

21 The inherent hierarchy of money

Perry Mehrling

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* (the means of final settlement) from *credit* (a promise to pay money, or a 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 from the standpoint of the level above.

I

To see this point more clearly, think about a monetary system under a gold standard, and think not about money and credit in the abstract but rather about concrete financial instruments: gold, currency, bank deposits, and securities (Figure 21.1).

In such a world, *gold* is the ultimate money because it is the ultimate international means of payment, and 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 as a reserve, but that doesn't mean that these currencies represent gold or are at the same hierarchical level as gold. They are still promises to pay – just more credible promises, because the presence of reserves makes it more likely that the issuer can fulfill his or her promise.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 (or deposits) over some time horizon in the future, so they are even more attenuated promises to pay. Here again, the credibility of the promise is an issue, and here again reserves of instruments that lie higher up in the hierarchy may serve to enhance credibility. In just this way, banks hold currency as reserve, but that doesn't mean that bank deposits represent currency or are at the same hierarchical level as currency.

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,



Figure 21.1 A simple hierarchy.

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 (or businesses), 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 the practice of textbooks when they speak of the money supply, although there remains 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, 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, or perhaps SDRs (Special Drawing Rights at the IMF) in the modern case. (The United States 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 that 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? It is best therefore not to reify the concepts of money and credit, and to stay instead with the more general idea that the system is hierarchical in character. As a corollary, it is best also to avoid sterile disputation between theories that arise from different levels of the hierarchy. The points of view of a central banker and a local grocer cannot be expected to coincide, unless both are willing to step out of their own positions within the hierarchy, and instead embrace the hierarchical character of the system as a whole.³

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. A few paragraphs earlier, I used currency and high-powered money as synonyms, 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 (Money Market Mutual Fund 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 talking about the hierarchy as a matter of the qualitative difference between various financial *instruments*. It is illuminating now to shift focus and see the hierarchy as a matter of the relationship between the various financial *institutions* that issue those instruments (Table 21.1).

To keep things simple, I have included only the instruments we have already been talking about, so there are important entries missing (and, as a consequence, my balance sheets do not balance). Missing are: government debt as the most important asset of the central bank, bank loans as an asset of the banking system, and the entire “shadow banking” system that finances its holdings of securities by borrowing in the wholesale money market. 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 would 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 (or credit), an analytical distinction first proposed by Gurley and Shaw in their seminal *Money in a Theory of Finance* (1960).⁴ But Gurley and Shaw treated currency as a form of outside money, i.e. so-called “fiat” money, because they aggregated only over the private economy, not including the government sector. From their point of view, currency as well as gold appears to be an asset that has no liability counterpart. In this chapter, 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

Table 21.1 The hierarchy in balance sheets

<i>Central bank</i>		<i>Banking system</i>		<i>Private sector</i>	
<i>Assets</i>	<i>Liabilities</i>	<i>Assets</i>	<i>Liabilities</i>	<i>Assets</i>	<i>Liabilities</i>
Gold	Currency	Currency	Deposits	Deposits Securities	Securities

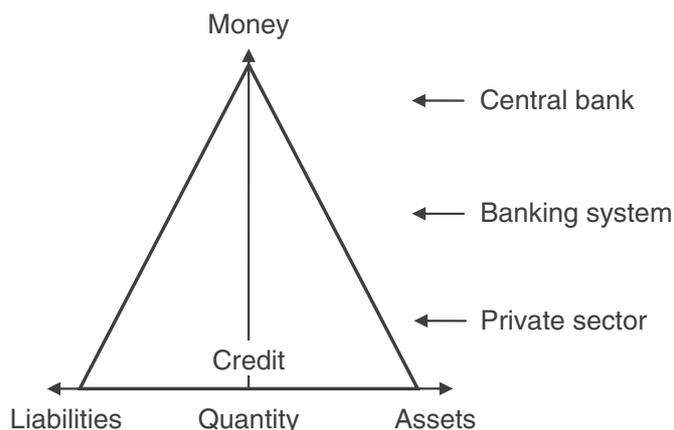


Figure 21.2 The hierarchy of money and credit.

point of view. We are taking a global view, appropriate for the modern financially globalized economy.

To consolidate the idea of an “inside” asset, and to aid intuition for what follows, it may help to visualize the hierarchy as a symmetric pyramid rising on a credit-to-money axis centered on zero, so that net outstanding credit at any level of the hierarchy is zero (Figure 21.2). 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. That means that the entire pyramid would disappear if we consolidated balance sheets, as most standard aggregative macro models do, to one extent or another. The point of the present chapter, by contrast, is to leave open the possibility that important macroeconomic variables, such as interest rates and GDP, are affected both by the gross quantity of inside credit and also by who is issuing it and who is holding it, which is to say by the precise location of that inside credit within the hierarchy of money and credit.

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 we examine, we see constant motion, expansion, and contraction. Focus on daily settlement, on the business cycle frequency, or on the longer term secular scale. You’ll see daylight overdrafts cancelling at the end-of-day clearing, credit cycles, or wars and depressions.

We can distinguish two aspects of this dynamic fluctuation. First, and most simple, is the expansion and contraction of the *quantity* of credit – at each level of the hierarchy, there is fluctuation in both assets and liabilities. In expansion mode (Figure 21.3), borrowers find that they can borrow more, and credit becomes

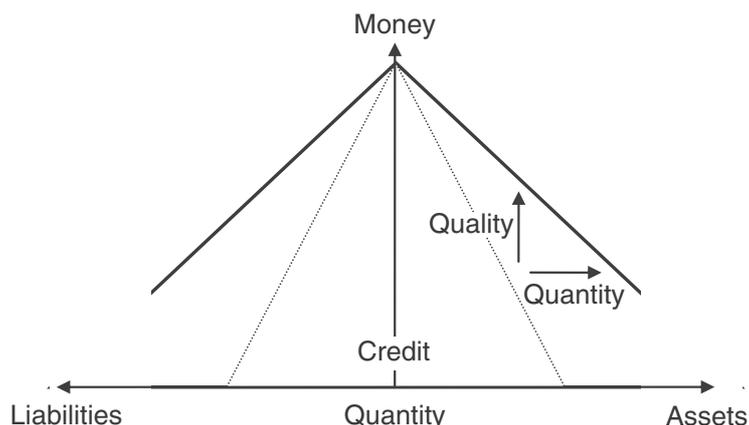


Figure 21.3 Expansion mode.

available even to marginal borrowers, while in contraction mode borrowers must repay and only the best borrowers are able to refinance their positions. Second, and more subtle, is the fluctuation of the *quality* or “moneyness” of any given type of credit. In expansion mode, the best borrowers within any level of the system find their liabilities treated as close substitutes for the liabilities of those one level higher, i.e. as money, not credit. And then, in contraction mode, differentiation returns.

At the business cycle frequency, the phenomena surrounding this expansion and contraction are grouped under the headings “irrational exuberance” in the expansion phase and “financial crisis” in the contraction phase. During the expansion phase, the qualitative difference between credit and money becomes attenuated; the quantity of credit expands while the qualitative hierarchy flattens. During the contraction phase, the distinction between more money-like and less money-like forms of credit is re-established; credit contracts while the hierarchy steepens. (One symptom of this fluctuation in quality is fluctuation in relative price, i.e. fluctuating spreads between the rates charged to qualitatively different borrowers.)

Whatever the underlying cause of fluctuation may be, we can usefully think of the fluctuation as involving a swing from one extreme to another, from *scarcity* to *elasticity* and back again. And all along the spectrum bounded by these extremes, the monetary system can be characterized as a shifting balance between these two aspects of the system. The task of monetary theory is to characterize this balance at a given moment in time, as well as the underlying dynamics that drive that balance from one extreme to another.

This two-sided character of the system under study is the reason that the history of monetary theory so largely consists of a dialog between two points of view, often distinguished as the Currency Principle versus the Banking Principle, which emphasize, respectively, the importance of scarcity and the importance of elasticity. From the point of view I have been developing, both of these

intellectual traditions have part of the truth but neither has it all, because liquidity is at the same time both scarce and elastic. How can this be so?

The scarcity of (ultimate) money ↔ **Currency Principle**

The elasticity of (derivative) credit ↔ **Banking Principle**

The 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 their own. Just so central banks cannot increase the quantity of gold, and ordinary banks cannot increase the quantity of central bank currency. At every level of the system, the availability of money from the level above serves as a disciplinary constraint that prevents expansion; credit is payable in money, but money is scarce.⁵

The elasticity comes from the fact that agents at any particular level in the hierarchy can, by their own actions, increase the quantity of credit at their own level, and also the quantity of money for the levels below them. Even at the bottom of the hierarchy, if you and I want to make a trade and you are willing to accept an IOU from me then we can trade; what makes the trade possible is an expansion of credit. The elasticity of credit thus offers a degree of freedom that relaxes the constraint posed by the scarcity of money.

This balance of discipline and elasticity applies also to banks. By trading among themselves, banks can and do relax the constraint of central bank reserves, but not forever. Credit is a promise to pay money, and relaxation today comes at the price of greater constraint some time in the future. The important point is that the system involves at all times a balance between discipline and elasticity, with sometimes one and sometimes the other aspect serving as the more dominant feature.

The fluctuation of the system from scarcity to elasticity, and back again, has its intellectual counterpart in the fluctuating dominance of different traditions of monetary thought, from the Currency Principle to the Banking Principle, from monetarism to Keynesianism, from Metallism to Chartalism, and back again. At the extreme of elasticity, calls for reassertion of scarcity gain prominence; at the extreme of scarcity, calls for reassertion of elasticity gain prominence. Neither tradition ever wins out completely, however, because the system both traditions are trying to understand has both aspects at all times.

IV

I have used the word “inherent” 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 both deep features of the system that emerge organically from the logic of its normal functioning. That is to say, the hierarchy is *not* something simply imposed from the outside, e.g. by the power of government, or the force of law. Rather, monetary systems are inevitably hierarchical, from the

inside, by the logic of their internal operations. This is probably a controversial point of view, so I had best be careful to explain what I mean.⁶

The key is to appreciate the institutions that, at each level of the hierarchy, act as market makers exchanging credit for money and money for credit. A security dealer is a kind of market maker, standing 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 in terms of currency (see Hicks, 1989). And I propose to extend the same idea 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, quoting prices at which they are willing to convert credit into money and vice versa.

Returning to 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 (Figure 21.4). And for each market maker, there is an associated price of money. There are three prices in the simple hierarchy: 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). These prices are the *quantitative* link between layers of *qualitatively* differentiated assets. The market makers who quote these prices in effect straddle the layers of the hierarchy, using their own balance sheets to knit those differentiated layers into a coherent whole.

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 differentiation between the prices

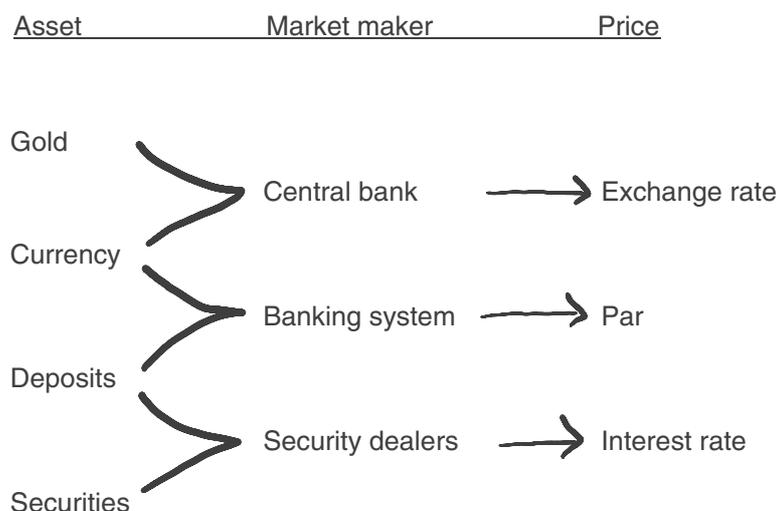


Figure 21.4 Simple hierarchy of market makers.

of various financial assets. 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 in financially underdeveloped economies), or when they are overwhelmed by the task at hand (such as under the extreme stress of war finance or during periods of financial crisis).

Even in less extreme times, the normal fluctuation of the hierarchy regularly puts strain on market making institutions. In expansion mode, it is an easy business, and highly profitable. But a contraction of credit, or steepening of the hierarchy, means an increased qualitative differentiation between credit and money, which is to say between the instruments the market maker holds as assets and the instruments it holds as liabilities. In the course of that differentiation there is bound to be pressure on the price of credit in terms of money, which pressure shows up as a solvency challenge. And there is bound also to be pressure to make good on the promised monetary qualities of the liabilities, which pressure shows up as a liquidity challenge.

The point to emphasize here is the inherent limits of price adjustment for addressing such challenges. Interest rates can and do move to reflect system stresses, but such moves are not necessarily equilibrating, and might just make the strain worse. Further, the whole point of deposit 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 is especially vulnerable whenever the hierarchy steepens, because it is bound to defend a fixed price between layers of an increasingly differentiated hierarchy. Reserves of money are the first line of defense but are soon exhausted in any significant contraction.

The problem is that, in a liquidity crisis, everyone wants money and no one wants credit. Fortunately, what counts as money at one level in the system is merely credit for the level above, and this means that higher levels of the system can generally solve the crisis of levels below them. Small crises can be solved by monetary expansion at the immediately higher level; large crises, however, may require involvement of the very highest levels. Just so, in the recent 2007–9 US financial crisis, the private banking system initially tried to absorb the brunt of the crisis, but when the crisis proved too big the central banking system had to get involved (Mehrling, 2011).

V

So far, I have been talking about market makers as a reactive bunch who respond to fluctuations that are outside their control. I have also been implicitly assuming that their behavior is driven by the dictates of profit maximization. By contrast, the whole concept of *monetary policy* rests on the idea that the central bank has some ability to influence the course of fluctuation, and some other objective than profit maximization in mind. Just so we can understand monetary policy as an attempt by the central bank deliberately to manage the dynamic fluctuation of the

system for the general good, somehow defined, rather than for its own private profit (Young, 1927).

So, for example, at a time when the elasticity principle is dominant (meaning that credit is expanding and the hierarchy is flattening) instead of waiting for a financial crisis to reassert the scarcity of money, the monetary authorities might try themselves to assert the scarcity principle 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 scarcity principle is dominant (meaning that credit is contracting and the hierarchy is steepening) the monetary authorities might try to assert the elasticity principle by lowering the Fed Funds rate (“pushing on a string” in the immortal words of John Maynard Keynes).

Observe that, in both cases, I have taken the main policy instrument of the central bank to be the overnight borrowing rate, as indeed it most often has been. However, there can be considerable slippage between that rate and the longer term interest rates on which important investment and spending decisions depend, so other policy instruments may also be necessary. It is best therefore to abstract from the precise instrument of control and instead to emphasize, more generally, that the central bank deliberately intervenes, in any number of ways, in the dynamic process, driving the fluctuation of the hierarchy between the extremes of scarcity and elasticity. In doing so, the central bank exploits its position at the top of the hierarchy, i.e. the fact that its own liabilities are more money-like than the liabilities of anyone below it.

The central bank is thus structurally in a position to intervene, but what should be the objective of its intervention? Historically, central banking can be said to have begun with the narrow objective of stemming crises, and so with an emphasis on lender of last resort intervention. Indeed, such is more or less the exclusive focus of Bagehot’s celebrated *Lombard Street* (1873); he took for granted that the proper function of a central bank was to protect deposit par and the exchange rate. If ordinary banks were having difficulty defending par, the central bank could help them out by providing additional reserves; that is the “lend freely” part of the Bagehot Rule. On the other hand, if the problem is defending the exchange rate, the central bank could raise the discount rate in order to attract additional international reserves (gold); that is the “high rate” part of the Bagehot Rule. Both of these interventions seemed natural objectives for a central bank tasked with backstopping the market making system under a gold standard.

But Bagehot was just the starting point from which sprang all the refinements of modern monetary policy. Once having taken on responsibility as crisis backstop, it was natural for the central bank to extend its remit to consider how to ward off crisis before it came. Here is the origin of the ambition of central banks to manage not only the hierarchy of money and credit, narrowly conceived, but also the broader underlying economy. Wartime economic management, when conservative central banks everywhere gave over the reins to more aggressive state Treasuries, showed what could be done under the most extreme stress. As a consequence, when the reins were handed back, central bank ambition for policy intervention in peacetime ratcheted up. The point to emphasize is that, even as

policy ambition expanded, the source of the central bank's leverage remained the same, namely its position at the top of the hierarchy of money and credit.

Inevitably this enlarged policy ambition came into conflict with the discipline of the gold standard (relaxed at Bretton Woods in 1944) and then the discipline of fixed exchange rates (relaxed in 1971 when the Bretton Woods system ended). But the relaxation of these historic disciplinary institutions did not mean the end of discipline (nor the end of scarcity). Rather, it merely shifted the origin of discipline to internal central bank practice. Further, mechanisms of central bank cooperation, devised initially as ad hoc responses to insupportable discipline, transformed into a disciplinary mechanism in their own right. The survival constraint was not eliminated, only shifted.

Today, in the aftermath of the US financial crisis of 2007–9 and in the very midst of the European financial crisis of 2011–? all eyes are on the central banks, not only the Fed and the European Central Bank, but also the Bank of Japan, the Bank of England, and the Swiss National Bank. “Forget the G7, Watch the C5” is the watchword of the day. Both crises have proved too large for the banking system where they started, and too large as well for any individual national central bank; both became global crises and consequently both rest today in the hands of the collectivity of central banks. As always, expansion was followed by contraction, and flattening of the hierarchy was followed by reassertion of differentiation between money and credit. This time is different in detail because of financial globalization and the attendant replacement of a bank-loan-based credit system by a capital-market-based credit system. But the big picture remains the same: central banks using their position at the top of the hierarchy to manage the inherent instability of credit.

Notes

- 1 As a case in point, consider Ralph Hawtrey's *Currency and Credit* (1923). My title is an homage to Hawtrey's famous emphasis on the “inherent instability of credit.” The substance of the following chapter is the product of 25 years thinking, writing, and most importantly teaching about how the modern monetary system works. My testing ground has been the classroom of Barnard and Columbia students, having in mind the wise words of Duncan Foley that, if you can't explain something to an audience of bright and interested undergraduates, then probably there remains some deficiency in your own understanding.
- 2 Put another way, 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.
- 3 Just so the perennial dispute between Banking School and Currency School has its origin in the different points of view that arise from different levels of the hierarchy. See Section III.
- 4 Stimulated by the so-called Radcliffe Report of the British Committee on the Working of the Monetary System (1959), Gurley and Shaw were primarily interested in how banking intermediaries helped to finance the private investment needs of the economy.
- 5 Minsky (1957) called this constraint the “survival constraint” in order to differentiate it from the intertemporal budget constraint of standard economics.

6 I do not mean to suggest that the monetary system is self-organizing or self-regulating in the sense emphasized by Austrian thinkers, such as Carl Menger (1892), who in rejecting top-down or outside-force theories of money also tend to miss the inherently hierarchical character of the system.

References

Bagehot, W. (1906) [1873] *Lombard Street: A Description of the Money Market*, New York: Charles Scribner's Sons.

Committee on the Working of the Monetary System (1959) [Radcliffe] *Report*, London: H.M. Stationery Office.

Gurley, J. G., and Shaw, E. S. (1960) *Money in a Theory of Finance*, Washington, D.C.: The Brookings Institution.

Hawtrey, R. (1923) *Currency and Credit*, London: Longmans, Green, and Co.

Hicks, J. (1989) *A Market Theory of Money*, New York: Oxford University Press.

Mehrling, P. (2011) *The New Lombard Street, How the Fed became the Dealer of Last Resort*, Princeton, NJ: Princeton University Press.

Menger, C. (1892) 'On the Origins of Money', *Economic Journal* 2: 239–255.

Minsky, H. (1957) 'Central Banking and Money Market Changes', *Quarterly Journal of Economics* 71: 171–187.

Young, A. A. (1927) 'The Structure and Policies of the Federal Reserve System', in *Economic Problems New and Old*, Boston: Houghton Mifflin.