and discipline. In finance, the future determines the present, but no one knows the future, so there can be multiple views of what the future will look like. **How does it happen that one path gets chosen over other possibilities; how does it happen meanwhile that diverse views get coordinated?** One way is by market pricing of different views, and by the effect of that pricing on behavior that operates through the survival constraint.

At one extreme, if the market changes its mind about your view of the future, you may have difficulty rolling your funding. The current survival constraint is thus a key mechanism through which one future path gets chosen over all the others. But there are subtler paths at work as well, through which ideas about the future cause changes in cash flows today, which make the survival constraint looser for some people and tighter for others. Today, we explore one of them, namely the cash flow consequences of changes in futures prices.

Forwards and Futures

Suppose a firm has ordered a machine for delivery three months from now. The firm plans to pay for the machine by borrowing, but is concerned that interest rates three months from now might be higher. The firm can lock in a borrowing rate by engaging in a forward contract with a bank. We can think of that forward contract as a swap of IOUs, as follows:

Firm A		Bank	
Assets	Liabilities	Assets	Liabilities
3 month deposit	6 month loan	6 month loan	3 month deposit

We have seen in a previous lecture that the swap of IOUs will have zero present value if the deposit and the loan both pay the forward rate of interest F[3,6] defined by forward interest parity

$(1+R[0,3])(1+F[3,6]) = (1+R[0,6])$

We have also seen that, because of the failure of the expectations hypothesis, in general

$F[3,6] > E_0R[3,6]$

This empirical regularity provides incentive for the bank to enter into the forward contract. The forward loan is more profitable on average than the spot loan.

This swap of IOUs solves the problem of the firm, but creates a problem for the bank, since in three months the bank will have to come up with the money to lend to the firm, and at that moment it is entirely possible that R[3,6] will be greater than F[3,6], so leaving the bank with a loss. Ideally the bank (or the banking system as a whole) has another client with exactly offsetting needs, i.e. a firm that wants to lock in a lending rate by engaging in a forward contract as follows.

Firm B		Bank	
Assets	Liabilities	Assets	Liabilities
6 month deposit	3 month loan	3 month loan	6 month deposit

You can see how these two contracts exactly offset each other on the balance sheet of the bank. The bank is borrowing and lending for 3 months, and borrowing and lending for 6 months. So the bank has no net exposure. The important point to notice is that this combination of forward contracts have essentially cleared today a future payment from Firm B (the ultimate lender) to Firm A (the ultimate borrower). That's the sense in which the forward market can be considered to be about advance clearing. There will be cash flows in three months between surplus and deficit firms, but they are all pre-arranged today.

Now in general there is no reason to expect that forward contracts all net out in this way on the balance sheet of any single bank. Even when banks trade their forward exposure with one another (using FRAs) there is no reason to expect that forward contracts all net out in the banking system as a whole. That means that the banking system will be left with a net exposure to the risk that the future spot rate of interest will be higher than the current forward rate. Banks will not hold this risk unless they are compensated by an expectation of profit. The source of this profit is movement in the forward rate away from expected spot.

As we have seen repeatedly in this course, this is one possible explanation for the empirical failure of the expectations theory of the term structure. The difference between the current forward rate and the current expectation of the future spot rate is just the expected profit from an unhedged forward exposure. The point is that the imbalance between future cash flows and cash

commitments shows up as distortion of the current forward interest rate away from the expected future spot interest rate.

There is more to it than this. If the forward imbalance is large, then the current price distortion will be large, and that means that the expected profit from an unhedged forward exposure will be large. This expected profit can be expected to attract speculators in the larger economy, outside the banking system, to hold the exposure that the banks cannot or are unwilling to bear. **Conceptually we will think of the futures market as the place where the banking system sells off its excess forward exposure to speculators in the outside economy.** Futures are forwards that are marked to market daily, with any changes in value settled daily. Distortions that affect forward rates will also affect futures rates, and hence current cash flows.

Chain of Hedges

Client	Bank	Banking System	Futures Exch.	
F[3,6]	Forward contract	FRA	Futures	ER[3,6]

We know from the failure of EH that forward interest rates tend to be upward biased forecasts of future spot rates. The bias is often thought of as a kind of liquidity premium, but people have had a hard time explaining just what risk that premium is compensating for. More generally we observe the following pattern

Forward rate > futures rate > expected spot

This pattern gives everyone a profit incentive to enter into the trade. To understand these effects, we now backtrack and build up a somewhat more formal account of forwards and futures, now using the more typical language of finance.

Forwards

Start with forwards. Forward contracts are promises to deliver goods at future time T at a given price K. The classic example is that of the wheat farmer who has a natural long position in wheat and the baker who has a natural short position. Both face price risk. If they can arrange a forward contract, however, they can lock in the future price of wheat, and both can be made better off. We say they each **hedge** their natural forward exposure by taking an opposite position in a formal forward contract.

When the time comes to settle, the spot price of wheat is likely to be different from the contracted delivery price. In this sense one side "wins" ex post. In fact, we can track these "winnings" over the life of the contract as the value of the contract changes. But in forward contracts these winnings are only notional. No matter what happens to the spot price of wheat, at delivery date the short delivers the contracted good to the long, and the long delivers the contracted price to the short in money.

For our purpose we want to think about the case where the underlying is not a physical commodity like wheat but a financial instrument like a Treasury bond. (Or a bank time deposit, such as a Eurodollar deposit.) It's easiest to think about the case where the underlying is a zero

coupon riskless bond that yields no cash income and has no carrying cost. We can rewrite our Forward Interest Parity condition (from p. 129) in price terms as follows:

$[1/1+F(3,6)] = [1/1+R(0,6)][1+R(0,3)]$

where the first term is the forward delivery price, the second is the current spot price, and the third the interest rate between now and the forward date.

Now think about how the forward price changes over time. The equation above is the forward rate at time zero. At time 1, 2, 3 we have the following

$[1/1+F_1(3,6)] = [1/1+R(1,6)][1+R(1,3)]$

$[1/1+F_2(3,6)] = [1/1+R(2,6)][1+R(2,3)]$

$[1/1+F_3(3,6)] = [1/1+R(3,6)]$

There is no reason at all to expect that these forward rates are the same as the period zero forward rate. That means that the forward contract established at time zero will change in value throughout time. To help us think about that change, and to connect this discussion up with standard finance treatments, it will be useful to recast the discussion in continuous time by introducing some new notation:

At time 0 when the contract is written, we have the following formula relating the forward delivery price K to the current spot price S_0:

[1] $K = S_0 e^{rT}$. (See equation 3.5 in Hull 5th ed.)

Don't let this equation scare you. It is nothing more than a version of our familiar forward interest parity condition. Think of the forward price K as 1/(1+F[3,6]), the spot price S_0 as 1/(1+R[0,6]) and the interest rate term e^{rT} as (1+R[0,3]).

To see how this is an arbitrage condition, think about how you would make money if the condition does not hold:

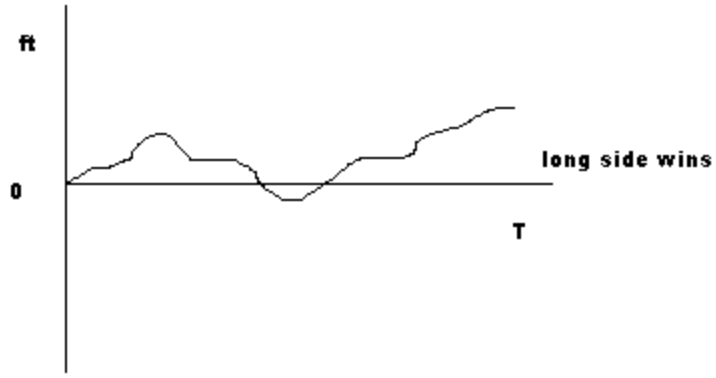
$K > S_0 e^{rT}$, buy the bond spot and sell it forward. In this way you lock in a rate of return greater than the rate of interest r at which you can borrow money to finance the trade.

$K < S_0 e^{rT}$, sell the bond spot and buy it forward. In this way you lock in a borrowing rate lower than the rate of interest r at which you can lend the money you receive from selling the bond.

The important point to emphasize is that, for forward contracts, the delivery price K is fixed for the life of the contract. Hence, over the life of the contract, the <u>value of the forward</u> contract will change,

[2] $f_t = S_t - Ke^{-r(T-t)}$, for $0<t<T$. (See equation 3.9 in Hull)

Note that there is a time subscript on both f and S, but K is fixed. At t=0, we have $S_0 = Ke^{-rT}$, so $f_0=0$ when the forward contract is signed. At t=T, we have $f_T = S_T - K$. This is the notional winning we talked about above. In between time 0 and time T, the value of the forward contract fluctuates, depending mainly on the fluctuating spot price of the underlying zero coupon bond.



In most forward contracts, at the final date the long side pays the short side the agreed price K and receives the agreed underlying, which is worth S_T. In interest rate forward contracts however, "cash settlement" is the rule. Instead of delivering the bond for K, the short side delivers the current spot price of the bond in return for the payment K. This means net cash payment of the final value $f_T=S_T-K$ from short to long if positive and from long to short if negative. In cash settlement, the notional winnings become real cash flows at time T.

Futures

A futures contract is like a forward except that all changes in the value of the contract f_t are instead absorbed in changes in the delivery price, which is therefore called the futures price, F_t. F_t is reset every day so that f_t is zero. In other words, the **futures price is that price at which the analogous forward contract has a current value of zero**.

$$0 = S_t - F_t e^{-r(T-t)}$$

[3] $$F_t = S_t e^{r(T-t)}$$

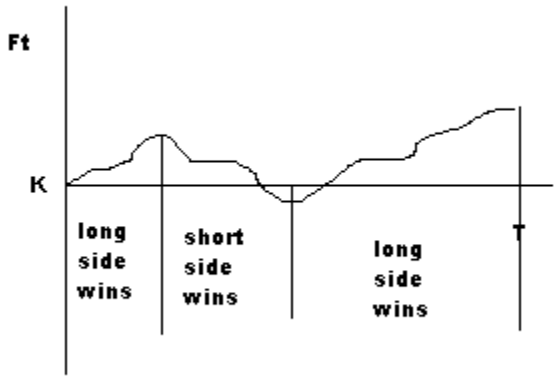
This equation is sometimes called **future-spot parity**.[1] Notice that at origination the futures price is equal to the contract price in a forward contract. (Compare equation [3] and equation [1]). And at expiry the futures price is equal to the spot price. In algebra this means that at t=0, $F_t=K$, and at t=T, $F_T=S_T$.

The big difference between forwards and futures is that for futures the daily "winnings" that come from changing spot prices are not at all notional. In fact they are actually paid out daily over the life of the contract. In a forward contract, the only payment flow is at the end, and the amount of that payment is fixed by the contract from the very beginning. In a futures contract, payments

[1] I am abstracting from interest paid or received on the fluctuating balances as the futures price changes. That is one reason my formula may be a bit simpler than the formula you may have learned in a previous finance course.

are being made all along the life of the contract, whenever the futures price changes. This is called "mark to market".

Concretely, these payments involve additions and subtractions from "margin accounts" held at the futures clearinghouse. It is significant that both the long and short side have to put up margin, because at the moment the contract is entered, both are in a sense equally likely to lose and so equally likely to have to make a payment to the other side.[3] You can think of these margin accounts as similar to bank deposits, but in fact the clearinghouse will accept securities for the purpose. They have to be liquid securities however, and at the end of the day the securities are repriced to reflect any change during the day. Thus the collateral underlying the futures contract as well as the futures contract itself are both marked to market every day.



The cumulative payment on the futures is the same as the final payment on the forward, but for the futures the cash flows come about every day during the life of the contract.[4] This is a very concrete way in which views about the future are settled today.

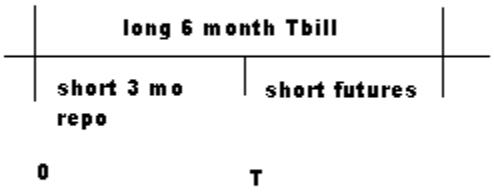
Monetary Issues

Stigum (718-722) talks about a trade involving spot 6 month bills, the 3 month ahead futures contract on the bills, and the 3 month riskfree repo rate. It starts with buying the 6 month bill for spot price S_0 using money borrowed at the repo rate r. Borrowing short term in order to lend long term however exposes you to price risk, since you don't know the rate at which you will be able to finance the second three months of the bill. To hedge that price risk, you short 3 month Treasury bill futures at the futures price F_t. Then, whatever happens to the Tbill price is exactly countered by whatever happens to the futures price. In this way you hedge all price risk. The question then is, given that you have hedged all price risk, why would you ever expect to profit from this trade, and hence why would anyone do it?

Here is a diagram showing the trade in question:

[3] Not exactly equally likely, given failure of EH, but you get the idea.

[4] Not exactly the same, given interest paid on the daily winnings in the futures contract, but here we are abstracting from that.



One way to understand this trade is that the trader is long a synthetic forward contract (the combo of the Tbill and repo) and short the corresponding futures contract. This way of putting the matter makes it even more puzzling why the trade would ever make a profit. To see why such a trade might be profitable, let's return to the relationship between futures and current spot prices that we talked about previously, under the name futures-spot parity. The same relationship Stigum calls "full carry pricing". Deviations from full carry pricing offer opportunities for arbitrage profit, as below:

$F_t = S_t e^{r(T-t)}$ full carry pricing, no arbitrage profit

$F_t > S_t e^{r(T-t)}$ cash and carry arbitrage:
1) Short futures, long underlying at S_0, finance by borrowing at r
2) At futures expiry, deliver underlying for spot S_T, repay loan

$F_t < S_t e^{r(T-t)}$ reverse cash and carry arbitrage:
1) Long futures, short underlying at S_0, invest proceeds at r
2) At futures expiry, pay spot S_T for underlying

Stigum uses a somewhat different language when she talks about the difference between an "implied repo rate" and the actual repo rate, but she is talking about exactly the same thing. We understand what the actual repo rate is—it is the rate paid on the short term repo (the "carry" part of the trade) that is used to purchase the Treasury bill (the "cash" part of the trade). The implied repo rate is the short term return that is locked in by the combination of the cash bill and the short futures position. Observe that we know the price at which we can buy the bill, and we know the price at which we can sell the bill (i.e. the futures price).

Define the implied repo rate ρ by the equation

$$F_t = S_t e^{\rho(T-t)}$$

The implied repo rate ρ is thus the borrowing rate that would have to hold in order for futures prices to satisfy full carry pricing. Now we can express the arbitrage opportunities as a deviation of the implied repo rate from the actual repo rate:

$\rho = r$ full carry pricing, no arbitrage profit

$\rho > r$ cash and carry arbitrage

$\rho < r$ reverse cash and carry arbitrage

The arbitrage profit in the cash and carry trade arises from the fact that you can borrow at a lower rate than you can lend. Put that way, it is astonishing that such a relationship would ever hold for more than an instant. Why doesn't everyone do it, and in volume sufficient to eliminate the arb?

The cash and carry arbitrage is long forward and short futures. What is the risk in that position that might command a premium for bearing it? If the forward rate is typically greater than the expected spot, that means we can expect to gain by borrowing short and lending long. Our long forward interest rate position should be increasing in value. But at the same time our short futures interest rate position should be decreasing in value. These two positions more or less net out in terms of value, **but not in terms of cash flow**. Futures are marked to market whereas forwards are not. **This means that the cash and carry trade typically involves negative cash flows throughout the life of the contract, plus a large positive cash flow at maturity.** The profit comes from the fact that the positive cash flow is larger than all the negative flows added up, but the fact remains that the timing is inconvenient.

In order to get the positive cash flow, we have to hold the position for three months, and that means surviving a series of (expected) negative cash flows. Not only that, but these negative cash flows might all come at once. Thus the volatility of the spot price of the underlying bill creates liquidity risk for the cash and carry trade.

The fact that markets typically violate the expectations hypothesis is well-known, even if not well-understood. The profitability of the cash and carry trade is not well-known, but perhaps we can understand it as a reward for bearing liquidity risk. What I suggest is that these two anomalies have the same origin, namely a mismatch in the natural forward interest rate positions emerging from the real economy.[2]

[2] A former TA in this course, Daniel Neilson, has done important work showing how deviations between forward and futures rates serve as an observable proxy for unobservable deviations between forward rates and expected spot rates. This is a kind of test of the liquidity premium theory of the term structure.
http://www.peri.umass.edu/fileadmin/pdf/conference_papers/d_arista/liquidity_darista.PDF

19. Interest Rate Swaps

Reading: Stigum 19 on Swaps. See also Hull who builds from the idea (mentioned in Stigum) that swaps are like a portfolio of forward contracts. Daily Financial Times includes bid-ask of swaps in 5 currencies, all maturities from 1 year to 30, "swap curve" vs. yield curve.

Lots of kinds of swaps. I'm going to focus on interest rate swaps, both medium term and short term. Basis swaps, currency swaps, are easy to understand by analogy. Ignore swaptions because we are ignoring options.

A Swap is a swap of IOUs, or a parallel loan

A good place to start is p. 880, *Street Speak in Swap Land*, which is about interest rate swaps, fixed for floating. If we think of the notional principal as actual, we can understand what is going on better. Suppose two counterparties AA and BBB, one is able to borrow fixed rate but wants to borrow floating rate, the other is able to borrow floating rate but wants to borrow fixed. They can help one another by swapping IOUs, fixed for floating, in a parallel loan structure as follows.

	AA	BBB
	Seller of swap	Buyer of swap
	Short swap	Long swap
	Payor of Floating	Payor of Fixed

	Assets	Liabilities	Assets	Liabilities
Original position		Fixed rate borrowing		Floating rate borrowing
Parallel Loan	Fixed rate	Floating rate	Floating Rate	Fixed Rate
Swap		Swap	Swap	

In the parallel loan arrangement AA still pays its original creditor a fixed rate over the life of its original loan, and BBB still pays its original creditor a floating rate over the life of its original loan. The swap of IOUs between AA and BBB means that AA receives a fixed rate payment from BBB, while BBB receives a floating rate payment from AA; these receipts match the promised payments on their original loans. In this way, AA achieves floating rate financing while BBB achieves fixed rate financing.

The parallel loan structure solves the problem for the two companies, but it does so by expanding both balance sheets, and so also the apparent leverage and counterparty risk exposure on both balance sheets. Since money is going both ways, it is natural to net the two payments and pay only the net, from AA to BBB or from BBB to AA, whichever is larger. (In most cases, BBB will be paying AA because the short term interest rate is lower than the long term interest rate, due to failure of EH.) This netting goes

some way toward reducing counterparty risk. The swap contract goes even farther by netting the principal payments as well, both at the beginning and at the end of the contract. The net payment flows on a swap contract are the same as in the parallel loan structure, but now everything is off-balance sheet.

In the swap arrangement, note well the market lingo convention. A long swap position pays fixed and receives flex, and a short swap position is just the opposite. (One way to remember this is to observe that the long swap position increases in value when the floating rate of interest rises; another is to think of the swap as a kind of insurance contract that hedges floating rate risk, so a liability for insurer and an asset for the insured.) In deference to the market lingo, I'll treat long swap positions as assets and short swap positions as liabilities when we put them on the balance sheet. This should cause no confusion as long as we remember the parallel loan interpretation.

Swaps and Other parallel loans

To help link the swap idea to other things we already know, it might be helpful to think of "selling a swap," or taking a short swap position, as like buying a five-year fixed rate bond and financing the position by borrowing short term, using the bond as collateral for the loan. From this point of view, we see the swap as similar to a repo. The difference is that with swaps we are dealing with corporate liabilities not governments, and also with much longer maturities both of the bond (say five years rather than 3 month Tbill) and of the financing (say 6 month LIBOR rather than overnight repo).

Further, it is common in the finance literature to analyze the swap as a strip of forward interest rate contracts, one for every time-dated payment on the underlying notional fixed rate loan. (Note: forward not futures contracts, because payments are periodic, not marked to market.) BBB is in effect locking in the a future borrowing rate, and AA is in effect locking in a future lending rate. WARNING: being long a swap is like being short a portfolio of forwards, so you hedge a long swap with long futures, which is rather counterintuitive (to say the least!)

As always, keeping in mind the parallel loan interpretation will allay any confusion. When we were talking about interest rate forwards, we saw how a forward can be understood as a parallel loan as follows:

Long forward		Short forward	
Assets	Liabilities	Assets	Liabilities
6 month loan R(0,6)	3 month deposit R(0,3)	3 month deposit R(0,3)	6 month loan R(0,6)

If we compare this to the parallel loan interpretation of the interest rate swap, we see that exposure on the first payment of the short swap is just like the long forward. The later swap payments are analogously like more distant forwards.

Why swap? Comparative advantage

In Stigum's example (p. 874), BBB borrows floating and AA borrows fixed, then they swap. The reason they do this is that by assumption BBB can borrow relatively more cheaply in

floating, and AA can borrow relatively more cheaply in fixed (though absolutely more cheaply in all markets). Each however would like to be borrowing in the market that is relatively more expensive. Here is the structure of rates Stigum assumes:

	Term Loan floating	5 year fixed Eurobond
BBB	6 month LIBOR + 1/4	5.85
AA	6 month LIBOR + 1/8	5.375
Difference	12.5 bp	47.5 bp

There are 35 bp to play with (47.5-12.5). In Stigum's example (Table 19-1, p. 875) AA gets 25 of them and so achieves LIBOR – 1/8 (original floating – 25 bp), and BBB gets 10 and so achieves 5.75 Eurobond (original fixed – 10 bp). We can understand this better by putting her numbers into our parallel loan interpretation, as follows:

AA		BBB	
Assets	Liabilities	Assets	Liabilities
[5 year fixed, 5.50	5 yr bond, 5.375 LIBOR]	[LIBOR	LIBOR + ¼ 5 year fixed, 5.50]

Why does this apparent free lunch exist? One reason is market imperfection. Stigum tells the story about British capital controls that were evaded by parallel loans which were in effect currency swaps. Something like this might be happening, if Triple B is for some reason locked out of the Eurobond market. Stigum notes also that US interest rate swaps have their origin in 1981, in the midst of the Volcker tight money period, when some lesser credits would have been locked out of certain markets completely.

But another possible reason for this structure of rates is counterparty risk, and that suggests that the lunch might not be so free. A bank may be willing to lend short term to Triple B because it thinks it can reassess the situation every 6 months, perhaps raising the markup over LIBOR if BBB gets into trouble. The higher markup for longer term lending compensates for the fact that there is a lot more that can go wrong in five years than in six months. The swap gives Triple B long term financing, but leaves AA holding the credit risk. If things go bad for Triple B, it will roll over its term loan at higher and higher markups over LIBOR, but it is receiving only LIBOR flat. It will have to pay its bank and also Fuji. Possibility of default to both the bank and to AA.

Market making in Swaps

Whatever the reason for this apparent free lunch, the important point is that the 35 bp attract not only AA and BBB, but also brokers and dealers who take a few of the bp to set up and manage the swap. Thus AA could borrow in the Eurobond market and swap fixed for floating by selling a swap to a dealer. Triple B likewise could borrow in the floating market and swap floating for fixed by buying a swap. The dealer would take 1bp each way, subject to credit check.

From the dealer's perspective, the net result is a growing book of swap positions, hopefully matched book. Dealers make markets in swaps by taking positions and hedging them, either by taking opposite positions in the same or similar swaps or by taking opposite positions in related instruments (like forwards and futures).

AA		Investment Bank		BBB	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	AA swap	AA swap	BBB swap	BBB swap	

In general dealers will not be able to match longs and shorts perfectly with the business that naturally comes to them, so they must resort to imperfect hedging. The kind of position that works as a hedge gives us further insight into the nature of the swap contract.

Stigum likes to say that a swap is a synthetic corporate bond (p. 906). What she means is that a short position in interest rate swaps is similar to a long bond position financed by repo. Looked at in this way, it is clear that such a position can be hedged with a short bond position. Treasuries or Euros are the most liquid instruments, so one might look for a hedge there, either in the cash market or the futures markets. It is important however to recognize that such a hedge involves continued exposure to basis risk because the swap rate is a corporate not a government rate. Also, the dealer faces counterparty risk on both sides. Other hedges will be necessary to shelter from basis and counterparty risk.

Supposing that the position can be hedged satisfactorily, we can see the **swap dealer as doing essentially the same thing as a government security dealer**, but in corporate bonds instead of governments. There is a term structure in swaps, as a markup over Treasuries.

Just as we found it useful to distinguish private repo from Fed Funds, international Eurodollar futures from Treasury futures, so also we have the private swap term structure versus the Treasury term structure. At every maturity there is a rate of interest determined in the private credit markets and there is one determined in the government credit market. We can think of the swap term structure as a markup on Treasuries, or the Treasury term structure as a discount from the swap term structure. Speculators on the spread at various maturities keep the markets in touch.

The swap market is huge, now larger than the bond market itself (as measured by size of notional swap principal). The deep question raised by the enormous success of this market is whether in some sense it is coming to replace the government market. In good times, it seems there is just as much liquidity in swaps as in governments, but in bad times (tight liquidity) the hierarchical structure is restored. Swap dealers are apparently market makers at a lower level in the hierarchy of money than government bond dealers, who are themselves below the banks that make markets in money, who are themselves below central banks.

Money market swaps

The short term swap market is more of an interbank (rather than intercorporate) market. For example (p 895-6), a bank AA may take in a one-year deposit and swap fixed for 3mo LIBOR, by selling the IMM money market swap. Morgan buys the swap (and hence is short fixed) and immediately hedges with a long futures position. Why? 897 "A swap is a strip of FRAs, and a FRA is a single-set swap". A swap is like a portfolio of financial forward contracts. This is important, forwards not futures, so the futures hedge involves exposure to liquidity risk. Hence the futures contract is there only temporarily as a hedge, until Morgan can maybe make a one year loan to an LBO and close out the futures contract. LBO borrows at 3mo LIBOR, swaps flex for fixed, and we're done. In this way the one-year fixed rate deposit winds up financing Morgan's one year flex rate loan. The swap between the banks hedges interest rate exposure.

Read through the example in Stigum and try to work out the final balance sheet relationships, as follows:

AA		Morgan		LBO	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
[1 year fixed	1 yr deposit 3 mo floating] Swap at 4.44	[3 mo floating Swap at 4.44 [1 year fixed	1 year fixed]		3 mo floating 1 year fixed]
			3 mo floating] Swap at 4.47	[3 mo floating Swap at 4.47	

You can see the first swap between AA and Morgan (I show the implicit parallel loan in brackets above the swap), and then the second swap between LBO and Morgan. (I've left off the futures hedge which is there only temporarily in the time between doing the first swap and doing the second.) Basically AA is borrowing fixed but wants to borrow flex, and LBO is borrowing flex but wants to borrow fixed. Instead of doing a swap with each other,

they each separately do a swap with Morgan. Morgan winds up with equal and opposite swap exposures, i.e. matched book.

How does Morgan make money if it is short a swap at 4.47 and long a swap at 4.44? Here the parallel loan interpretation makes everything clear. Morgan is paying Libor and receiving Libor, so these flows net out. But on its long swap it is paying 4.44 fixed and on its short swap it is receiving 4.47 fixed. This is a 3 bp net profit.

Significance

What is going on here? In the marketplace as a whole, there are a range of borrowers and lenders, each with its own preferences about maturity (and currency), and each with its own market access. What is happening in the swap market is that they are each finding lowest cost financing by matching up their needs with the market as a whole.

The effect of the swap market is to spread stresses in one place and at one time, across the system and across time, and to unite the individual markets into one big market. In this way, any imperfections in the markets caused by regulation or intervention can be evaded. If central banks raise short rates in one area, making it unattractive for short term lending, then borrowing can happen elsewhere and swap into the desired currency.

There is a very significant passage on p. 900:

"Money market swaps occur in what could rightly be called arbitrage land. Traders arbitrage swaps against futures, swaps against cash, swaps against FRAs, FRAs against futures, and so on. Arbitrage opportunities keep arising because these related markets are constantly affected by many different events. Maybe an Asian bank does a big cash-and-swap arbitrage, which drives up the swap market; this creates profit in the swap-FRA arbitrage, so someone does the swap-FRA arbitrage, which drives up FRAs; this creates profit in the FRA-futures arbitrage, so someone does the FRA-futures arbitrage, which drives up futures. An event that moves one rate
causes a rate ripple that creates some basis points for every play except may the futures player if
he is an unhedged spec. Clearly, someone loses, usually the spec player in the futures pit."

Note the hierarchical structure of the sequence of trades portrayed here: swap market (OTC) to FRA (interbank forward) to futures. We have already seen that the flexibility of this system depends on ability to hedge immediately in highly liquid futures markets. Now we see also that, although in one sense the system works to spread stresses, in another sense it concentrates stresses. Everything comes back to futures.

20. Credit Derivatives

FT publishes various credit indices. See www.markit.com and click through products to find information on various indices, specifically CDX (a portfolio of corporate bonds), and ABX (a portfolio of top tranche mortgage CDOs).

To keep us focused on the big picture, I begin with a quotation from Fischer Black, from an unpublished paper way back in 1970. He writes:

> "Thus a long term corporate bond could actually be sold to three separate persons. One would supply the money for the bond; one would bear the interest rate risk; and one would bear the risk of default. The last two would not have to put up any capital for the bonds, although they might have to post some sort of collateral."
>
> **Fischer Black, 1970**
> "Fundamentals of Liquidity"

You can see that he is thinking about the kind of world that we live in today; the latter two instruments he mentions are today's interest rate swaps and credit default swaps, and the collateral he mentions is the margin that participants in these markets may have to post to ensure performance.

The main instruments we'll be talking about today are credit default swaps (CDS), and collateralized debt obligations (CDO). Both of these arose initially to handle problems of corporate credit risk, so the underlying assets were corporate bonds and bank loans. There are extensions however to the case where the underlying assets are sovereign bonds and loans, and commercial and household mortgages. It is the latter extension that is important to understand for the current subprime mortgage crisis, but we'll have to get to it in stages.

From a finance view (which is Fischer Black's approach), the key idea for understanding credit derivatives is the following:

Price of risky asset + Price of insurance on risky asset = Price of riskfree asset

Or

Yield on risky asset = Yield on riskfree asset + Credit Risk Premium

Thus, from this perspective, credit default swaps look like a kind of credit risk insurance.

From a money view, the instruments look a bit different. Translating Fischer Black into our parallel loan balance sheet construction, we can see how credit default swaps enter the picture analogously to interest rate swaps:

Me Buyer of Insurance Long Swap Short Credit Risk			Seller of Insurance Short Swap Long Credit Risk
Assets	Liabilities	Assets	Liabilities
Corporate bond [Treasury bond [Treasury bill Credit default swap Interest rate swap	 Corporate bond] Treasury bond]	 [Corporate bond [Treasury bond	 Treasury bond] Treasury bill] Credit default swap Interest rate swap

In brackets I am showing the parallel loan construction—one side promises to make the same payments the corporation makes (and to miss the same payments the corporation misses), while the other side promises to make the same payments that the Treasury makes on a bond of the same maturity. So long as the corporate bond does not default, this swap of IOUs involves a net cash flow from the long swap to the short swap, simply because the coupon on the corporate bond is larger than the coupon on the Treasury. In the event the corporate bond defaults, however, there is a large cash flow in the opposite direction; in effect the long swap delivers the defaulted bond to the short swap, and receives in return a perfectly good Treasury bond.

Corporate Bonds

Conceptually, it is easiest to introduce the basic ideas while thinking of the underlying as a corporate bond. Such bonds are often complicated instruments, because of the various warrants attached (call provisions and protection), but we'll abstract from that and think of them as simply promises to pay a certain **coupon** at regular intervals over the life of the bond, and the **face value** upon maturity. Standard valuation considerations suggest that the price of such a bond can be thought of in present value terms as

$$P(0) = \Sigma\delta^t C_t + \delta^T F_T$$

where C is the coupon, F is the face value, and δ is a discount factor that we can think of as (1/1+R) where R is some risk-adjusted interest rate. You can see from the formula that there is an inverse relationship between R and P. Fluctuations in the price of the bond after issue can be thought of as fluctuations in the risk-adjusted interest rate.

One reason for those fluctuations is fluctuation in the risk-free interest rate, but we've already talked about that in previous lectures. Here we want to focus on fluctuations that come from

(1) changes in the price of credit risk and
(2) changes in the quantity of credit risk.

Bonds are typically rated by one of the various rating agencies, such as Fitch. Ratings go from AAA to B-, to NR. Lower rated bonds sell at a discount, which means that the risk-adjusted interest rate is higher. Usually however there is an attempt to set the coupon at a level that counteracts this effect so that, at least when the bond is issued, bonds of various ratings all sell near par.

At any moment in time there is a pattern of credit spreads over Treasuries, or perhaps over the swap rate, small spreads for AAA and larger spreads for lower ratings. The important point to realize is that these spreads fluctuate over time, which is one source of risk (price), and that individual bond ratings can also change over time, which is a second source of risk (quantity). The basic idea of credit derivatives is to create an instrument that will allow these sources of risk to be carved off of the bond and priced (maybe even sold) separately.

Why do the spreads move as they do? In the industry, most of the attention focuses on default. As a matter of fact there is considerable dispute about what exactly constitutes default, but we'll put that aside. The point to hold on to is that the lower rated bonds tend to have higher default risk. Thus, instead of getting 10 years of coupons followed by the face value, you might only get 3 years of coupons, and the liquidation value of the bond. That's the risk for someone who owns the bond. Credit default swaps are a way to sell off that risk while retaining ownership of the bond.

CDS Pricing

Suppose that I own a bond that promises to pay a constant coupon C for 10 years.[1] Suppose further that I buy an interest rate swap in which I pay fixed and receive LIBOR. In effect, my combined portfolio now pays LIBOR + S%, where S is the credit spread over the fixed rate on the swap. Now we're ready to think about selling off the credit risk.

Assets	Liabilities
Corporate Bond (LIBOR + S%)	
If no default:	
[LIBOR	LIBOR + U%]
If default:	
[Face value of Bond, F	Liquidation Price of Bond, P]

As I mentioned at the beginning, a credit default swap can be understood as a swap of IOUs. I issue an IOU that promises to make periodic payments of LIBOR+U% on the face value of the bond, so long as the bond issuer keeps up his own payments. In the event of default, I promise to pay the liquidation value of the bond and then we are done. Call my IOU a "mirror bond", because the time pattern of payments exactly mirrors the corporate bond I'm holding. In return for this promise I accept an IOU that promises to

[1] This example is inspired by Duffie and Singleton, p. 180, though they don't do the full parallel loan analysis.

pay simply LIBOR as long as the issuer keeps up his payments, and the full face value of the bond in the event of default. Thus on net I am paying U until default, and then receiving the difference between the full face value of the bond and its liquidation value.

Here are the net cash flows from the swap of IOUs, assuming default at period 5:

1	2	3	4	5
-U*F	-U*F	-U*F	-U*F	+F-P(5)

I'm paying a small amount for four periods in order to receive a possibly big amount in period 5.

What is U? U is a number that makes the present value of the small payments exactly equal to the present value of the large payment, so that at inception the swap is a zero value instrument, i.e. a swap of IOUs that have the same exact value. In the finance view, we think of that U as a kind of insurance premium. So long as I pay U, I am insured against the risk of default on the bond I hold. Why so? If there is default, then the liquidation value of my bond is P(5), but my swap of IOUs pays F-P(5), so on net I recover F, the full face value of the bond.

At inception the CDS is a zero value instrument, but not after. Any change in the credit spread S%, whether market-wide or idiosyncratic to the specific bond, will change the value of the CDS. In theory, the value moves inversely to the value of the underlying bond. **In this way, CDS is not so much insurance against eventual default as it is insurance against change in the credit spread, and hence the price of the bond.**

Market making

Now let's think about the same transaction from the point of view of the seller of insurance, who might be a bank or an investment bank. The seller receives a stream of small payments but faces the possibility of having to make a single large payment in the event of default. It is possible to create portfolios of such swaps, which pool the idiosyncratic default risk so that the risk of the pool is less than the risk of any component. (This depends crucially on the correlation of the individual risks. The less correlation, the better, and the law of large numbers can help us out.) This pooling then allows the seller of the swap to charge a lower spread to the buyer.

Buyer of Insurance		Dealer		Seller of Insurance	
A	L	A	L	A	L
Risky Bond CDS_i		CDS on index CDX	CDS_i CDS_j CDS_k		CDS on Index CDX

This pooling is very important. In a sense we can say that the incentive to create a swap comes from mis-pricing of credit risk in the original corporate bond, and so we should expect that the creation of a flourishing swap market will reduce the price of credit risk overall. That is exactly what happened.

Diversification reduces risk, but does not eliminate it. If the seller seeks matched book, this pool of OTC swaps might then be hedged against a general bond index, perhaps by trading an exchange traded swap such as CDX with a hedge fund. (In the case of subprime mortgages, the swaps on the ABX index were key.)

UBS Example: Market Making and Liquidity Risk

In fact, if U<S, then the buyer of the swap can swap out the credit risk and wind up with an investment that pays a small spread (S-U) over LIBOR. That is what UBS was doing in its most important risk arbitrage trade, as follows

Assets	Liabilities
AAA CDO tranche, floating rate Credit default swap (AIG)	Money Market Funding (ABCP, RP)

UBS was doing something it called a Negative Basis Trade in which it paid AIG 11 bp for 100% credit protection on a supersenior CDO tranche, and financed its holding of that tranche in the wholesale money market. In its report to shareholders[2], to explain why it lost so much of their money, it states that this trade netted an apparently riskfree arbitrage profit of 20 bp. Because it was apparently riskfree, they did massive amounts of it. The risk turned out to be liquidity risk, when money market funding dried up and they could not sell their AAA tranche. Their CDS hedge did them no good since they could not use it to raise funding. (To make matters worse, the CDS hedge was typically only against the first 2% loss, leaving UBS exposed for everything more than that.)

Goldman Sachs Example: Hedging CDS with CDS

The recent SIGTARP report on AIG[3] explains the relationship between Goldman and AIG, which I translate into our balance sheet language as follows

Goldman Sachs AIG

Assets	Liabilities	Assets	Liabilities
CDS, AIG	CDS, clients		CDS

Goldman Sachs was in effect acting as a CDS dealer, selling protection to clients but buying protection from AIG. AIG was a naked seller of protection. When the referenced

[2] Google "UBS Shareholder Report" and you will find it.

[3] http://www.sigtarp.gov/reports/audit/2009/Factors_Affecting_Efforts_to_Limit_Payments_to_AIG_Counterparties.pdf

risky asset started to fall in price, the value of the insurance rose. This is a liability of AIG, so it cut into their capital buffer (AIG had no dedicated reserves against these CDS because it thought they were essentially riskfree). Not only that, but AIG had agreed to post collateral, and mark the CDS to market, so these losses were not just book losses but involved payments into a segregated account that Goldman Sachs controlled, about 30 billion at the time AIG failed.

AIG failed because it was no longer able to meet these collateral calls; instead the government took over, lending 85 billion. There has been a lot of loose talk about how the government paid off Goldman at par instead of forcing Goldman to take a loss. This is not exactly what happened. Rather, because the CDS was marked to market, Goldman already had possession of the collateral, it had already been paid. The government money was used to acquire the referenced securities at liquidation value in order to end the swap. In terms of our algebraic example, AIG had already paid F-P(5); what the government did was to lend AIG money to pay P(5) for the bond which the government then took onto its own (the Fed's) balance sheet as Maiden Lane 2 and 3.

Goldman Abacus example: Synthetic CDO as Collateral Prepayment

Paulson, the hedge fund manager, paid Goldman to help him bet against subprime, and the way he did it was by establishing a so-called "synthetic CDO".[4] (The very first CDS was established in this way for JP Morgan to hedge tail risk its portfolio of corporate loans. See Gillian Tett, Fool'sGold.) Here is a simplified version of the balance sheets:

Paulsen		Abacus		IKB	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
CDS on RMBS		T bills	CDO tranche CDS on RMBS	CDO tranche	

Abacus is a synthetic CDO, not a cash CDO, because its exposure to credit risk comes from its CDS position, not from any actual holdings of RMBS (residential mortgage backed securities). It sells that exposure to IKB in the form of bonds (CDO tranche), and invests the proceeds in Treasury bills. (The point is that the combination of long riskfree securities and short CDS is equivalent to an outright position in the referenced risky security. Refer back to the arbitrage relation I mentioned at the beginning of lecture.)

How was this supposed to work? So long as the RMBS was not in default, Paulsen paid a regular premium to Abacus, which added that premium to the Tbill return to pay interest to IKB on its bond holding. Once the RMBS was in default, however, Abacus paid Paulsen by delivering its holding of Treasury bills, while Paulsen delivered the underlying RMBS which he bought at liquidation value. The difference between face

[4] See The Greatest Trade Ever, by Greg Zuckerman

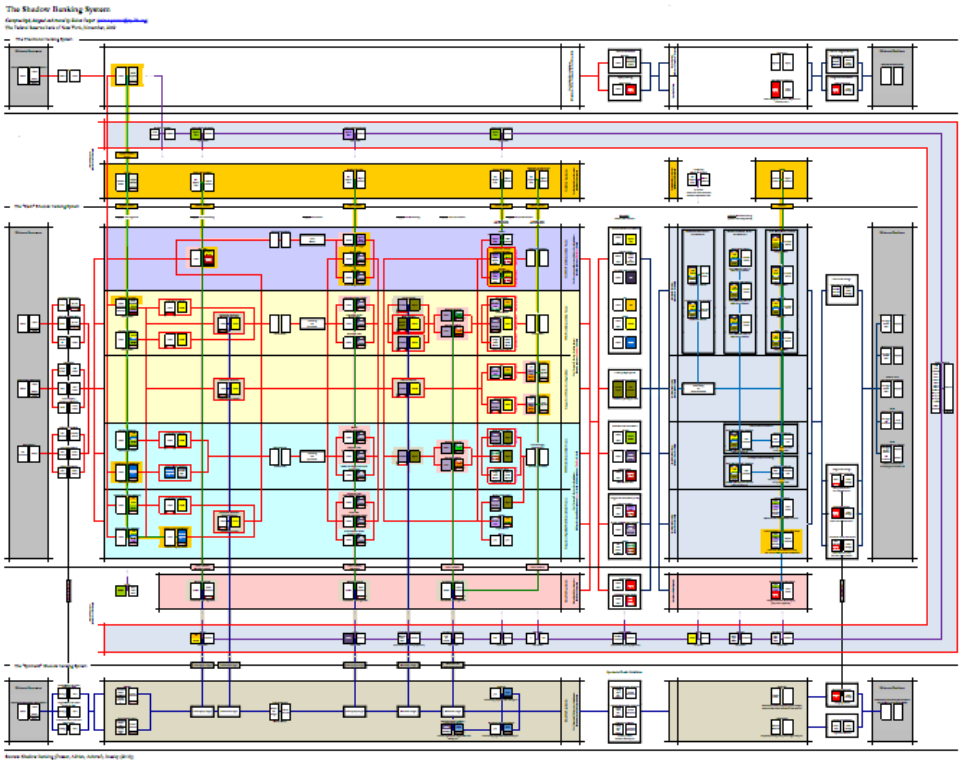
value and liquidation value is absorbed in the value of the CDO tranche bonds owned by IKB.

For this class the important point is that, in effect, the Abacus arrangement went one step farther than the AIG mark to market CDS. In the AIG case, falling value of the referenced securities forced collateral payments to Goldman Sachs. In the Abacus case, the collateral payments were all made at the very inception of the contract when IKB bought the bonds, so IKB did not have to come up with any additional collateral (which it could have refused). Instead, the falling value of the referenced securities merely caused a transfer of the collateral, already collected, from Abacus to Paulsen.

Darell Duffie and Ken Singleton, Credit Risk: Pricing, Measure, and Management. Princeton UP, 2003
David Lando, Credit Risk Modeling, Theory and Applications. Princeton UP, 2004.
Gunter Meissner, Credit Derivatives, Application, Pricing, and Risk Management. Blackwell 2005.

21. Shadow Banking, Central Banking, and the Future of Global Finance

What is a bank?



1. Shadow Banking as …. Market-based Credit

Securitization

Assets	Liabilities
RMBS	Hi tranche Mid tranche Lo tranche

Shadow Bank

Assets	Liabilities
Hi tranche	RP

MMMF

Assets	Liabilities
RP	“deposits”

Traditional Bank

Assets	Liabilities
FDIC backstop	Deposits
Reserves	Capital

kstop

Immature Liquidity Backstop

Securitization		Shadow Bank		MMMF	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
RMBS	Hi tranche Mid tranche Lo tranche	Hi tranche Liquidity put	RP	RP Liquidity put	"deposits"

Traditional Bank

Assets	**Liabilities**
Loans Reserves	Deposits Capital Liquidity put

kstop

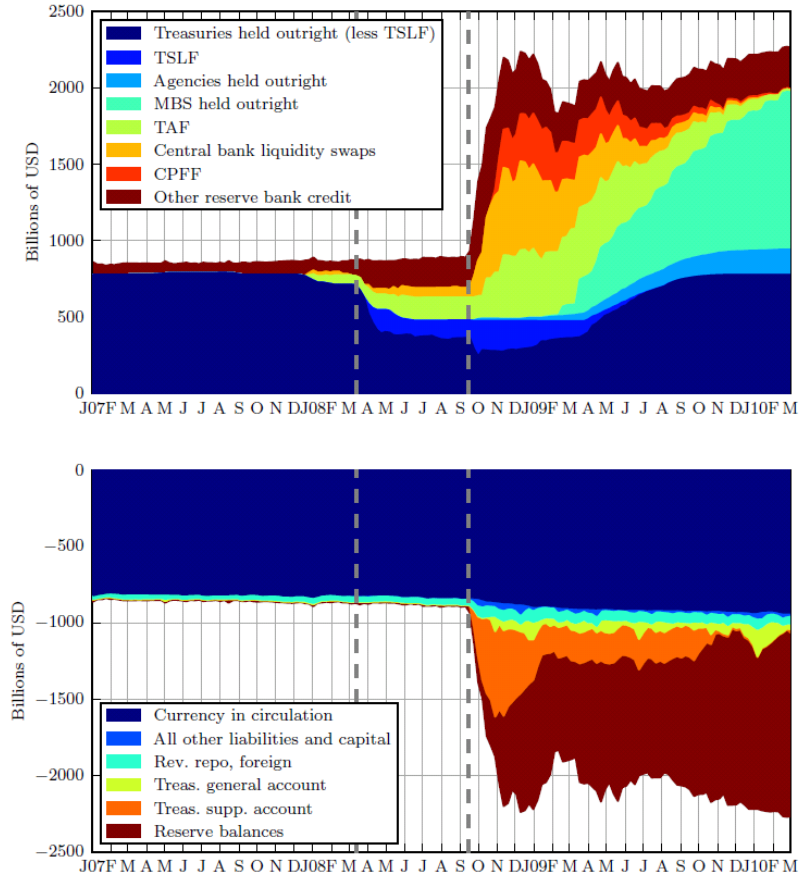
FDIC backstop

Immature Solvency Backstop

CDO		Invest. Bank		Insurance	
Assets	Liabilities	A	L	A	L
RMBS	Hi Tranche Mid Tranche Lo Tranche	**Hi CDS** Loans	**Mid** **CDS Lo** **CDS**		**Hi CDS** Capital

Shadow Bank		Pension/Insurance		Hedge Fund	
Assets	Liabilities	A	L	A	L
Hi Tranche **Hi CDS**	MM Funding	Mid Tranche **Mid CDS**	DB/Annuity Capital	Lo Tranche **Lo CDS**	Loans Capital

Fed as Dealer of Last Resort



2. Shadow Banking as…. Global Banking

Korean Bank		Global Bank		Dollar Bank	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
Domestic currency loans FX swaps $ reserves	Dollar funding	Dollar lending	Wholesale money market	Wholesale money market	"deposits"

RMBS		Shadow Bank		MMMF	
Assets	**Liabilities**	**Assets**	**Liabilities**	**Assets**	**Liabilities**
Mortgage loans	Hi tranche Mid tranche Lo tranche	Hi tranche	ABCP	ABCP	"deposits"

From Domestic to International LOLR

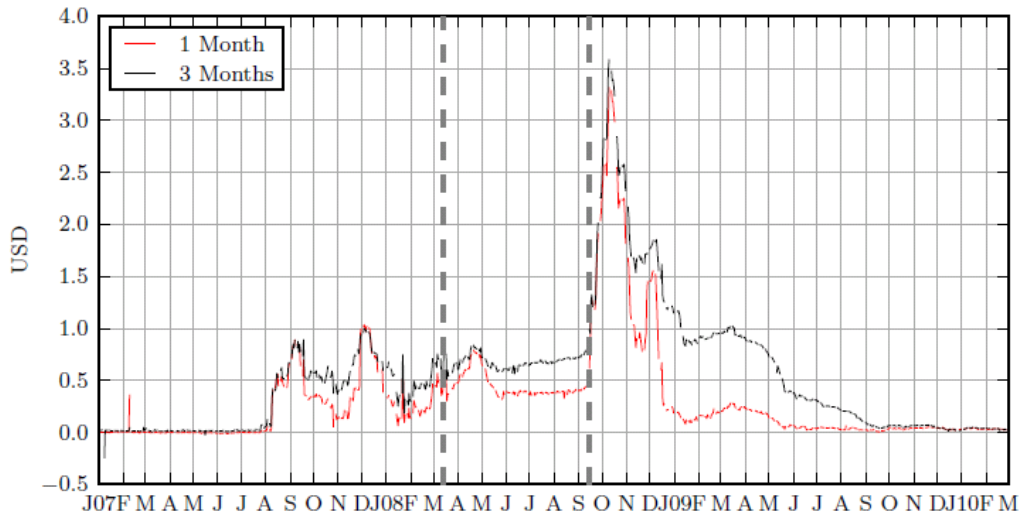


Figure 6: USD OIS–LIBOR spread, 1 and 3 months.

3. Shadow Banking as …. Modern Finance

• Modern Asset Management
–**Financial Globalization** and Securitization
–**Capital Markets** and Money/Derivatives Markets

Shadow Bank		Asset Manager	
Assets	**Liabilities**	**Assets**	**Liabilities**
RMBS **Derivatives**	**Money**	**Money**	Capital **Derivatives**

What is shadow banking?

"Money market **funding** of Capital market **lending**"

•**Global** funding of **local** lending
•**Market pricing**, both money and capital
•Key role of **market-making** institutions

Other definitions

•Received views…
–Gorton and Metrick (2011) "Regulating the Shadow Banking System"; Adrian and Ashcraft (Oct 2012) "Shadow Banking: A Review of the Literature"; FSB (Nov 2012) "Strengthening Oversight and Regulation of Shadow Banking"
•Emphasize "shadow" versus "traditional" banking
–Bank vs. non-bank
–Regulated vs. non-regulated (evasion)
–Direct government backstop vs. indirect (unauthorized)
•Versus our emphasis on banking evolution
–Financial globalization: dollar funding, reserve currency holding
•Cf. Shin (2011) "Global Banking Glut and Loan Risk Premium"
–Financial revolution: derivatives, risk pricing markets
•Imagine a world of ONLY shadow banking

"Making" Markets, Money and Risk

Capital Funding Bank | Global Money Dealer | Asset Manager

Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
RMBS CDS IRS FXS	MM funding	MM funding	"deposits"	"deposits"	Capital CDS IRS FXS

Derivative Dealer

Assets	Liabilities
Credit Default Swaps Interest Rate Swaps FX Swaps	Credit Default Swaps Interest Rate Swaps FX Swaps

Idealization

•We **abstract from**…
–Retail depositors, security investors, traditional banks
–Securitization process (underwriting, legal basis)
–Liquidity reserves (Tbills and cash)
–Capital reserves (Haircuts)
–Proprietary dealing (price risk, "leverage")

•In order to **focus on**….
–System interlinkages and system behavior
–Normal liquidity risk, not tail solvency risk
–Tail liquidity backstops, both funding liquidity (money) and market liquidity (capital)

Financial Stability Board (WS5)

•"Securities lending and repo markets play crucial roles in supporting price discovery and secondary market liquidity for a variety of securities issued by both public and private agents."

Backstopping Market-making

- Global Bank (funding)

Assets	Liabilities
MM funding Liquidity put (market)	"deposits"

- Derivative Dealer (risk)

Assets	Liabilities
CDS IRS FXS Liquidity put (market)	CDS IRS FXS

- **Central Bank (or C5)**

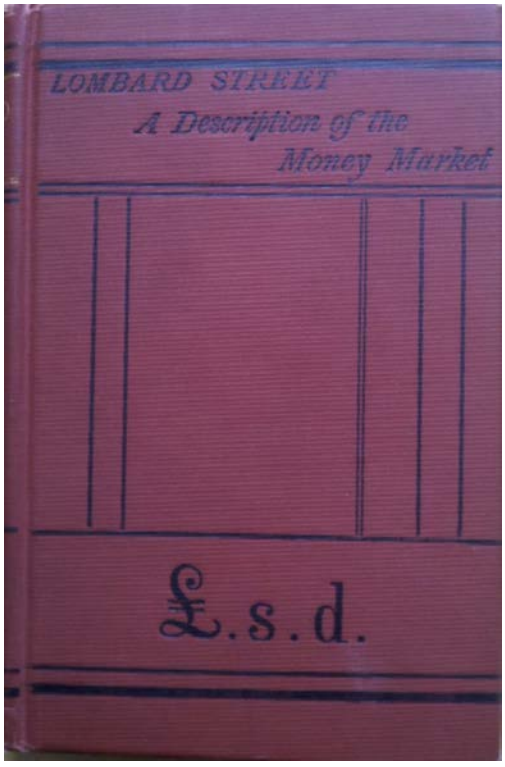
Assets	Liabilities
	Liquidity put (funding) Liquidity put (market)

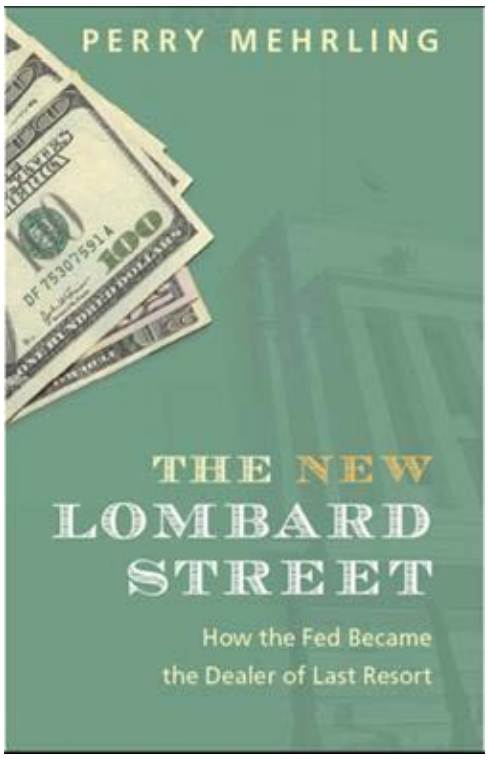
Fed Balance Sheet Transformed

Dec 15, 2011				Dec 15, 2011 (restated)			
Assets		**Liabilities**		**Assets**		**Liabilities**	
Treasuries	$1.7	Currency	$1.0	[TBill	$2.6	Curr./Res.	$2.6]
MBS/GSE	.9	Reserves	1.6	[Tbond	2.6	Tbill	2.6]
Other	.3	Other	.3	[Risky Secs	.9	Tbond	.9]
				Other	.3	Other	.3
TOTAL	2.9	TOTAL	2.9				
				TOTAL	6.4?	TOTAL	6.4?

•Fed as Dealer of Last Resort
–Global Money Dealer (OIS) [and FX swap]
–Global Derivative Dealer (IRS, CDS)

A Bagehot Moment



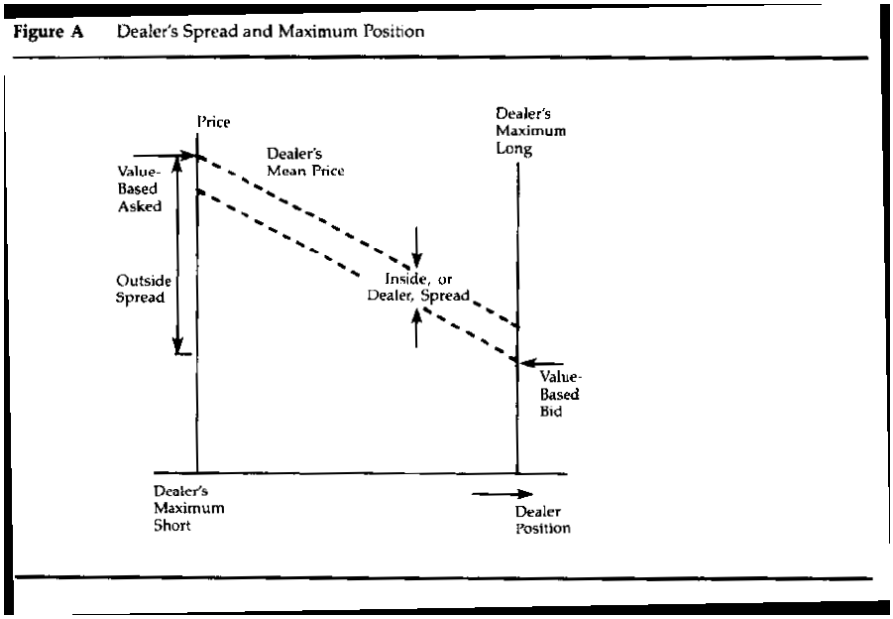


Regulation I: Systemic Risk "Inherent Instability of Credit"

•Hawtrey (1919) Currency and Credit

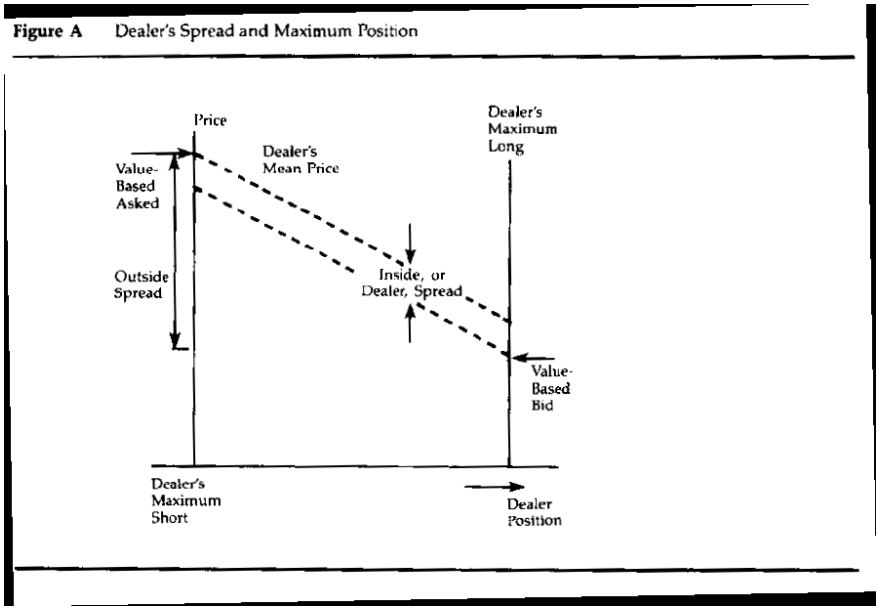
•Minsky (1986) Stabilizing an Unstable Economy
•Adrian and Shin (2010) "Liquidity and Leverage"

Centrality of the Dealer Function

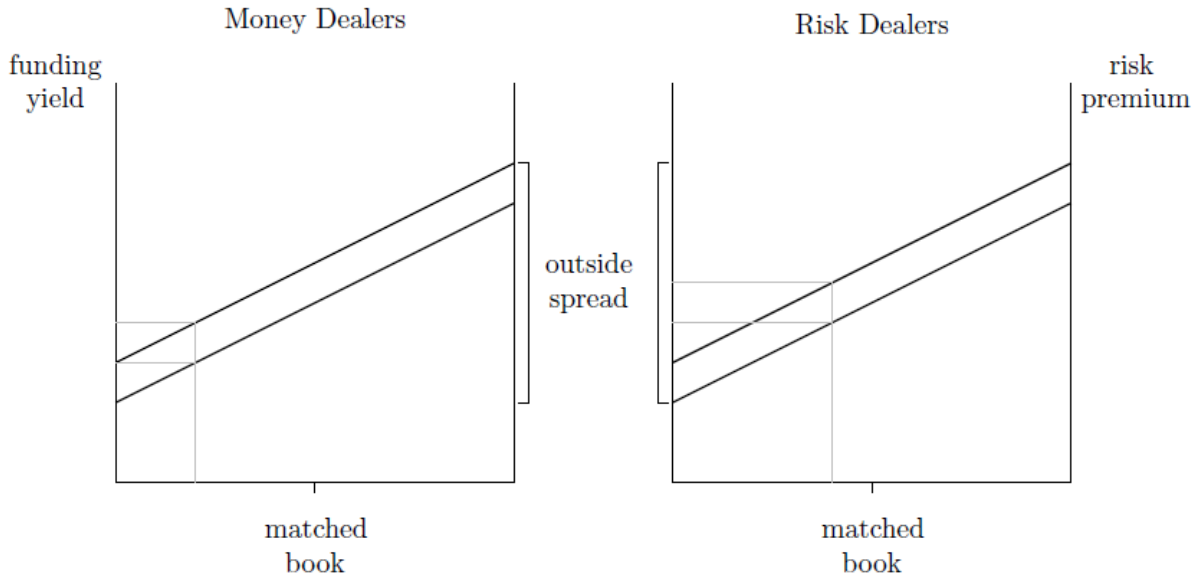


•Treynor, 1987

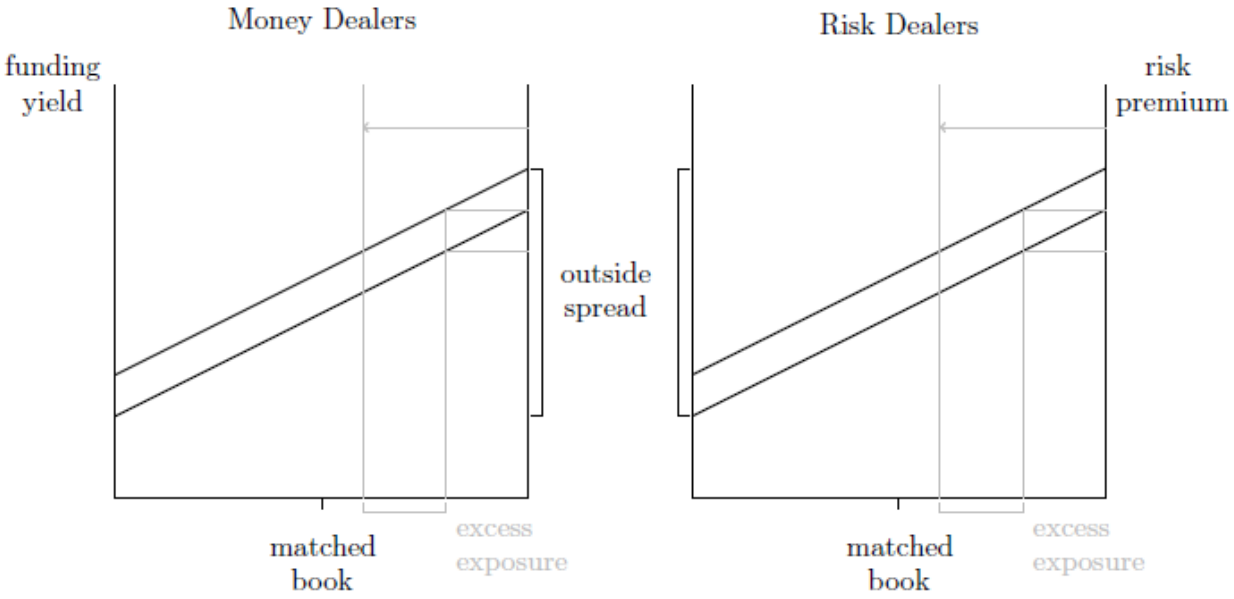
Funding Liquidity and Market Liquidity



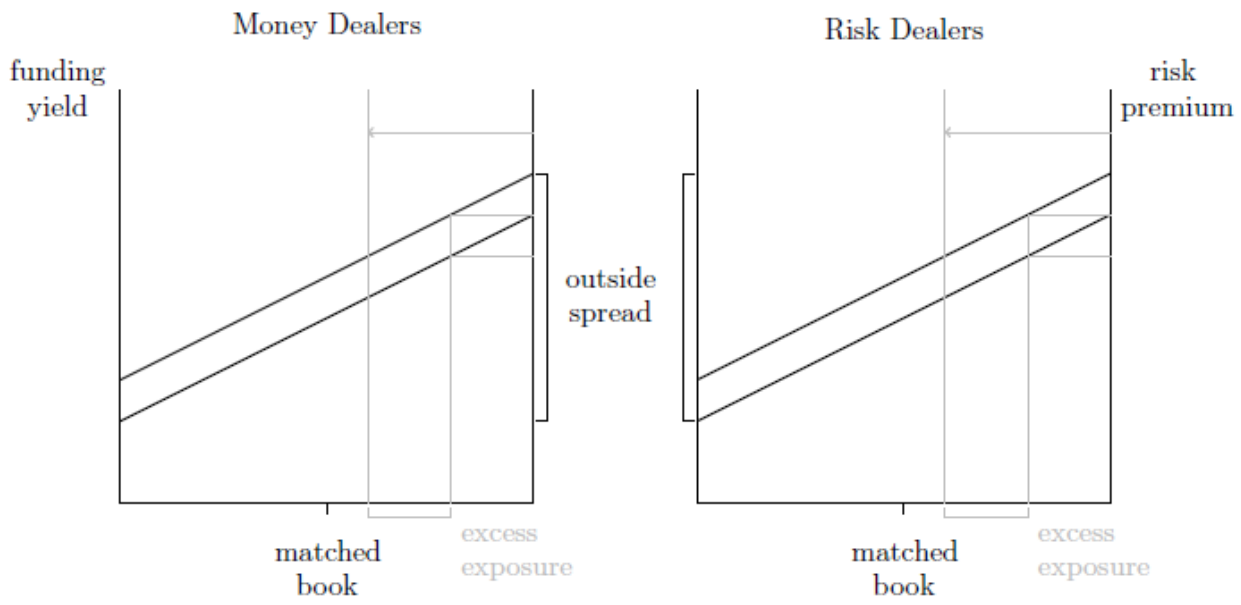
Boom Order Flow: Asset Manager > Capital Funding Bank



Bust Order Flow: Asset Manager < Capital Funding Bank



Bust Order Flow: Dealer Funding Stress



Fed Balance Sheet Transformed

Dec 15, 2011 Dec 15, 2011 (restated)

Assets		Liabilities		Assets		Liabilities	
Treasuries	$1.7	Currency	$1.0	[TBill	$2.6	Curr./Res.	$2.6]
MBS/GSE	.9	Reserves	1.6	[Tbond	2.6	Tbill	2.6]
Other	.3	Other	.3	[Risky Secs	.9	Tbond	.9]
				Other	.3	Other	.3
TOTAL	2.9	TOTAL	2.9				
				TOTAL	6.4?	TOTAL	6.4?

•Fed as Dealer of Last Resort
–Global Money Dealer (OIS) [and FX swap]
–Global Derivative Dealer (IRS, CDS)

Regulation II: Plumbing Collateral and Payment Flows

•Aitken and Singh (2010) "The (sizable) role of rehypothecation in the shadow banking system."
•Minsky, "survival constraint", settlement in the payments system

Backstop for Normal Times?

CFB

Assets		Liabilities	
RMBS	100	MM	100
CDS	0		
IRS	0		

GMD

Assets		Liabilities	
MM	100	Deposit	100

AM

Assets		Liabilities	
Deposit	100	Capital	100
		CDS	0
		IRS	0

DD

Assets		Liabilities	
CDS	0	CDS	0
IRS	0	IRS	0

•**Abstracting from Counterparty Risk**
–Lending (RMBS collateral as security)
–Derivatives (Pre-funding and hedging)

Value Fluctuation and Collateral Flows

CFB

Assets		Liabilities	
RMBS	90	MM	100
CDS	10		
IRS	0		

GMD

Assets		Liabilities	
MM	100	Deposit	100

AM

Assets		Liabilities	
Deposit	100	Capital	90
		CDS	10
		IRS	0

DD

Assets		Liabilities	
CDS	10	CDS	10
IRS	0	IRS	0

• **Blocking a liquidity spiral (firesale of RMBS, CDS)**
– Lender of Last Resort to GMD, DD
– Dealer of Last Resort for RMBS, CDS

Value Fluctuation and Payment Flows

CFB GMD AM

Assets		Liabilities		Assets		Liabilities		Assets		Liabilities	
RMBS	90	MM	100	MM	100	Deposit	100	Deposit	100	Capital	90
CDS	10									CDS	10
IRS	0									IRS	0

DD

Assets		Liabilities	
CDS	10	CDS	10
IRS	0	IRS	0

• **Blocking a liquidity spiral (firesale of RMBS, CDS)**
– Lender of Last Resort to GMD, DD
– Dealer of Last Resort for RMBS, CDS

“Making” Markets, Money and Risk

Capital Funding Bank | Global Money Dealer | Asset Manager

Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
RMBS CDS IRS FXS	MM funding	MM funding Liquidity Reserves	“deposits” Capital Reserves	“deposits”	Capital CDS IRS FXS

Derivative Dealer

Assets	Liabilities
CDS IRS FXS Liquidity Reserves	CDS IRS FXS Capital Reserves

Liquidity vs. Solvency

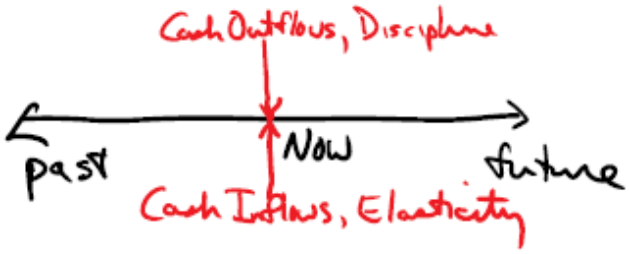
•Key players in market-based credit system are dealers, not shadow banks per se
•Key backstop for matched book dealers is liquidity, not capital
•Key backstop for speculative dealers is capital, not liquidity
•Survival constraint is about collateral flows, not just payment flows

22. Touching the Elephant: Three Views

Back in Lecture 1, The Four Prices of Money, I drew two pictures on the board, IS-LM to represent Intermediate Macro and the Fisher diagram to represent Intermediate Micro, both prerequisites for this course but neither of which I said would be of much help. These pictures capture what I call the "economics view", and I wanted to put them aside for a while in order to explore what I like to call the "money view" of the world. Today I want to introduce yet a third view, the "finance view", and in doing so begin to consider the connection between all these three views. The argument I will present is largely historical, as in history of economic thought, but the point is to help us understand the current intellectual landscape, and in particular to situate the present course in the broader intellectual discourse of both economics and finance.

Three World Views

For the last 21 lectures we have been living in the money view world, a world where almost everything that is important happens in the temporal present. We have been talking about a world in which cash inflows must be adequate to meet cash outflows (the survival or liquidity constraint) over the period of a single day. This is a period much too short for there to be any elasticity in production or consumption, the usual subject matter of economics, so we have abstracted from them. In the money view, we are looking at a world where "the present determines the present".



In the **economics view**, by contrast, we are typically looking at a world where "the past determines the present" in the sense that the current flow of goods being produced is the consequence of capital investments made over many generations in the past. The sale of those goods is the source of current income, most of which is consumed but some of which is saved in the form of additions to the capital stock available for future production.

$Y=F(K,L)$ Production

$K_t = K_{t-1} + I_t$ Capital Accumulation

Most of the debate in economics, during the last twenty years, has been not about the contrast between the economics view and money view, but rather about the contrast between the economics view and the **finance view**, which is diametrically opposite. In finance, "the future

determines the present" in the sense that current capital values are a consequence of ideas about future income flows, which we discount back to the present. Changing ideas cannot change the physical quantity of capital currently in existence, but they can very easily change the valuation of that capital, as well as the valuation of capital investments not yet made. Mere ideas thus change the current world by drawing a path that links it with some possible future world.

Economics View	**Money View**	**Finance View**
Past determines present	Present determines present	Future determines present
Capital stock, K	Cash Flow	Capital valuation, p_k
Y=f(K,L)	Survival Constraint	$P_K = E_0\Sigma(1/1+r)^tC_t$

The economics view and the finance view meet each other in the present, but the present is the realm of the **money view**. Cash flows emerging from past production meet cash commitments engaged with an eye on future production, and the balance or imbalance between the two poses the problem that is solved every day by the monetary apparatus.

Most of the history of monetary thought is about the conversation between the money view, which emerges organically from experience with the monetary system, and the economics view in the academy. But the last few decades have been an exception. Some thirty years ago, when I was starting out in monetary economics, the most significant intellectual challenge to monetary economics came from finance, not the money view. Thirty years later, that challenge has been more or less resolved (albeit with remaining unease on both sides) with the acceptance of finance as a proper subfield of economics. My own interest has been the challenge of finance to the money view, and my career has been all about building up the money view is a response to that early challenge.

To get a flavor of that challenge, listen to Fischer Black in his 1976 manifesto "What a Non-monetarist Thinks":

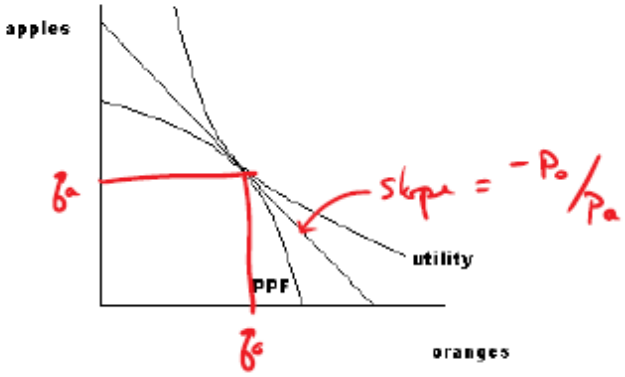
> I believe that in a country like the US, with a smoothly working financial system, the government does not, cannot, and should not control the money stock in any significant way. The government does, can only, and should simply respond passively to shifts in the private sector's demand for money. Monetary policy is passive, can only be passive, and should be passive. The pronouncements and actions of the Federal Reserve Board on monetary policy are a charade. The Board's monetary actions have almost no effect on output, employment, or inflation.

Strong words, so strong that we might be inclined to dismiss them out of hand as the ravings of the uninformed. Today I propose to resist that natural inclination, and instead try to understand the conception of the world in which such words might make sense.

Imagine A World Without Money: Commodity Exchange

Let's start with economics, which we all know, and the world of traditional value theory. When economists build theories about why goods have the prices they do, we quite typically abstract from money. This practice has allowed us to build the elegant modern Walrasian general equilibrium theory of value. It is however a theory without any place in it for money.

Just so, consider the familiar static two-good equilibrium. Here we have production possibilities and consumer preferences jointly determining the relative price of two goods. If the price of oranges is p_o and the price of apples p_a then the slope of the budget line is $-p_o/p_a$, the relative price of the two goods.



Nothing is said about money. Indeed nothing is said about the price level either.

Into this silence on the matter of money and prices enters the quantity equation, which we can write as

$MV = PQ$, or $MV = p_aq_a + p_oq_o$

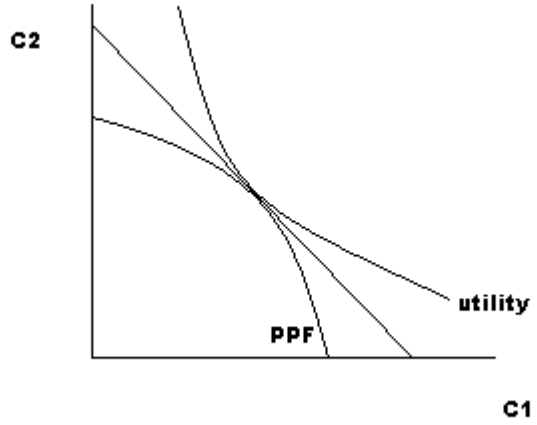
The right hand side tells the value of goods sold, and the left hand side tells how money turnover facilitated that sale. The equation as stated is an identity really, but economists have always been tempted to read causality into it. One story is that the quantity of money determines the price level—this is reading the equation left to right, which gives us the quantity theory of the price level. Another story is that the scale of transactions determines the scale of money flow (including elastic credit substitutes)—this is reading the equation right to left. We leave this debate aside, and pursue other matters for the moment.[1]

[1] Most economists, following Frank Hahn (e.g. Money and Inflation 1982), see the absence of money as a weakness of the general equilibrium theory. The theory doesn't have anything to say about the monetary phenomena and institutions that are apparently so important in the real world. The response of economists has been to try to

In the 20th century, the great Irving Fisher moved debate in a more constructive direction by expanding the idea of exchange to include intertemporal exchange. Consider thus the familiar one-good two-period equilibrium. Here we also have production possibilities and consumer preferences jointly determining the relative price of two goods. What is new here is the conception of the rate of interest as the relative price of goods between two different time periods. That rate of interest, notice, has nothing to do with money. Denoting production by $\{C_1, C_2\}$, the wealth of the representative consumer can be written

$W = p_1C_1 + [p_2/(1+r)]C_2$.

In equilibrium, the price ratio $p_1/[p_2/(1+r)]$ will be equal to the marginal rate of substitution for each consumer, and also the marginal rate of transformation for each producer.



---

develop models in which markets are less perfect, so there is a place for money. Three approaches can be distinguished:

1) Overlapping generations, money as a store of value

(Wallace "The Overlapping Generations Model of Fiat Money"1980)

2) Search and matching models

(Kiyotaki and Wright "On Money as a Medium of Exchange" 1989)

3) Walrasian GE with transactions costs

(Starr "Monetary General Equilibrium with Transactions Costs" 2005)

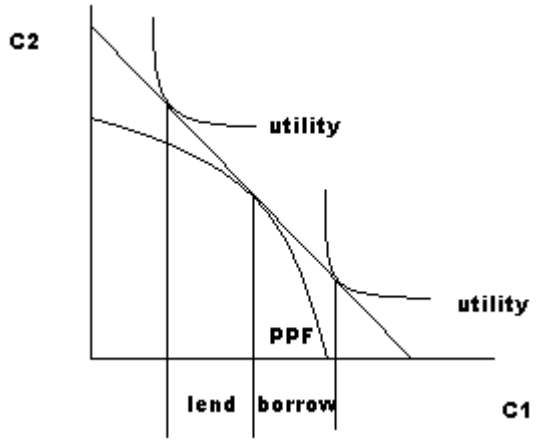
All of these take as their problem the exchange of commodities in a world without any financial assets except money (i.e. an extension of our first model), and they show that money plays a role. The problem is that, quite generally, these results do not generalize to a world with financial assets because money, paying no interest, is a dominated asset.

Again nothing is said about money or the price level, and we are tempted to try to insert some version of the quantity equation into the vacuum. But we have a problem. Future consumption is not really exchanged today, so we can't really put it on the right hand side along with present consumption. Fisher's brilliant innovation was to realize that the **market for securities is a current market with current prices where claims for future consumption are exchanged**.

How so?

To make room for loans (credit) we assume that there are two kinds of people in this world, both with the same production possibilities, but one that prefers consumption today and one that prefers consumption tomorrow. In equilibrium, both kinds of people produce the current and future consumption goods in the same proportions, but the present-oriented consume more of the present good and the future-oriented consume more of the future good.

As before, the relative price of goods today and goods tomorrow includes the rate of interest, which looks like the pure price of time. However, instead of thinking of the two consumers as trading goods across time, we think about what happens in period 1 separate from what happens in period 2. In period 1, the present-oriented consumer A consumes more of the present good than he produces, and in period 2 the future-oriented consumer B consumes more of the future good than he produces. How is this achieved? By borrowing and lending at the rate of interest. We might imagine this being facilitated by some financial intermediary. The quantity of outstanding credit then depends on the difference between people. Maybe people use these credit balances to make payments, but there is nothing about that use that affects prices.



Prefers C1 Bank Prefers C2

Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
	+loan	+loan	+deposit	+deposit	

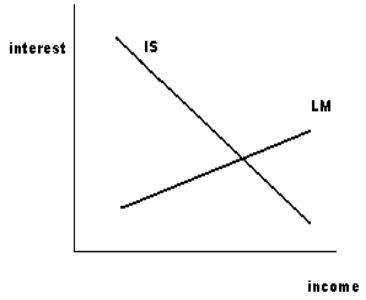
So now we know how to expand the quantity theory to an intertemporal equilibrium, but we still have a problem, no money.

One way to proceed would be simply to add securities transactions and the prices at which they are made to the commodity transactions and the prices at which they are made. Irving Fisher's transactions version of the quantity theory did just that:

$MV = PT = p_cC + (1/1+r)F$

where F is financial transactions.

Again there is temptation to read this identity as a causal equation, and there is dispute about whether causation flows from left to right or right to left. For our purposes the important point is that the left to right view suggests that monetary manipulation by the central bank affects not only the price level (price of goods) but also the price of assets, P_K, and hence also the rate of interest r. This is the origin of the idea expressed in the Hicks-Samuelson IS-LM model that the monetary authority can affect the real economy by pushing around the money supply



It is this story about money that Fischer Black was attacking so emphatically back in1976. What did he think was wrong with it?

Imagine a World without Money: Risk

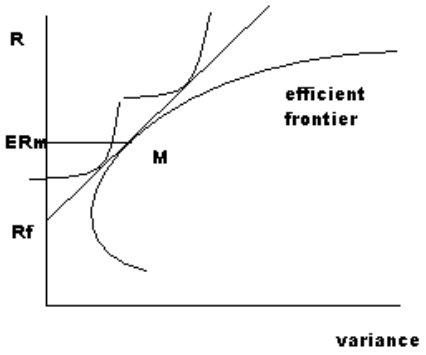
Fischer Black always started from the world of securities, not commodities, and from the Capital Asset Pricing Model. Here we have the "finance view" in its purest form. (We can think of Irving Fisher as a transitional figure between the economics view and the finance view.) We can think of CAPM by comparison to the Fisher model as involving the introduction of **risk**. In standard general equilibrium theory, we do this by distinguishing future states of the world and treating goods produced in each possible future state as different goods with different prices.[2] From a financial point of view, the important point is that the relative price of goods today and goods tomorrow includes not just the price of time but also the price of risk. The great achievement of the Capital Asset Pricing Model was to allow us to separate out these two prices and talk about them separately.[3]

Imagine a world of capital assets, each of which we characterize by the mean and variance of its

return over a short interval of time, and consider the set of expected returns and variances that can be achieved by holding a portfolio of these assets. To make the equilibrium interesting, we assume that there are two kinds of people in this world, one quite risk tolerant and the other not so risk tolerant. In equilibrium, everyone holds the same portfolio of the risky capital assets, but the risk tolerant hold more and the less tolerant hold less. This allocation is achieved by having the more tolerant borrow from the less tolerant, at the risk free rate of interest. The interaction of these two types determines jointly our two prices, the risk free rate R_f, and the price of risk ER_m-R_f, which is the expected return on the market portfolio minus the risk free rate. In equilibrium we get the famous CAPM equation

$ER_i = R_f + (ER_m - R_f)\beta_i$ Security Market Line

$ER_p = R_f + (ER_m - R_f)\sigma_p/\sigma_M$ Capital Market Line



Starting from CAPM, Fischer Black proposes that we think of banks first of all as essentially intermediaries between the risk tolerant and the less risk tolerant. As a first approximation, bank assets are the loans to the risk tolerant, and bank liabilities are the assets of the less tolerant. Thus, if each has wealth of 100, the tolerant borrow 50 in order to invest 150 in the risky portfolio, and the less tolerant lend 50 in order to invest 50 in the risky portfolio. The important point is that **the outstanding quantity of bank assets and liabilities is determined by private supply and demand, and the same is true of the interest rate**. In a CAPM world, monetary policy determines neither the quantity of money nor the price of money. Both are endogenous variables determined by private borrowing and lending behavior.

---

[3] Rubinstein has a nice paper sorting out the relationship between Arrow Debreu and CAPM.

If we think about this through the lens of the quantity equation, the point is that financial equilibrium requires us to think of the causation passing from right (PT) to left (MV). Think about what happens if the price of the market rises. Risk tolerant wants to borrow to buy more, risk averse wants to sell to lend more, so both loans and deposits change whenever the stock market changes.

Risk Tolerant		Bank		Risk Averse	
Assets	Liabilities	Assets	Liabilities	Assets	Liabilities
150 Market	50 Loan	50 loan	50 deposit	50 Market 50 Deposit	

That is the sense in which Fed monetary announcement is a charade, as Fischer Black asserts. From his point of view, the quantity of money (bank deposits) must be allowed to fluctuate freely to allow people to adjust their risk exposure as stock prices fluctuate. Similarly, the price of money (rate of interest) must be allowed to fluctuate freely to allow the market for riskless borrowing and lending to clear. In both respects there is no place for money or monetary policy.

In this world there is no role for government policy to control the money supply. (Black likes to say that there is nothing that we can call the quantity of money, and hence nothing to control.) It is equally true that there is no role for government policy to control the rate of interest. Of course the government has, historically, tried to control both of these things (p. 13-17). Black's position seems to be that such efforts produce inefficiency at the very least. In a financially developed economy, they don't even do that, since there will be multiple ways to evade control. This is the intuition behind the audacious quotation with which we started the lecture.[4]

The Education of Fischer Black

What Black observes is that, in this CAPM world, once people are using banks as their source of risk free borrowing and lending for investment purposes, it is inevitable that they would use them as means of payment as well (p. 12). People who need to make payments to one another do so simply by making book entries in the banking system. What is more, payments from a negative account (loan) seem just as possible as payments from a positive account (deposit). Such patterns of payments will affect the quantity of bank assets and liabilities, but not their price because the payments system is so efficient.

Just so, if a holder of a positive account in one bank makes a payment to a holder of a positive account in another bank, the quantity of outstanding bank credit does not change. But if a holder

[4] The economics profession has attempted to respond to the challenge of finance. To date the most successful account is Mike Woodford's Interest and Prices, which synthesizes the current consensus around some kind of Taylor Rule oriented toward inflation targeting. The focus however is entirely on inflation of goods prices, and not at all on asset prices. What remains to be done for the modern model is what Irving Fisher did for the 19th century model, i.e. expand it to asset prices by linking to finance. That's the step I am trying to point toward in this course.

of a negative account makes a payment to a holder of a positive account, credit expands. And if a holder of a positive account makes a payment to a holder of a negative account, credit contracts. (Work this out for yourself, or see previous lecture on payments system.)

This is the elasticity that we noted in an earlier lecture as being essential for smooth operation of the payments system. Note that in the finance view of the world, **the elasticity of the payments system arises from the elasticity of credit in capital markets, not vice versa as in the money view.** Private borrowers and lenders must be allowed free rein in order to achieve an equilibrium of supply and demand in their attempts to achieve desired exposure to market risk. Fischer starts at the bottom of the money credit hierarchy, whereas we started at the top, but we both arrive at a similar idea about the intertwining of the payments system with the credit system.

The big difference is that for Fischer, the only constraint is wealth, which people allocate between different assets depending on their tolerance for risk, in a market where price is efficient and there is no trouble adjusting portfolios at the margin. There is nothing here about the money view's emphasis on the survival constraint, or the role of dealers in marking markets. But wait….

Fischer started with CAPM, but he didn't stop there. On Monday, December 30, 1985, Fischer gave the presidential address to the American Finance Association which was meeting in New York, and stunned his audience with the following words:

> "We might define an efficient market as one in which price is within a factor of 2 of value; i.e. the price is more than half of value and less than twice value. By this definition, I think almost all markets are efficient almost all of the time. 'Almost all' means at least 90 percent."

Here we can detect, I think, the influence of Fischer Black's friend, Jack Treynor, who had originally introduced him to his own version of the capital asset pricing model but gone on to a life in the markets rather than academia, and in that life had produced the dealer model that we have been using in previous lectures. Think about what the dealer model says. It says, just as Fischer relates, that the price of a security fluctuates within bounds set by the value based trader, bounds that can be rather far from true value. At any moment in time, the price of the security will lie somewhere within those bounds, exactly where depends on the inventory of the dealer. This is the model in Fischer's mind when he spoke those words in 1985.

Once we bring on board the dealer model, we have the beginnings of a bridge between the finance view and the money view. It is not clear to me that Fischer walked very far along that bridge in the time before he died, but he certainly got started. Were he alive today, here is what I would urge on him as the crucial next steps.

Lesson 1: <u>Market</u> liquidity depends on the dealer system.

This is a crucial step away from the pure CAPM, and a step that reveals an important limitation of that model, namely its abstraction from liquidity. The model essentially assumes that all securities can be bought and sold in perfectly liquid markets, but this would be true only if the

outside spread, established by the value trader, was as tight as the inside spread, established by the dealer. This realization represents a first step in bringing the finance view into contact with the money view, and Fischer clearly took this step. When he died however, he was still thinking of the constraint on market liquidity as deriving fundamentally from dealer capital, not dealing funding.

The collapse of LTCM in 1998 taught a second lesson, or should have. Fischer Black died in 1995, but presciently warned about the risk involved in the relative value investment strategy that LTCM was using. In retrospect we recognize that their strategy was loading up on liquidity risk, and when liquidity dried up they could not refinance their position.

Lesson 2: The ability of dealers to provide market liquidity depends on their own funding liquidity.

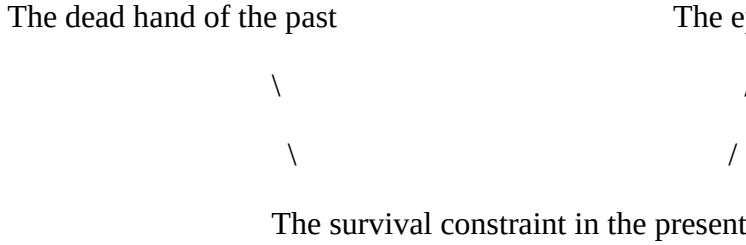
But what does funding liquidity depend on? Here we find the final link in the chain that brings the finance view into contact with the money view.

Lesson 3: The ultimate source of funding liquidity is the central bank, for the simple reason the ultimate means of payment is the liability of that bank which it can expand or contract.

By the time he died, the education of Fischer Black had proceeded through Lesson 2, but not yet Lesson 3. That is my extrapolation, and I wish Fischer had lived so I could suggest it to him. Because the implication of Lesson 3 is that monetary policy is not a charade. The answer to the challenge of finance is the money view.

The future of banking

In this course we have been conceptualizing banking as being essentially about **selective** enforcement (discipline) and relaxation (elasticity) of the survival constraint. The importance of this function is two-fold. On the one hand it means that agents who have accumulated obligations to pay that they cannot now meet can, if the banking system lets them, put off the problem until some time in the future when perhaps they will be better able to meet their obligations. On the other hand it means that agents who have brilliant plans for the future that they cannot now realize for lack of spending power can, if the banking system lets them, issue obligations to pay in the future.



In helping to postpone realization of failure, and to anticipate realization of success, banks take on risk. They make the payments that the agents are not (yet?) themselves able to make. In this sense the essential core of the banking/dealing function is the **selective bearing of liquidity risk**. So long as there is liquidity risk, there will be a role for banks, and not only that, a crucial allocative role.

From a policy perspective, the question of regulation revolves around "externalities", which is to say all the ways in which private profit motive for selective bearing of liquidity risk deviates from the larger social good. The problem is that, the more liquidity risk banks bear, the more they charge for an additional unit, until ultimately they are unwilling to bear any more and the system breaks down. We have seen how, in this instance, central banks can help by serving and lender and dealer of last resort. It follows that they can also help by intervening earlier on when the deviation between private and social begins to widen.

All of this begs the question, What is the optimal price of liquidity? The Economics View, assuming liquidity is a free good, essentially is also asserting that the optimal price of liquidity is zero. The Finance View, also assuming liquidity is a free good, is also essentially asserting that the optimal price of liquidity is zero. The Money View, by contrast to both, sees the price of liquidity as fluctuating over the boom-bust cycle. Sometimes it gets too low, sometimes it gets too high. That is where the central bank comes in. Managing the balance between discipline and elasticity means managing the price of liquidity.