**Re-theorizing Liquidity**[[1]](#footnote-1)

Perry Mehrling

Barnard College, Columbia University

January 20, 2016

 Why does “liquidity” need re-theorizing?

 Throughout the postwar period, economics emphasized a concept we can call “monetary liquidity”, which involved treating the supply of money as a decision variable of the central bank and the demand for money as a decision variable of households and businesses. In the short run, the money rate of interest was supposed to move to bring demand into line with supply; in the long run, the price level was supposed to move to bring the “real” supply of money into line with real money demand. This analytical strategy made a certain amount of sense for the time it was developed, but today it requires reconsideration in the face of far-reaching institutional change. “Funding liquidity” and “market liquidity” (Brunnermeier and Pedersen 2009), concepts that emerge from the field of finance, can be extended to the monetary sphere as a way of re-theorizing liquidity for modern times. In this frame, profit-seeking banks supply funding liquidity and profit-seeking dealers supply market liquidity, for a price.

**Settlement and the Moneyflow Economy**

 In a monetary economy everyone—households and businesses, but also governments and financial institutions—is a cash-inflow/cash-outflow entity, receiving payments from and making payments to the rest of the world. A key disciplining element faced by everyone is the constraint to settle their debts as they come due. Debts are promises to pay money, and money is the means of settling debts. This settlement constraint is asymmetric, requiring deficit agents (whose cash outflow exceeds their cash inflow) to find means of payment that surplus agents (whose cash inflow exceeds their cash outflow) are willing to accept.

 In most modern economies, a basic form of money is the issue of the state, so deficit agents can always settle by transferring accumulated cash hoards to surplus agents. Most importantly, deficit banks typically settle with surplus banks by transferring reserve balances, which are typically held as accounts at the central bank, which is typically the source of the state money issue. By contrast, the rest of us more typically accumulate our money balances in the form of bank deposits, bank liabilities that are promises to pay state money. And we typically make payments to one another by transferring this private credit money between banks; only the net transfer from one bank to another gets settled in reserve money.

 Because bank money is nothing more than a promise to pay, banks have the power to relax the settlement constraint that faces each of us, simply by swapping IOUs with us. In such an operation, our IOU is the bank’s asset (a bank loan) and the bank’s IOU is our asset (a bank deposit). Central banks similarly have the power to relax the settlement constraint that each bank faces, simply by swapping IOUs with them. In this case, the bank’s IOU is the central bank’s asset (a discount loan) and the central bank’s IOU is the bank’s asset (a reserve balance). In both cases, the settlement constraint is a source of discipline while the swap of IOUs is a source of elasticity in the system, but the elasticity is not automatic or unlimited. Banks (and also central banks) are well-advised to choose carefully when and with whom they swap IOUs, because every additional deposit is a promise to pay reserves on demand. In this way, the banking system’s settlement constraint disciplines its power to relax our own settlement constraint.

 But observe that the banking system’s reserve constraint binds only if surplus banks insist on being paid in reserve money. If surplus banks are instead willing to grant credit to deficit banks, then interbank promises to pay can substitute for scarce reserves, and hence relax the banking system’s reserve constraint. In practice interbank credit takes various forms—correspondent balances, Federal Funds borrowing, and even Eurodollars or RP (repurchase agreements)—but always the precise form depends on the surplus banks. Whatever is acceptable to the surplus banks can be used by the deficit banks to settle their immediate obligations. Even so, it must be emphasized that interbank credit is just a promise to pay, not final payment, and so must itself be settled at some future time. In effect, interbank credit shifts the banking system’s settlement constraint into the future, relaxing a current constraint at the cost of tightening a future constraint.

 The interbank rate of interest is the cost to a bank of putting off the day of reckoning, and it is therefore the key price that banks have in mind when considering whether to expand their balance sheets in order to help us, their retail clients, to relax our own settlement constraints. The interbank rate is a market price, but central banks can influence that price because reserve balances held at the central bank remain the ultimate means of interbank settlement. In effect, central banks can relax or tighten the banking system’s reserve constraint, so lowering or raising the cost of putting off the day of reckoning. That’s what monetary policy is all about. Central banks change the settlement constraint faced by the banking system, and the banking system changes the settlement constraint faced by the rest of us.

 Notwithstanding all of the above, it is nevertheless certainly possible, as a matter of analytical convenience, to abstract from the details of the monetary transmission mechanism, to imagine that households and businesses use state money directly to make payments to one another, and to treat the money rate of interest as determined directly by the supply and demand for state money. Because such an analytical strategy amounts to abstraction from the system of private credit money, it will make the most sense in cases where the state sector is in fact relatively more important than the private sector.

 In general the monetary system is hybrid, part public and part private, with sophisticated mechanisms, such as the interbank market and the discount window, for maintaining par between state money and bank money. But the Great Depression saw the collapse of private banking systems, and subsequent World War saw continued subordination of private banking to the exigencies of state. In crisis times the state sector rose to dominance, and analytical parsimony led to neglect of the mechanisms that manage hybridity. This explains why we find Keynes adopting just such an analytical strategy in his Depression-era General Theory (1936), and the American Keynesians following suit in subsequent wartime and immediate postwar.

 Thus it was that the particular concept of “monetary liquidity” got stabilized in postwar economic discourse. “Liquidity preference” got conflated with money demand, and so (presumably) became amenable to the standard economist’s toolkit. And money demand—Keynes distinguished transactions demand, precautionary demand, and speculative demand—got conceptualized as a matter of portfolio allocation in a simple world where bonds and money are the only assets. Subsequently, as private banking recovered along with private capital markets, the portfolio allocation decision was simply expanded to include these additional alternative assets. The result was the postwar monetary orthodoxy of Patinkin, Modigliani, and especially Tobin (1969).

 But the recovery of private banking and private capital markets brought more than a widened scope of assets in the portfolio allocation problem. Just as in crisis times the state sector rises to dominance, in peace times the private sector rises to dominance, and when that happens the abstraction from private credit money is no longer analytically defensible. Indeed, in the brave new world of private credit money, analytical parsimony suggests just the opposite strategy, that we might want to put private credit money at the very center of the analysis.

 It is not just a matter of paying due attention to interbank markets and the discount window, the mechanisms that ordinarily manage the hybridity between state and private money. Today, wholesale money markets are the marginal source of capital market funding in the new market-based credit system (Mehrling, Pozsar, Sweeney and Neilson 2014). And central bank backstops operate in a globally integrated system that is coordinated by arbitrage across domestic money markets (Mehrling 2016). Here, in the far-reaching institutional change of financial globalization, we find the fundamental reason and urgency for re-theorizing liquidity.

 From a Flow of Funds accounting perspective, dishoarding of money balances is only one way that deficit agents can fill the gap between cash inflow (source of funds) and cash outflow (use of funds). The other two ways are borrowing or selling a financial asset. In finance, the emphasis is entirely on these other two ways, which are called “funding liquidity” and “market liquidity” respectively (Brunnermeier and Pedersen 2009). Funding liquidity refers to the ability to raise cash by borrowing against future income, while market liquidity refers to the ability to raise cash by selling holdings of financial assets. For our purposes, it is necessary to go behind this logical distinction to the institutions that supply these different kinds of liquidity, namely banks and dealers, respectively.

 Banks supply funding liquidity by swapping their own liabilities (deposit accounts) for the liabilities of borrowers. In doing so, they expand their balance sheets on both sides and expose themselves to liquidity risk, because the borrower is free to spend his deposit account. Spending a deposit means transferring it to someone else, possibly someone with an account at another bank, so causing reserve outflows that must be replenished. For this purpose, banks rely on the wholesale money market, and especially the interbank Fed Funds market. Banks supply funding liquidity to their retail clients, expecting to be able to meet their own funding liquidity needs in the interbank market.

 Dealers supply market liquidity by standing ready to buy a specific financial asset at a posted (bid) price, adding the asset to their own inventory if need be. In doing so, they expose themselves to liquidity risk since additions to inventory must be funded, and that funding must be sustained until a buyer for the inventory shows up. Dealers rely for funding on the wholesale money market, especially the repo market where they can use their inventory of securities as collateral for short term borrowing. Not only dealers but also asset managers use the repo market as a source of funding. Not only banks but also corporate cash pools use the repo market as a short term store of monetary liquidity.

 Observe that, in providing the liquidity demanded (and supplied) by their customers, banks and dealers do not simply connect deficit agents with surplus agents; they are not mere brokers. Whether the deficit agent raises the needed cash by borrowing or by selling, what the surplus agent gets is cash, actual money not a promise to pay money. To be sure, that money is credit money not reserve money, new money created for the purpose by the expansion of bank/dealer balance sheets on both sides. In effect, the bank/dealer takes the deficit agent’s settlement problem onto its own balance sheet, promising to pay what the deficit agent cannot. The magic comes from the fact that the bank/dealer promise to pay is money, whereas the deficit agent’s promise to pay is not.

 But there is no free lunch, since banks and dealers face their own settlement constraint. In effect, banks and dealers make money by exposing themselves to liquidity risk, for a price. Banks charge higher interest on their loans than they offer on their deposits, and dealers pay less for financial assets than they expect to be able to sell them for at some later date. Expected profit is the compensation for bearing liquidity risk, but the profit is only expected not guaranteed. In crisis times, banks and dealers face their own settlement constraint, and de facto if not de jure their survival depends on the ability and willingness of the central bank to supply funding and market liquidity to them. De facto if not de jure, the central bank is both lender of last resort and dealer of last resort.

**Funding liquidity and banks**

 Suppose that 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, so we can’t use barter to make our exchange.

 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. This is the classic economists’ story about how money overcomes the improbability of a “double coincidence of wants”. See Figure 1 and Table 1.

Figure 1: Moneyflows Using Money



Table 1: Moneyflows Using Money

 Me You

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Assets | Liabilities |  | Assets | Liabilities |
| -ΔM+goods |  |  | +ΔM-goods |  |

 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, or by giving me back one of my own, so there is only net indebtedness between us. Note that the promise to pay is never actually settled, 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. See Figure 2 and Table 2.

Figure 2: Moneyflows Using Credit

 

Table 2: Moneyflows Using Credit

 Me You

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Assets | Liabilities |  | Assets | Liabilities |
| +goods | +ΔIOU  |  | +ΔIOU-goods |  |

 Now consider yet a third way we might organize our interaction. Suppose neither of us trusts the other 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 in Table 3.

Table 3: Moneyflows Using Credit Money

 Me Bank You

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Assets | Liabilities | Assets | Liabilities | Assets | Liabilities |
| +goods | +ΔIOU | +ΔIOU | +ΔM | +ΔM-goods |  |

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. In an economy with more than two traders, the cross-sectional pattern of trade similarly would cause bank credit to expand (if net debtors buy from net creditors) or contract (if net creditors buy from net debtors). The use of bank credit as a means of payment thus necessarily involves a certain elasticity in the quantity of money.

 In this sequence of examples, we have been thinking about retail trade and exploring alternative mechanisms of liquidity provision—from money, to bi-lateral credit, to bank credit. Much the same logic however applies higher in the system. Suppose there is not one bank but rather many banks so that the problem of me paying you becomes a problem of my bank paying your bank. By analogy to the above discussion, it is clear that such interbank payment could be arranged as transfer of reserves (money), correspondent balances (bi-lateral credit), interbank wholesale money market (multi-lateral credit), or a banker’s bank (lender of last resort). Similarly, at the level above that, suppose there is not one central bank but many, one for each country. Then by analogy, payment could be arranged as a transfer of reserves (gold), short-term capital flows, or some kind of international bank money.

 In both of these extensions to the case of decentralized banking, individual banks face their own settlement constraint, and interbank interest rates put a price on that constraint. This price gets passed down to retail depositors and borrowers, with a spread that compensates for bearing liquidity risk. A central bank that pays interest on reserves and charges interest for discount loans in effect puts bounds on the interbank rate of interest, and central bank operations in interbank markets push market interest rates around within those bounds.

**Market Liquidity and Dealers**

 A liquid market is "one in which an individual transaction does not disrupt the continuity of the market".[[2]](#footnote-2) 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 choose 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; they take market liquidity for granted.

 What is the institutional basis of a liquid market? Consider an example (inspired by Hicks 1989). 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 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.

 A security dealer is somewhat like Westside Market, more so than might first appear. A security dealer who acts as an agent in the issue of corporate securities buys securities from a corporate client 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 buy retail as well as to sell retail. Such 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.)

 In the securities market both bid and ask prices are retail prices, and wholesale (interdealer) prices lie 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. But the two sided dealer also buys, and so it needs also an inventory of cash. Conceptually, we can think 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”. The consequence of these inventories is a liquid market for retail traders. Our hypothetical dealer uses inventories of cash and securities to absorb fluctuations in demand and supply; any increase in one inventory is matched by a decrease in the other, leaving total inventories unchanged. The result is that imbalances in the time pattern of demand and supply show up as balance sheet changes, not price changes. See Figure 3.

Figure 3: Dealers Absorb Order Flow Imbalances



 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, and/or force ultimate suppliers and demanders to wait for the opposite side of the trade. The security market would be like the housing market in which individual buyers have to find individual sellers and negotiate individual deals. Instead, in dealer markets, 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). In this way dealers operate to stabilize price and to facilitate transactions when supply and demand are not balanced.

 Dealers supply market liquidity by offering options to trade. If you want to trade, the dealer is willing to trade at the prices he quotes, either way, but if you don’t want to trade that’s okay too. It’s up to you, not the dealer. This way of conceptualizing the dealer reveals that anyone who offers an option to trade, for example 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.) Our interest is in the professional dealer complex that supplies market liquidity in the most important financial markets.

 Actual dealers differ from my hypothetical dealer in one important respect. They actually hold almost no inventories of either cash or securities. Instead they borrow cash by using securities as collateral, and lend cash by taking in securities as collateral. Typically real world dealers wind up net long, but they move as close as they can to the ideal balance sheet shown in Table 4. The key point is that market-making activity is not limited by the dealer’s own holding of inventories, but rather by the sum total of inventories held anywhere in the economy, and by the ability of dealers to access those inventories as needed.

Table 4: Ideal Dealer Balance Sheet

|  |  |
| --- | --- |
| Assets | Liabilities |
| Reverse | Repo  |

 In this table, we can think of the dealer’s repo position as borrowing money (and lending collateral) to finance the dealer’s long security positions, and the dealer’s reverse position as lending money (borrowing collateral) to finance the dealer’s short positions. Instead of holding inventories, dealers cultivate access to inventories of cash and securities that reside on other people’s balance sheets, and use the repo and reverse markets to tap these outside inventories as needed. The dealer’s clearing bank is one such source of cash; the dealer’s asset manager clients are one such source of securities. In effect, dealers operate a just-in-time inventory system, which allows them to operate as though they do have inventories even though the inventories are actually out in the market some place.

 In actual practice dealers tend to be net long securities and net short cash as we can see in Tables 5 and 6. Table 6 shows the net outright positions in a range of securities. Negative numbers are short positions, but in general the numbers are positive and the total inventory of securities held is $222 billion. But that net number is an order of magnitude smaller than the overall size of the dealer balance sheet as shown in Table 5. Primary dealers are borrowing $2,552 billion (lending securities) and lending $2,370 billion (borrowing securities), ten times larger than the net, in effect moving inventories of cash and securities from one balance sheet to another, using their own balance sheet as the intermediary. When deficit agents raise cash by selling an asset, it is the dealer system that provides that cash, not from its own inventory of cash but rather by tapping sources of cash elsewhere in the economy. This is modern liquidity.

Table 5: Financing by Primary Dealers, January 6, 2016, (in millions of dollars)

|  |  |
| --- | --- |
| Securities In | Securities Out |
| Reverse Repurchase Agreements | 1,755,221 | Repurchase Agreements | 2,092,309 |
| Other Securities Borrowed | 615,020 | Other Securities Lent | 459,953 |

Source: New York Federal Reserve Bank, author calculations

Table 6: Primary Dealer Positions, January 6, 2016 (in millions)

|  |  |
| --- | --- |
| Type of Security | Net Outright Position |
| U.S. Government |  |  | 74,511 |
|  | Treasury Bills | 18,723 |  |
|  | Treasury Bonds, <2 yrs | 19,207 |  |
|  |  2 to 3 years | -6,676 |  |
|  |  3 to 6 years | 11,880 |  |
|  | 6 to 7 years | 13,503 |  |
|  | 7 to 11 years | -3,164 |  |
|  | > 11 years | 11,554 |  |
|  | Floating Rate Notes | 1,911 |  |
|  | Treasury Inflation Protected Securities (TIPS) | 7,573 |  |
| Federal Agency and GSE |  |  | 15,012 |
|  | Discount Notes | 3,770 |  |
|  | Coupon Securities | 11,242 |  |
| Mortgage-Backed Securities |  |  | 88,880 |
|  | Federal Agency and GSE | 68,448 |  |
|  | Non-Agency | 20,432 |  |
| Corporate Securities |  |  | 17,227 |
|  | Commercial Paper | 7,382 |  |
|  | Investment Grade Securities | 7,731 |  |
|  | Below Investment Grade Securities | 2,114 |  |
| State and Municipal |  |  | 14,215 |
| Asset-backed Securities |  |  | 12,271 |
|  | Credit card-backed | 537 |  |
|  | Student loan-backed | 3701 |  |
|  | Automobile loan-backed | 1080 |  |
|  | Other | 6,953 |  |
| TOTAL |  |  | 222,116 |

**The Inherent Instability of Credit**

 The discussion above provides the institutional and analytical foundation for re-theorizing liquidity for the modern world. Contra Keynes, and also contra the postwar orthodoxy that built on him, it is an inside credit story, not an outside money story. It is a story about match and mismatch between the cross-time and cross-section pattern of income and expenditure for each agent, not a story about money demand. It is a story mainly about private financial intermediaries—banks and dealers—not the Treasury and the central bank.

 That said, the story remains very much within the broader line of monetary thought to which Keynes belongs, a line that stretches back before Keynes to Bagehot and Hawtrey, and after Keynes to Minsky and Kindleberger. In this line of thinking, credit is seen as an inherently destabilizing element in our modern economy, one that has both good and bad aspects. By allowing people to spend today in anticipation of income yet to come in the future, we relax the survival constraint that would otherwise hold back economic growth. But because income is uncertain, the ability to repay is uncertain as well. More generally, credit extension always involves a speculative element—that is the whole point of it—but no speculative tree ever grows to the sky, and no one knows where the boom will stop until it stops.

 Today, as in the past, the central task of the regulatory authorities is to manage the balance between discipline and elasticity in the system as a whole. Too much discipline is bad, but so is too much elasticity. In the past, management focused on price of money, meaning the policy rate of interest, as the key point of leverage. In the future, as we take on board the implications of the shift to a market-based credit system, management will likely focus additionally on the price of risk, and specifically tail risk in three dimensions: duration risk, credit risk, and foreign exchange risk. De facto if not de jure the modern central bank serves as both lender of last resort and dealer of last resort (Mehrling 2015).

**References**

Brunnermeier, Markus and Lasse Pedersen. 2009. “Market Liquidity and Funding Liquidity.” Review of Financial Studies 22 (6): 2201-2238.

Harris, Larry. 2003. Trading and Exchanges: Market Microstructure for Practitioners. New York: Oxford University Press.

Hicks, John. 1989. A Market Theory of Money. New York: Oxford University Press.

Keynes, John Maynard. 1936. The General Theory of Employment, Interest, and Money. London, Macmillan.

Mehrling, Perry. 2011. The New Lombard Street, How the Fed Became the Dealer of Last Resort. Princeton NJ: Princeton University Press.

Mehrling, Perry. 2014. “Why Central Banking should be Re-imagined.” Pages 108-118 in Re-thinking the lender of last resort. BIS Paper #79. Monetary and Economic Department, Bank for International Settlements.

Mehrling, Perry. 2016. “Elasticity and Discipline in the Global Swap Newwork.” International Journal of Political Economy.

Mehrling, Perry, Zoltan Pozsar, James Sweeney, Daniel Neilson. 2014. “Bagehot was a Shadow Banker: Shadow Banking, Central Banking, and the Future of Global Finance”. In Shadow Banking Within and Across Borders, edited by Stijn Claessens, Douglas Evanoff, George Kaufman, and Luc Laeven. World Scientific Publishing.

Tobin, James. 1969. “A General Equilibrium Approach to Monetary Theory.” Journal of Money, Credit, and Banking 1 No. 1: 15-29.

1. This paper draws heavily on material developed for my class “Economics of Money and Banking” but has been developed further in interaction with members of the IGLP Workshop, initially in June 2011 but continuing on after in multiple locations. [↑](#footnote-ref-1)
2. This way of conceptualizing liquidity is developed in more detail in Larry Harris, Trading and Exchanges: Market Microstructure for Practitioners (2003). [↑](#footnote-ref-2)