Shadow Banking:
Why Modern Money Markets are Less Stable Than 19th c. Money Markets But Shouldn't Be Stabilized by a "Dealer of Last Resort"

Carolyn Sissoko*

First Draft: March 14, 2014
This Draft: September 11, 2014

Abstract:

An important policy question that is currently being discussed by central bankers and academics is whether the "shadow" banking system should have a permanent backstop from the central bank akin to the extraordinary support that was provided by the Federal Reserve to the shadow banking system in 2008. I answer this question by (i) using a macroeconomic analysis of banking (ii) to explain the role played by the first lender of last resort in 19th c. Britain and (iii) applying this analysis to the modern shadow banking system. I conclude that because of its heavy reliance on collateralization, the shadow banking system is a poor substitute for the traditional banking system and does not merit the support of a "dealer" of last resort.

My macroeconomic framework explaining the role played by banks in economic growth builds on Cavalcanti and Wallace (1999) and was developed in Sissoko (2006) and Sissoko (2007). This framework explains that in an environment with liquidity constraints banks are special because their history of default is public, so unsecured bank borrowing is incentive compatible when bankers profit from their special characteristic by underwriting the debt of the non-banks in the economy. This debt circulates as bank liabilities and can resolve the liquidity constraints faced by the economy entirely, facilitating economic growth (and generating fees for bankers). Because being liquidity constrained is a significant penalty, a simple trigger strategy played by banks against defaulting non-banks can support an environment where borrowers limit their own debt, credit is generally available, and the economy grows to its maximum potential. The distinctions between this model and the most important competing model of banks as issuers of "information insensitive" assets in Gorton and Ordonez (2014) are first that safe assets take the form of unsecured debt, and second, that they are actually safe, because borrowers themselves do not want to borrow more than they can pay. Crises take place in my environment when there is a loss of confidence in the credit system.

To explain the role of the lender of last resort, I turn to the actual historical environment on which my model of banking is based, 18th – 19th c. Britain. In Britain the privately-issued bills that circulated in the money market were so safe that collateralizing them was viewed as clearly superfluous. Several institutions made these privately-issued bills safe: they were guaranteed by at least two parties, the issuer, the bank-acceptor, and anyone who chose to sell the debt, bank owners faced capital calls for bad debts the bank had guaranteed, the bills were short-term, and the credit system itself was protected by the

* I thank Richard Sylla for his comments. All errors are, of course, my own. This paper was completed while I was a fellow at the USC Center for Law and Social Science.
Author email: sissoko@usc.edu.

existence of a lender of last resort. The lender of last resort made the provision of banking services incentive compatible for bankers who effectively faced unlimited liability by providing high-powered money in liquidity crises and thereby averting the threat of a transition to a no-credit equilibrium. The lender of last resort played a second very important role: in order to maintain the "safe" quality of the money supply the central bank withdrew support from any bank that it believed to be overissuing debt. The words "last resort" themselves refer to this management of moral hazard by the central bank.

Shadow banking developed in an environment where the second duty of the lender of last resort had been forgotten, and solvency for large banks was effectively redefined to incorporate public sector support. Shadow banking has been described as market-based short-term finance of long-term lending, but in fact refers to a bank-guaranteed system of finance based on commercial paper and repurchase agreements, that provides little funding for private sector assets and significant funding for investment banks. In other words, to the degree that shadow banking has disintermediated commercial banks, it has done so by reintermediating investment banks – using repurchase agreements, which are a form of funding that is even more unstable than deposits, due to their reliance on collateral that is remargined daily. Furthermore, the growth of the collateralized money market that shadow banking represents is probably destabilizing the incentive structure that is the product of centuries of institutional evolution and that undergirds the traditional unsecured money markets upon which the past 250 years of economic growth have been founded.

Although it is often claimed that the purpose of a "dealer of last resort" is to support liquidity on asset markets, in practice, the policy is designed to give the largest dealer banks access to central bank credit and to support them though a crisis – as we saw in 2008. I explain that this policy only protects asset markets from fire sales of assets belonging to the select group that has access to central bank lending, not from fire sales in general, because dealer banks do not extend credit in the same way that commercial banks do. Finally, commercial banks traditionally bear risk for the economy, whereas dealer banks traditionally avoid bearing risk over time themselves but instead facilitate the allocation of that risk to others. As a result, dealer banks should not receive support similar to that of commercial banks, because they do not play the same role in the economy that commercial banks do.

Table of Contents:

An important policy question that is currently being discussed by central bankers and academics is whether the "shadow" banking system – which is currently little more than a funding system for investment banks – should have a permanent backstop from the central bank akin to the extraordinary support that was provided by the Federal Reserve to the shadow banking system in 2008. This paper finds that such a backstop would be a mistake, because it would further entrench the collateralized money market that developed over the past decade or two when extraordinary protection in bankruptcy for certain forms of collateral was enacted,[1] that played an important role in the failures of 2008, and that is likely to continue to undermine stability in the financial system.

This paper argues that policy decisions about the boundaries of lender of last resort activity should be based on a macroeconomic understanding of the role played by the banking system in the economy and of the role of the central bank. This paper takes a historical approach to the questions of What do banks do? and What does a lender of last resort do? Monetary theorists who wrote at the birth of modern banking believed that the economic growth that was taking place around them was founded on the "science" of credit that allowed banks to make unsecured credit widely available across the economy, on the banks as monitors of that credit, and on the central bank as both monitor in chief and a source of liquidity when there was a crisis of confidence in the credit system. (Formal macroeconomic analysis of this credit system is presented in other work.[2]) In short, this paper argues that unsecured credit is the instrument that the banking system uses to make the phenomenon of modern economic growth possible. Thus, the collateralized money market must be evaluated in the context of its effect on the stability of the unsecured bank-based credit system that has been the foundation of modern economic growth for almost two centuries. This paper argues that the collateralized money market is probably destabilizing the traditional banking system.

Despite significant reforms to the banking system over the past five years, central bankers recognize that the shadow banking system still plays an important role in the economy and that the fundamental question of how to stabilize this system has yet to be answered. William Dudley has found that there are two broad options: the first is to reduce the shadow banking system's importance by curtailing its use of short-term money market finance, and the second is to provide a central bank backstop to the shadow banking system by, for example, giving dealer banks access to lender of last resort facilities.[3] Many other central bankers have raised the possibility that the central bank will in the future be the "market maker of last

[1] Carolyn Sissoko, The Legal Foundations of Financial Collapse, 2 J. Fin. Econ. Pol'y 5 (2010); Enrico Perotti, The Roots of Shadow Banking, CEPR Policy Insight No. 69 (Dec. 2013). See also William Dudley, Fixing Wholesale Funding to Build a More Stable Financial System, Speech at the New York Bankers Association, Feb. 1, 2013.

[2] Carolyn Sissoko, An Idealized View of Financial Intermediation, Economics: The Open-Access, Open-Assessment E-Journal, 2007-5 (2007). Carolyn Sissoko, Short-term credit: A Monetary Channel Linking Finance to Growth (2006). This paper uses formal economic modeling to find that in a world where liquidity constraints are ubiquitous and constrain economic performance, banks can make economic growth possible by issuing and monitoring the use of overdraft accounts. The mechanism by which banks make credit generally available is, first, by being creditworthy and able to borrow unsecured themselves, and, second, by issuing and monitoring individual accounts. In short, by making access to credit an economy-wide norm banks overcome the liquidity constraints that are inherent in an unbanked economy and make possible the phenomenon of modern economic growth with which we are all familiar.

[3] William Dudley, Speech at the New York Bankers Association, Feb. 1, 2013.

resort," providing liquidity to financial markets as well as banks. Usually these discussions present some costs and benefits of such a policy without coming out firmly in favor of or against the proposal.[4]

One central banker, Mark Carney, Governor of the Bank of England, has, however, come out strongly in favor of expanding the concept of a lender of last resort to a more general backstop of "private markets" in order to "catalyse more efficient and effective private collateral management."[5] William Dudley finds that the decision of the central bank to support the shadow banking system will "depend in large part on the social value of the capital markets-based activities presently being financed in unstable short-term wholesale markets and the utility of short-term wholesale funding for lenders." It is remarkable that neither of these central bankers sees the need for a macroeconomic understanding of the role played by the collateralized interbank lending in modern banking as compared to the unsecured interbank lending of the previous two centuries of banking history. Instead, both central bankers appear to believe that microeconomic analysis will be sufficient to determine the appropriate scope of collateralized lending and of the central bank's support for such lending.

This paper argues that in order to evaluate the scope of the lender of last resort's role in modern collateralized markets a macroeconomic framework is needed. Whether or not private markets should be backstopped by the central bank is a question that cannot be answered by focusing on the microeconomic efficiencies of the specific activities currently being financed or of the current value to lenders of such funding. Instead, it is necessary to take a macroeconomic view of the banking system, of its role in the economy, and its relationship to the central bank. Only if the backstop of private markets makes sense in the context of a coherent and credible model of the banking system should it be entertained as a policy option.

Because of the relationship established in the theory between unsecured credit and economic growth, I argue that in order for the banking system to function it must be able to borrow on an unsecured basis and that modern regulators err when they promote the collateralization of money market debt. Instead modern regulators should promote an institutional structure where banks borrow on an unsecured basis and where market forces will weed out banks that are not creditworthy enough to borrow unsecured. In short, this paper argues in favor of the first option proposed by William Dudley, the size of the shadow banking system should be reduced by curtailing the use of collateralized short-term money market instruments to finance the banking system.

A secondary goal of this paper is to interpret the history of banking, and more particularly to demonstrate how utterly unfounded are many commonplace assertions about that history. Financial institutional development matured in the 19th c. and I argue that it set off a wave of economic growth in what are now known as the "developed" countries that has continued more or less up through the present day. The foundational institutional structure that made it possible for 19th c. economies to overcome the financial constraints that had restrained economic growth for millennia provides very important lessons about economic incentives and the means by which institutions can be used to align them. The lessons of 19th c. finance are important, in my view, first, because we in the 21st century have inherited an institutional structure the key components of which we have yet to study and understand, and second, because changes

[4] See, e.g, Dudley; Paul Tucker, Regulatory Reform, Stability, and Central Banking, Hutchins Center on Fiscal and Monetary Policy (2014).

[5] Mark Carney, The UK at the heart of a renewed globalization, Speech, Oct. 24, 2013.

that have been made to this structure in recent decades may well be undermining its integrity. If I am correct, then structural damage to the financial system, if it continues, has the potential to lead to a complete collapse of the financial system – and a return to the growth rates that were experienced before the modern era.

I find that in 19th c. Britain, the unsecured money market instruments that circulated were of extremely high quality and there was virtually no risk of creditor losses. How was it possible to establish such a safe system of unsecured credit? First, the instruments that circulated bore at least two promises to pay, that of the issuer and that of the acceptor, and in order to sell the instrument the seller had to promise to pay it too. This legal structure had the effect of aligning the incentives of everyone who circulated the debt. Second, bank owners, even joint share owners, were subject to capital calls in the event that creditors were losing money in the liquidation of a bank. Once again the legal structure aligned the incentives of the market participants. Third, the assets that were financed were short-term, limiting their price volatility and exposure to unexpected events. Finally, whenever there was a crisis of confidence in the system, the Bank of England as lender of last resort stood ready to expand the money supply and prevent the panic and general withdrawal of access to credit from causing needless defaults. Because the bankers had to guarantee all the debt they circulated, the lender of last resort protected not only the economy, but also the bankers from having to liquidate their personal assets in order to make good on all the defaults that would have taken place in a general collapse of credit. In short, the lender of last resort made the continued offering of banking services incentive compatible despite the bankers' liability for bad debts.

One concept underlying this analysis is a recurring theme in the banking literature. Theoretic models of credit typically have an equilibrium with credit and an equilibrium without credit – where people have no trust in the credit system. I argue that the role played by a lender of last resort is to alleviate growing distrust in the credit system and make it clear that there will be no transition to a no-credit equilibrium over the near term.

It is remarkable that none of these lessons about the incentives faced by bankers that can be drawn from the 19th c. British experience are recognized in the current literature on banking. On the contrary, the process of securitization employed a complex sequence of transactions in order circumvent the traditional bank guarantee on the assets sold – only to meet with the surprised discovery that the traditional rule was designed to align incentives. Instead of recognizing that the foundations of modern banking were laid in an environment where bankers were liable for the debts they circulated and where the existence of a lender of last resort was necessary to reduce the risks faced by bankers sufficiently that banking services could be provided throughout the economy, it is commonly argued that banks are "always" bailed out in crises.[6] Instead of analyzing why the finance of long-term assets on money markets creates extraordinary risks for the economy, the modern literature assumes that this is what a money market does. In short, until we make the effort to understand the value of the institutions that we have inherited, we are in danger of destroying them out of ignorance.

[6] See, e.g., Michael Pettis, The real cost of Chinese NPLs, Jan. 21 2011, available at https://www.creditwritedowns.com/2011/01/the-real-cost-of-chinese-npls.html ("Throughout modern history . . . there has really only been one meaningful way to resolve banking crises . . . wealth is transferred in sufficient amount from households to borrowers or banks."). Interestingly, Reinhart and Rogoff, even as the argue that bailout costs are too difficult to measure accurately, also appear to embrace the view that banking crises inevitably entail bailouts. Carmen Reinhart & Kenneth Rogoff, This Time is Different 163-64 (2009).

Part I introduces the shadow banking system, and finds that in its current form its primary purpose is to provide an unstable source of funding for investment banks. Part II asks the question: What do banks do? and uses historical analysis and the work of early monetary theorists to find that by underwriting the issuance of unsecured debt by members of the broader economy banks support economic growth. Part II also observes that banking grew up in an environment where the private sector debt which backed the money supply was virtually risk-free, and thus that "bailouts" are not an unavoidable consequence of maintaining a banking system. Part III presents the elements of a model that would be consistent with the 19th c. British experience and compares these elements to the Gorton and Ordonez model of "information insensitive" debt. Part IV asks: What does a lender of last resort do? and again using a historical approach finds that letting elements of the banking system fail was equally as important to the last resort lender's role as supporting the banking system through a crisis. Part V argues that the movement to collateralized money markets is destabilizing the unsecured foundations on which the banking system is built, and that extending the scope of the lender of last resort to support private markets would work to entrench the forces that are destabilizing the financial system. Part VI concludes.

I. The Shadow Banking System is an Unstable Funding System for Banks, Not Assets

There are many definitions of shadow banking. A New York Federal Reserve Bank monograph effectively equates shadow banking to securitization, or the process by which individual loans are packaged into bundles, used to issue a wide variety of collateralized assets, and sold to investors. The New York Fed monograph is often used to demonstrate how complicated and virtually incomprehensible the shadow banking system is – it includes a "map" of the shadow banking system that, for legibility, the authors recommend printing as a 36" by 48" poster.[7]

More commonly, however, the term shadow banking refers to the use of money market instruments to provide short-term finance to long-term assets,[8] and thus focuses attention on bank runs and on the fact that shadow banks can face such runs, just as traditional banks do. For this reason securitization should not be equated with shadow banking, because a significant portion of private sector securitized assets were financed on a long-term rather than on a short-term basis.[9] This paper will limit its focus – as does most of the literature on shadow banking – to the role played by money markets in the finance of long-term assets.

This paper finds that our current money markets play only a very small role in the direct finance of private sector long-term assets and for the most part are used as a financing system for investment banks. In short, the "market-based" credit system that some equate with the shadow banking system,[10] is very small – and relies heavily on commercial bank guarantees. To the degree that a substantial shadow banking system continues to exist, it does not fund long-term assets directly, but instead provides wholesale funding for investment banks, and to a lesser degree commercial banks.

---

[7] Zoltan Pozsar, Tobias Adrian, Adam Ashcraft, & Hayley Boesky, Author's Note in Shadow Banking, NYFRB Staff Rep. No. 458 (July 2010).

[8] Perry Mehrling, Zoltan Pozsar, James Sweeney, and Daniel Neilson, Bagehot was a Shadow Banker (Nov. 2013).

[9] For example, although only $35 billion of private label residential mortgage-backed securities have been issued since 2008, at the end of 2013 more than $1 trillion of such securities remained outstanding. Data from SIFMA: http://www.sifma.org/uploadedFiles/Research/Statistics/StatisticsFiles/SF-US-Mortgage-Related-SIFMA.xls?n=47986.

[10] Perry Mehrling, Zoltan Pozsar, James Sweeney, and Daniel Neilson, Bagehot was a Shadow Banker 2 (Nov. 2013).

To be clear, the focus here is on finance of private sector banks and assets. Thus, although Fannie Mae and Freddie Mac played a very important historical role in the development of the shadow banking system by pioneering the practice of financing long-term mortgage debt on money markets through the issue and roll over of short-term debt that was at least nominally a private-sector obligation,[11] they now officially have government support, and, for the purposes of this paper their debt is treated not as part of the shadow banking system, but as a government obligation.

This paper provides a simple framework for understanding the shadow banking system that is organized around the two instruments, commercial paper and repurchase agreements, that play an important role in money markets and that are, very roughly, comparable to deposits. Studying how these instruments are used not only allows a distinction to be drawn between the direct finance of assets and the finance of assets that sit on bank balance sheets, but also makes clear why the shadow banking system is unstable.

This analysis finds that the money market instruments have in the past played three roles: they have funded banks and non-financial firms directly, they have funded assets that lie off bank balance sheets, and in order to play these roles, they have created a need for commercial bank guarantees that induce lenders to lend off-balance-sheet or in the case of tri-party repo to investment banks. In practice, the direct funding of assets now takes place only on a very small scale.

Because the two money market instruments, commercial paper and repurchase agreements (repos), are both short-term, it is easy for those who invest in them to "run," or to decide that they no longer wish to invest their funds with a specific issuer or, indeed, in privately issued money market assets at all. Because these investors can always choose to put their money in Treasury bills or bank deposits, runs in the money markets are associated with unmanageably sudden shifts in investor preferences across short-term assets. In short, a fundamental attribute of the shadow banking system is that the decisions of money market investors can destabilize the money markets.

Money market mutual funds and enhanced cash funds (that promise liquidity, but are less regulated than money market funds) are the most obvious money market investors, but the buy-side of the money market is composed of a huge array of institutional investment funds, corporations, and government bodies that have funds they wish to keep in liquid form. All of these entities can be part of a run in the shadow banking system. In addition, as will be explained in detail below, in the repo market it is possible for the recipients of funds, such as prime brokerage clients and banks in the interdealer market, to run.

Now that the basic instability of the money markets has been established, the next step in understanding the shadow banking system is to understand the different ways in which commercial paper and repo-based instruments are used; this is discussed in sub-part A. The following sub-parts evaluate what shadow banking does, discuss why it is more unstable than traditional banking, address why it is not a market-based credit system, and, finally, finds that the key shadow banking question that regulators must address is: Should lending on collateralized wholesale funding markets be curtailed?

[11] See Frank Fabozzi & Michael Fleming, U.S. Treasury and Agency Securities 11 (April 2004), available at http://www.newyorkfed.org/cfcbsweb/Treasuries_and_agencies.pdf.

A. Shadow Banking Instruments

1. Commercial Paper

a. Unsecured

Commercial paper is traditionally an unsecured obligation to make a payment that has a maturity of one year or less. It is analogous to the commercial bills that were used to finance economic activity in 19th c. Britain, and indeed has existed in one form or another for centuries.

i. Issued by financial institutions

A little over half the commercial paper issued in the United States, or approximately $550 billion, is issued directly by financial institutions.[12] Because this market-based funding source is much less stable as a funding source than retail deposits, it is categorized along with other bank funding sources that are prone to runs as wholesale funding. The case of Lehman Bros. illustrates the instability of this form of funding. When Lehman declared bankruptcy, its commercial paper went into default, and set off a run by investors who feared money market mutual fund losses on money funds that invested in commercial paper; as a result the commercial paper market itself faced a run.

ii. Issued by non-financial corporations

Approximately one quarter of commercial paper is unsecured and issued by non-financial corporations. Because non-financial corporations have less access to liquidity than banks, there is a risk that when their commercial paper is due they will be unable to roll it over into a new issue and will be unable to honor their commercial paper obligations due to this liquidity risk. For this reason, almost all non-financial commercial paper is protected by a liquidity facility provided by a bank, which promises to retire the commercial paper if the issuer is unable to do so. Observe that when Lehman failed, the run on commercial paper was not carefully targeted to financial commercial paper, and as a result non-financial commercial paper was subject to a run as well.

b. Collateralized: Asset Backed Commercial Paper

In recent decades, sponsoring banks have moved assets that they originated into financing vehicles that are "bankruptcy-remote," or not available to the sponsor's creditors in the event that the sponsor declares bankruptcy. In addition, in theory any support that would be provided by the sponsor to the vehicle was defined in a contract, so the sponsor had contractually limited exposure to the vehicle's liabilities.[13] Thus, asset-backed commercial paper (ABCP) vehicles were designed as a means of removing assets from the sponsoring bank's balance sheet.

The ABCP market was one of the key markets that collapsed in the early days of the financial crisis – from $1.2 trillion outstanding in early August 2007 to $905 billion three months later. Since then the market has continued to decline slowly, and it now hovers around $250 billion.

---

[12] Federal Reserve Commercial Paper Release, Outstanding

[13] In practice, banks sometimes supported these vehicles even in the absence of a contractual obligation to do so, and sometimes did not.

Because these vehicles finance long-term assets they face the same liquidity risk as non-financial issuers when issuing commercial paper. In addition these vehicles face credit risk in the event that the value of the assets falls below the value of the commercial paper, and the vehicle is no longer fully collateralized. Both liquidity and credit risk must be addressed before the vehicle can receive a credit rating that is high enough for it to issue ABCP. The three principal means by which liquidity and credit risk were resolved are discussed below.

i. Bank supported ABCP: Conduits

Prior to the financial crisis most ABCP was issued by ABCP conduits that were sponsored by banks. The banks typically provided both a liquidity facility, which guaranteed that the commercial paper would be retired even if it could not be rolled over, and a credit facility, which promised to honor some fraction of the commercial paper in the event that the value of the collateral fell too low to cover the costs of repaying the commercial paper.

In August 2007 when the crisis started there was a sudden loss in confidence in the ABCP market and many conduits could not roll over their commercial paper. The banks had to step in and honor the liquidity guarantees that had been made – and in order to do so they had to seek regulatory exemptions that are documented by the Federal Reserve.[14]

ii. Liability structure supported ABCP: SIVs, LPFCs, etc.

Some ABCP-issuing vehicles guaranteed the payment of ABCP by funding the assets with a combination of bonds, medium-term notes and ABCP. These vehicles took many forms; the most common were called structured investment vehicles (SIVs).

The concept behind these vehicles was that, in the event that the commercial paper could not be rolled over or the value of the assets fell below a trigger point, assets would have to be sold to pay off the ABCP and any losses would fall to the longer term debt holders. In 2007 most SIVs hit their triggers and were unwound. Because of the losses that were incurred by both longer-term and commercial paper investors (after lawsuits determined the allocation of proceeds), they are no longer a popular investment product.

iii. Repo Conduits – discussed below

2. Repurchase Agreements

A repurchase agreement (repo) is a simultaneous agreement to sell an asset today and to repurchase it a specific date and time in the future. It has the same economic effect as a collateralized loan. Typically the amount lent is less than the value of the collateral;[15] the percentage difference is called a haircut.

[14] See the letters granting JPMorgan Chase & Co., Citigroup Inc., and Bank of America Corp. Regulation W exemptions that are dated August 20, 2007, available at the Federal Reserve website: http://www.federalreserve.gov/boarddocs/legalint/FederalReserveAct/2007/.

[15] Note that in securities lending, where institutional investors provide high-quality, high-demand collateral like Treasuries to the market, haircuts frequently go in the reverse direction. That is, more money must be lent than the value of the collateral in order to induce the securities lenders to lend.

There are two repo markets: the bilateral repo market and the tri-party repo market. In the bilateral repo market the lender must have the capacity to receive and manage the collateral, whereas in the tri-party repo market the tri-party clearing banks, JP Morgan Chase and Bank of New York Mellon, provide collateral management services for the lenders. Money market investors like mutual funds lend only on the tri-party repo market where the principal borrowers are the dealer banks (although a few hedge funds and private institutions are credit-worthy enough to be accepted as counterparties on this market).[16]

The tri-party repo market since its early days has relied heavily on bank guarantees of liquidity provided by the clearing banks. Up until the end of 2012, there was a daily unwind of every trade in the market so that the broker-dealers borrowing on the market could have access to the collateral they had posted during the day. This process relied on a massive extension of credit by the two tri-party clearing banks during the day until the trades were rewound in the late afternoon.[17] Thus the tri-party clearing banks provided a guarantee to the market and bore the risk of a broker-dealer failure during the day.[18] Tri-party repo market reform has been high on the Federal Reserve's agenda and has finally been realized in the past few months: by October 2013 intraday credit had been reduced to 70% of the market, and by February 2014 to 20%.[19] It is predicted that broker-dealers will receive intraday credit only on a capped, committed basis by the end of 2014.[20] How well this newly restructured market will perform has, however, yet to be established.

On the bilateral market, where the lender must manage the collateral, the dealer banks are the lenders. The borrowers are prime brokerage clients, such as hedge funds, and other dealers.

As a result of this structure, funding generally enters the repo market via tri-party repo and the dealer banks, then, distribute this funding more broadly to their prime brokerage clients on the bilateral repo market. Thus, when a hedge fund buys an asset on margin, it borrows a significant fraction of the purchase price from the dealer bank that is its broker and posts the asset as collateral for the loan in a repo transaction. The dealer bank can then repo the asset on the tri-party repo market so that the dealer bank is effectively intermediating lending from the tri-party market to its client and earning an interest rate spread for the intermediation services. When the asset is of a type that cannot be used as collateral in the tri-party repo market, the dealer may choose to use the asset to raise funds on the inter-dealer segment of the bilateral repo market.

The dealer banks also hold collateral that is posted against derivatives contracts by other dealers and by prime brokerage clients. Whereas the inter-dealer derivatives contracts may have symmetrical collateral posting requirements, prime brokerage clients have typically been required to post collateral without having the right to require that dealer bank follow the same rule when the balance on the derivatives contracts is in the brokerage client's favor. As a result a dealer bank is almost certain to receive collateral

[16] Tobias Adrian, Brian Begalle , Adam Copeland , Antoine Martin, Repo and Securities Lending, Federal Res. Bank of NY Staff Report No. 529, Feb. 2013 at 5-6.
[17] William C. Dudley, speech, Introductory Remarks at Workshop on "Fire Sales" as a Driver of Systemic Risk in Tri-Party Repo and Other Secured Funding Markets, Oct. 4, 2013; FRBNY Press Release, Update on the Tri-Party Repo Reform, Feb. 13, 2014.
[18] Adam Copeland, Darrell Duffie, Antoine Martin, and Susan McLaughlin, Key Mechanics of The U.S. Tri-Party Repo Market, 18 FRBNY Economic Policy Review 17, 22, 24 (2012).
[19] FRBNY Press Release, Update on the Tri-Party Repo Reform, Feb. 13, 2014.
[20] FRBNY Press Release, Update on the Tri-Party Repo Reform, Feb. 13, 2014.

from its prime brokerage services when its client accounts are aggregated. The collateral posted by prime brokerage clients can then be used by the dealer to borrow in the tri-party repo market. As a result of this structure collateral posting by prime brokerage clients on their derivatives liabilities is also a form of financing for the dealer banks.

Thus, dealers often finance their own inventories, their prime brokerage clients' assets, and any collateral that is posted against derivatives liabilities by other dealers or prime brokerage clients on the tri-party repo market.

The repo market is very different from the ABCP market and from commercial paper markets in general, because a run in one of the latter markets can only be caused by end investors. In the repo market a run can be started either by end investors or by other dealers and/or prime brokerage clients. Darrell Duffie has explained the many channels by which funding can be withdrawn in a repo market. These include: brokerage clients can move their accounts – together with all the collateral they have posted – to another dealer; dealers or brokerage clients who are derivatives counterparties can seek a novation (i.e. transfer) of a derivatives contract in order to post collateral to or expect payment from a more creditworthy dealer; dealers or brokerage clients may seek to reduce new exposures by entering into derivatives contracts that will require a dealer to post collateral; or repo lenders may increase haircuts or stop lending entirely to the dealer.[21] In short, the repo market is subject to inter-dealer and brokerage client runs, as well as to runs by repo investors.

In 2008 it is very clear that both Bear Stearns and Lehman faced a withdrawal of funding from other dealers, from brokerage clients, and from end investors in the repo market.[22]

3. Repo Conduits

A repo conduit is a bankruptcy remote financing vehicle. The vehicle issues commercial paper that is backed by a repo with a maturity that matches the commercial paper. Thus, a repo conduit is backed primarily by the credit of the repo counterparty. Only if the repo counterparty fails to pay, can the repo conduit foreclose on the repo collateral. Because the term of the repo matches the term of the commercial paper, rating agencies do not require that a repo conduit have a backup liquidity facility.

The credit rating of a repo conduit typically is based entirely on the credit of the repo counterparty.[23] For this reason, repo conduits can be used – by institutions with high credit ratings – to finance assets that would not be eligible for tri-party repo financing.

---

[21] Darrell Duffie, How Big Banks Fail 23 – 42 (2011). *See also* William Dudley, More Lessons From the Crisis, Remarks at the Ctr. for Econ. Policy Studies Symposium, (Nov. 13, 2009), *available at* http://www.newyorkfed.org/newsevents/speeches/2009/dud091113.html; Adam Copeland, Antoine Martin & Michael Walker, *The Tri-Party Repo Market before the 2010 Reforms* 56-58 (Fed. Res. Bank of N.Y. Staff Rep. No. 477, 2010).
Duffie observes that when there is a repo market run, the *coup de grace* is almost always given by a clearing bank when it responds to concerns about a firm's financial position by exercising its right to offset aggressively, by for example demanding collateral for intraday exposures or refusing to give access to deposits. Duffie, supra note 9, at 41¬42. See also Tobias Adrian & Adam Ashcraft, Shadow Banking Regulation 17 (Fed. Res. Bank of N.Y. Staff Report No. 559, 2012).

[22] Duffie, at 23-42.

[23] Moody's Revises Approach To Counterparty Rating Actions In Repo ABCP Conduits, Oct. 21, 2009, available at http://www.cranedata.com/archives/all-articles/2541/

B. What Does Shadow Banking Do?

1. Shadow Banking is a Funding Mechanism for Banks

The most important role of the shadow banking system is to provide wholesale funding for banks. Unsecured wholesale funding is provided when a bank issues commercial paper. Secured wholesale funding is provided when an investment bank uses the tri-party repo market to finance inventories, the assets of brokerage clients, and any collateral posted by counterparties in derivatives transactions.

As of Dec. 31, 2013, financial institutions raised $550 billion unsecured on financial commercial paper markets and the dealer banks used the tri-party repo market to borrow on a secured basis close to $1.6 trillion. 80% of the collateral posted is Treasuries and Agencies. Only $330 billion of private sector assets are financed on this market.

2. Shadow Banking is a Funding Mechanism for Assets

Before the crisis, the shadow banking system played an important role in funding assets with liabilities that were secured by assets that were held off of bank balance sheets in bankruptcy remote vehicles. When this secured asset funding relied on bank support, it was usually provided by ABCP conduits. When this secured asset funding was made possible by a tiered liability structure, it was provided by SIVs and similar vehicles. When this secured asset funding relied on a maturity-matched repo, it was provided by a repo conduit.

Before the crisis the ABCP market was the most important source of shadow bank funding of private sector assets. (Not only did the tri-party repo market fund private sector assets that were for the most part on dealer bank balance sheets, but it was dominated by Treasuries and Agencies and thus played a relatively small role in financing private sector assets even indirectly.[24]) In post-crisis markets vehicles like ABCP and repo conduits are financing far fewer assets than they did before the crisis. The ABCP market is continuing its slow but steady decline over time and now hovers in volume around $250 billion.

3. Shadow Banking Allows Money Market Issuers to Rent Bank Credit and Allows Banks to Avoid Capital Requirements

When assets were directly financed by the shadow banking system, it was usually because financing vehicles paid a small fee to "rent" a commercial bank's credit rating by purchasing a guarantee of the vehicle's liabilities. Because these guarantees were off-balance sheet, the bank was able to avoid the capital requirements that would have been imposed if the bank had done the lending itself. The role played by the clearing banks in the tri-party repo market was similar: they provide intraday credit in order to give dealer banks access to their assets during the day, but faced no capital charge for the credit. Thus, a key function played by shadow banking is the arbitrage of capital regulations.[25]

The liquidity and credit facilities provided by banks to ABCP conduits are examples of unsecured bank guarantees.[26] By contrast, the tri-party clearing banks provide secured guarantees. The intra-day credit

[24] Arvind Krishnamurthy, Stefan Nagel & Dmitry Orlov, *Sizing Up Repo* 22 (NBER Working Paper No. w17768, 2012).
[25] Carolyn Sissoko, Note, Is financial regulation structurally biased to favor deregulation, 86 Southern California Law Review 365 (2013). Sissoko also has a discussion of the broader literature on the role of regulatory arbitrage in the ABCP market.
[26] See id. for details.

that the clearing banks provide to the dealer banks is secured by the collateral that has been posted on the tri-party repo market. Banks may also issue guarantees in the form of swaps that offset the market risk of collateral; these guarantees may be secured or unsecured depending on the derivative contract.

The collapse of the ABCP market since regulators have become attuned to the problem of regulatory arbitrage of capital requirements is just another piece of evidence that the vast majority of financing on the ABCP market at its peak was not driven by economic efficiencies, but by regulatory arbitrage as banks used liquidity and credit facilities to take on credit risk, while avoiding capital requirements. Indeed, the industry reaction to the 2004 Final Regulation governing such liquidity facilities – which resulted in a "reinterpretation" of the regulation that effectively gutted it – is also evidence of the importance of regulatory arbitrage to this market.[27]

C. Collateralized Money Markets Are More Unstable Than Traditional Banks

The use of collateral in repo markets makes them particularly unstable for two reasons: leverage and the fact that not just lenders, but borrowers, can start a run.

When the price of the collateral in a repo contract falls, the borrower is typically required to post more collateral within a day, and, in the event that the collateral call is not met, the collateral that was posted can be liquidated immediately. While this description shows how quickly market price changes can be reflected in the sale of collateral on repo markets, it does not take the leverage that is ubiquitous on repo markets into account. Because of leverage small changes in the market price of an assets can force the borrower to sell off a large fraction of the borrower's holding of that asset.

An example (drawn from a Fitch Ratings report) will make the instability inherent in repo market finance more clear.[28] Consider a borrower with a $5 million equity stake, which uses repo markets to finance the purchase of a $105 million portfolio of corporate bonds on which the lender imposes a 5% haircut, so that $1 can be borrowed for every $1.05 in collateral repo'd. The borrower will therefore have a leverage ratio of 21 to 1. A 2% decline in the value of the portfolio would reduce the total portfolio value to $102.9 million, reducing the equity in the portfolio to $2.9 million. If we assume that the borrower has no additional equity to contribute, the borrower can now only finance a $60.9 million portfolio at a 5% haircut. In short, because of the leverage inherent in using repo markets to finance assets, a 2% drop in portfolio value can force a sale of 42% of the assets held. Note that this example doesn't take into account the possibility that the lender increases the haircut on the repo, which would mean that even more of the assets had to be sold. In short, once a borrower has maximized the use of leverage on repo markets – whether the borrower does this intentionally in order to "maximize" returns or simply ends up in this situation after the collateral has declined in price – very small declines in price can force the borrower to sell a significant fraction of the assets. If the borrower is a large market participant, such as an investment bank, this is likely to be the first step in a liquidity spiral, where asset sales further reduce the value of the collateral and trigger additional assets sales.

Not only does leverage make repo markets inherently unstable, but, in addition, a key characteristic distinguishing the repo market from unsecured credit markets generally is that not only the lenders, but also the borrowers, can start a run. The use of collateral in bilateral repo markets makes a borrower run

[27] See Sissoko, Deregulatory Bias at.

[28] Fitch Ratings, Repo Emerges from the "Shadow" 8 (Feb. 3, 2012).

possible, because the collateral can be rehypothecated, or posted as collateral in a subsequent loan by the recipient of the collateral. In short, the collateral posted by borrowers in the bilateral repo market is a source of liquidity for the lender.

When borrowers decide that they don't want to be exposed to a troubled lender that may not be able to return the borrowers' collateral in the event that it fails, the borrowers may seek to transfer their accounts to a lender who is not troubled. When the borrowers' accounts are transferred, the collateral they have posted it transferred with the accounts, and the troubled lender loses the liquidity that was provided by that collateral.

As a result of this property of the repo market, the dealer bank failures of 2008 were characterized by "runs" by both prime brokerage clients and other dealers, none of whom wanted to be exposed to a failing bank. In fact, Krishnamurthy, Nagel, & Orlov conclude that the evidence supports the view that the 2008 crisis looks more like an inter-dealer credit crunch than a run by end investors on the two firms.[29] For these authors one factor distinguishing the two types of runs is the fact that the dealers are well-informed market participants, whereas end investors typically must decide whether to pull out of the market based on very limited information.[30] In short, it is possible that, far from being comparable to bank runs, the runs that took place in 2008 were runs that started with the most informed participants in financial markets.

Thus, there are two very important differences that make the repo market more unstable than unsecured funding markets. Not only does leverage mean that a small decline in price can easily force a large sale of assets, but in the bilateral repo market a run can be started not only by lenders, but also by borrowers.

D. Shadow Banking is Not "Market-Based" Lending

The "market-based" credit system is often contrasted with the "bank-based" credit system to distinguish environments where firms raise funds by issuing securities on markets from those where firms raise funds by borrowing from banks.[31] This distinction is clear when we focus on long-term capital markets, such as bond markets where established companies can and do raise money on a regular basis.

When it comes to money markets, however, the line between market-based and bank-based systems cannot be clearly drawn, because the so-called market-based systems rely heavily on guarantees provided by the banking system. In commercial paper markets non-financial companies, including ABCP conduits, can only borrow on these markets if they have liquidity support, usually in the form of a liquidity facility, repo or swap provided by a bank. As for the tri-party repo market: first, it is not a market for the direct funding of assets, but a market for funding assets that sit on dealer bank balance sheets; and second, it was also backstopped by guarantees provided by the tri-party clearing banks, which bore the credit risk of the dealer banks during the day. In short, in the money markets the "market-based" credit system might as well be called the "bank-guaranteed" credit system.

Another sense in which money market instruments are only nominally "market-based" is that these assets do not trade on secondary markets. Commercial paper is placed and almost never resold before maturity.

---

[29] Arvind Krishnamurthy, Stefan Nagel & Dmitry Orlov, *Sizing Up Repo* 19, 22 (NBER Working Paper No. w17768, 2012).
[30] Id. at 6.
[31] See, e.g., Michael Woodford, Financial Intermediation and Macroeconomic Analysis, 24 J. Econ. Perspectives 21, 21 (2010).

Repo obligations, similarly, are not traded actively, but held until maturity.[32] By contrast, in 19th c. London, the archetype of a traditional bank-based credit system, there was an active secondary market in the bills that were the primary tool by which central bank policy was implemented.

E. The Regulatory Question: Should Wholesale Funding Markets Be Curtailed?

It is misleading to describe the shadow banking system that exists today as "money market funding of capital market lending" and to focus on it as a means of financing assets,[33] because at present by far the most important use of shadow banking instruments is to provide wholesale funding for investment banks and through them indirect financing of assets that sit on their balance sheets. Although the view that shadow banking finances assets directly may have held some truth prior to the crisis when $1.2 trillion of ABCP financed bankruptcy remote vehicles, today, to the degree that shadow banking disintermediates commercial banks, it does so by reintermediating investment banks – using a form of funding that is even more unstable than deposits.

The financial crisis made obvious the dangers of the wholesale funding that is provided by the shadow banking system. Money market fund investors can run from financial commercial paper just as quickly as bank depositors can run on a bank. The repo market is even more unstable than commercial paper, as was discussed above, because the leverage in the market means that fire sales can be triggered by small declines in price, and because both borrowers and lenders can participate in a run.

Regulators have effectively addressed the role played by shadow banking in making it possible for banks to avoid capital requirements while supporting the issue of ABCP by off-balance sheet vehicles. By doing so they appear to have smothered shadow banking as a means of financing assets directly. The forms of shadow banking that have yet to be suppressed by regulators are the secured and unsecured wholesale funding systems.

For this reason, an important question that regulators have yet to answer is whether these wholesale funding markets are a valuable addition to the financial system or whether the risk of instability that accompanies them is so great that lending on these wholesale markets should be curtailed.[34] The remainder of this paper will argue that the repo markets that provide collateralized wholesale funding to investment banks undermine economic growth by tending to reverse the revolution in unsecured lending that took place when modern banking was first established and by exacerbating a decline in the credit quality of our financial institutions.

II. What Do Banks Do? A Historical Perspective

This paper argues that the birth of modern banking was a revolution in unsecured lending that fostered modern economic growth by overcoming the liquidity constraints that are ubiquitous in an unbanked economy. The key elements of the revolution are (i) private banks that, first, are trusted not to default, and that, second, use their own ability to borrow to make it possible for the other members of the economy to

[32] The collateral posted against a repo can often be rehypothecated, but this is very different from the resale of the debtor's obligation that takes place in secondary markets.

[33] Perry Mehrling, Zoltan Pozsar, James Sweeney, and Daniel Neilson, Bagehot was a Shadow Banker (Nov. 2013).

[34] See Dudley, Wholesale Funding.

borrow by offering them overdraft accounts and monitoring their behavior, and (ii) a central bank that can provide liquidity to protect this credit system from collapsing at the first hint of trouble.

In order to understand this financial revolution, this paper first explains the context of the monetary system in which the revolution took place, and then the development of the banking system itself. Then, the relationship between banking and economic growth, as it was understood by contemporary monetary theorists, is explained, as are the elements of the banking revolution. The next subsection explains how this banking revolution created a money market that was both based on unsecured private sector instruments and virtually risk-free by requiring every seller of the debt to guarantee it, by imposing capital calls on joint stock bank owners to cover the debts of a bank in liquidation, by circulating only short-term instruments, and by providing a lender of last resort to deal with liquidity crises. The final subsection why this understanding of the banking revolution is closer to the facts than the current conventional wisdom regarding banking in 18th and 19th c. Britain.

A. Money in 18th-19th c. Britain

1. Real Bills and the Decentralized Early Modern Monetary System

As the middle ages were coming to a close Europe developed a financial instrument called the bill of exchange that was managed by a network of wealthy merchant bankers. In its initial form the bill of exchange was used to finance international trade between Europe's leading cities. The bill was a short-term debt contract that was payable in a foreign country. Clearing mechanisms enabled trade in these bills to minimize the transport of gold and silver across Europe.

Starting in the sixteenth century the bill of exchange evolved into a very different instrument. Endorsement allowed bills to circulate from hand to hand before being redeemed, and domestic bills became the norm in highly developed commercial economies.[35] Local bankers managed local networks and stood ready to discount bills before they were due. Thus, a tradesman with a local bank account could write a bill in the name of a supplier, who could then choose to hold the bill, endorse the bill over to a creditor of his own, or cash it – less a discount – at the bank. The bill was a form of commercial paper that was endorsable and effectively allowed banks to underwrite a system of trade credit for the local community. The result was that trade in urban economies began to take place on the basis of a paper monetary system that was supported by a network of banks.

In order for the system to work, standards had to be put into place to prevent the local tradesmen from writing too many bills. In practice a single principle was used to regulate this credit system: A bill was valid only if it was issued in exchange for goods. Bills that were written in the absence of a real exchange were described as "fictitious" or "accommodation paper." Any tradesman who was caught issuing fictitious bills was considered a fraud and excluded from the financial network. Suspicion of such fraud could also derail a tradesman's career.

The principle that bills were valid only when they were issued in exchange for real goods – or when they were "real bills" – was the standard our early modern ancestors put in place to ensure that finance served

[35] Herman van der Wee, *Monetary, Credit and Banking Systems*, Cambridge Economic History of Europe vol. V at 322-32 (1977).

the needs of trade. It had the advantage of being applied at the individual level, creating a completely decentralized means by which the issue of financial paper could be controlled.

Nowadays the term real bills is used only in the context of "the real bills doctrine," which played an important part in the debate over monetary policy that took place in England in the early 19th century, and is associated with Adam Smith. The Currency School argued that the Bank of England should be constrained to issue bank notes in an amount that did not exceed the amount of gold it held in its vaults, while the Banking School argued that, since the bills circulating were real, this fact was sufficient to control the money supply and that the Bank needed to have the flexibility to issue an indeterminate quantity of bank notes when discounting real bills. It's worth noting that the idea that the only valid bills were real bills was so fundamental to the 18th and early 19th century concept of financial stability, that no one questioned it.[36] The issue in the debate was whether or not limiting the circulating money supply to real bills was alone sufficient to ensure financial stability.

The denouement of this controversy took place when the Bank Charter Act of 1844 was passed. This was effectively a compromise. Only the Bank of England was allowed to issue bank notes and the Bank's issue was fixed by the amount of gold in its vaults; however, the Act was subject to suspension by executive order. In practice, this meant that the Bank of England's note issue was restricted – unless economic circumstances required a greater supply of notes. The Act was temporarily suspended in 1847, 1857 and 1866.

2. Acceptance Finance and the Rise of Banks as Arbiters of Credit

In the meanwhile, the British economy was steadily outgrowing the restriction to circulating only real bills. By the start of the 19th century in England the system of domestic bills had evolved into acceptance finance. A country tradesman who regularly shipped his wares to a London middleman for sale would draw on his account with the middleman when making purchases in his own local community. The tradesman would write a bill drawn on the London middleman to pay his local supplier. The supplier would go ahead and circulate the bill through endorsement. However, until the bill was discounted at the local bank, sent by the banker off to his London correspondent for settlement and formally accepted by the London middleman as an obligation, there was no certainty under the law that the middleman would pay.[37]

In short, acceptance finance was a prototype for the checking account system that would develop decades later – just like a checking account system it required that (i) bad bills or checks be passed infrequently and (ii) middlemen or bankers could be relied on to honor their obligations. When one recognizes the sophistication of the financial system in Britain at the turn of the 19th century, one begins to understand why Henry Thornton considered the "science" of credit to be the fundamental source of British growth at the time.[38]

Now here is the question: Is the bill drawn by the country tradesman on the London middleman a real bill or a fictitious bill? Assuming they have an ongoing relationship is there anything wrong with a middleman accepting the bill before he has received a delivery of goods? Is there anything wrong with a

[36] James Rogers, *The Early History of the Law of Bills and Notes*, Cambridge University Press, 1995, pp. 232.
[37] James Rogers, *The Early History of the Law of Bills and Notes*, Cambridge University Press, 1995, pp. 113, 171-173, 188-189.
[38] Henry Thornton, *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*, 1802, pp. 175 – 176.

middleman extending an overdraft to a tradesman? It was probably inevitable that the practice of acceptance finance broke down the cultural barriers that had supported the restriction to the circulation of real bills only. Henry Thornton's *Paper Credit* makes it clear that by the early years of the 19th century, some British bankers were beginning to realize that a "good bill" could be backed by nothing more than an individual's personal credit.

Legal cases demonstrate that the use of accommodation paper was growing – and becoming more acceptable – through the first decades of the 19th century.[39] The Banking Act of 1844 started a different trend: banks that were no longer allowed to issue bank notes found another way to create money, the checking account. These two trends combined to create a new financial system centered around banks as the arbiters of credit.

The 19th century witnessed a transition from a decentralized system of paper money that was controlled by the principle that only real bills were valid to a more complex system in which short-term monetized credit was allocated by banks. A new approach had to be found to control the growth of the new monetary system based on checking accounts. The Banking Act of 1844 had been a first effort at direct control of the money supply. It was unsuccessful in many ways: In the first quarter century after it was passed, it had to be suspended three times in order to protect the economy from the ravages of liquidity crises. And the growth of checking accounts effectively neutered the Act.

In the meanwhile, however, the Bank of England had a discovered a new tool for controlling the money supply. In the 18th century the Bank's discount rate had remained fixed at 5%. As a consequence in normal times competing banks took most of the trade, and the Bank's discount business was relatively small. In a liquidity crisis, however, the Bank's discounts would increase astronomically for a few days or even weeks only to fall back to normal when the panic had eased.

In the first half of the 19th century a major concern of the Bank was the maintenance of its gold reserves. Thus, the outflow of gold that was associated with crises and strong demand for discounts at the Bank caused concern. It didn't take long for the Directors of the Bank to realize that by raising the discount rate, they could moderate the outflow of gold.[40] In the 1820s Bank Rate, or the discount rate of the Bank of England, started to be used as a policy tool. By the middle of the century Bank Rate was the principal policy tool that the Bank used to control the flows of gold to and from the Bank and to moderate the growth of credit and of the money supply.

3. The Banking System Gives Birth to Fiat Money

England developed a paper monetary system in the late 18th century. The monetary system was not uniform across the country. In commercial regions a large fraction of the circulating currency took the form of domestic bills. Local bank notes were often an important part of the currency too, especially in agricultural districts. Bank of England notes were issued in large denominations and were important for settling interbank accounts, but circulated very little in the countryside.

---

[39] James Rogers, *The Early History of the Law of Bills and Notes*, Cambridge University Press, 1995, chapter 10.

[40] In fact, this obvious possibility had been raised in 1802 by Thornton who makes it clear that usury laws interfered with the operation of this mechanism. *Paper Credit*, p. 254.

This state of affairs changed dramatically in 1797. The finance of the Napoleonic Wars had put an enormous strain on the financial system and the Bank of England risked running out of gold. The solution was the suspension of the convertibility into gold of the Bank of England note. This suspension lasted for almost a quarter of a century.

As the local banking networks had relied through their London correspondents on the gold reserves of the Bank England in order to meet the demands of their own customers, it was no longer possible for local banks to pay out their notes in gold upon request. To resolve the settlement problem in the countryside, the Bank of England began to issue notes in small denominations – making it possible for the local banks to pay out Bank of England notes instead of gold.

Thus, at the turn of the century the British economy shifted very smoothly from a gold standard to a fiat money standard. During the war the economy experienced a moderate level of inflation with the result that when the war finally ended in 1815 it was not immediately possible to resume convertibility of the Bank of England note into gold at the rate that prevailed in 1797. Policymakers, however, were committed to resumption at the original exchange rate. Thus, in the years following the Napoleonic Wars the British economy was put through a severe recession and in 1821 convertibility of the Bank of England note was restored.

Despite the fact that gold was now readily available, country banks continued to settle their obligations in Bank of England notes with frequency for the simple reason that Bank notes were accepted by almost everyone. Bank of England notes displaced gold as a means of settling trades, because they were in practice "good as gold".

Thus, the foundations of a modern banking system were laid in 19th century Britain. Paper bank notes were universally accepted in final settlement of debt. The banking system offered checking accounts to the general public and short-term credit to those that met the criteria of the bankers. And finally the whole system was moderated by the Bank of England's control over the short term interest rate on bills discounted at the Bank, and, as will be discussed in detail below, supported by the Bank of England's practice of acting as a source of liquidity for the whole banking system during a financial crisis.

B. The Relationship Between Banking and Growth as Understood by Early Monetary Theorists

In the late 18th and early 19th centuries Adam Smith and Henry Thornton analyzed the role played by the banking system in the economic growth of Britain. One chapter of *The Wealth of Nations* is devoted to explaining how banks contribute to the recent increase in trade and industry of Britain. Both authors saw the growth that was taking place around them and that made Britain exceptional.[41] They both understood that banks played an important role in this growth by expanding the money supply. They both recognized the role that the Bank of England played in supporting the banking system. These two authors explained that Britain's growth on the cusp of the 19th century took place because Britain had banks in most towns that were trusted not to default on their unsecured obligations, that used this trust to lend broadly on an unsecured basis within the local community, that monitored the borrowers, and that relied on a central bank to support this system of unsecured credit when its existence was at risk due to a panic.

[41] Smith, II.2.40-41; Thornton at 175-76.

For Smith, the mechanism by which banking contributed to growth was by expanding the metallic money supply and conserving on the circulation of gold and silver.[42] Thornton also emphasized the importance of the expansion of the money supply that the use of bills entailed, although he criticized Smith for failing to recognize the full degree to which bills could expand the money supply.[43] Thornton makes it clear that the bills that circulated in Britain were not only unsecured, but also were likely to circulate in amounts greatly in excess of the goods that made them "real." He writes:

> it may be observed, first, that the notes given in consequence of a real sale of goods cannot be considered as, on that account, *certainly* representing any actual property. Suppose that A sells one hundred pounds worth of goods to B at six months credit, and takes a bill at six months for it; and that B, within a month after sells the same goods, at a like credit, to C, taking a like bill; and again that C, after another month, sells them to D, taking a like bill, and so on. There may then, at the end of six months, be six bills of £100 each existing at the same time; and every one of these may possibly have been discounted. Of all these bills, then, one only represents any actual property.[44]

In short, circulating bills could play the role that they did in increasing the money supply because they were not collateralized either by an assets like gold or by goods.

Both Thornton and Smith recognize that individual banks were able to receive deposits and issue circulating liabilities in this environment where almost all credit was unsecured, because they were trusted by their communities. Smith writes:

> When the people of any particular country have such confidence in the fortune, probity, and prudence of a particular banker, as to believe that he is always ready to pay upon demand such of his promissory notes as are likely to be at any time presented to him; those notes come to have the same currency as gold and silver money, from the confidence that such money can at any time be had for them. [45]

Thornton, too, describes a system where small-town bankers were able to circulate liabilities within their local communities because of trust.[46]

Only Thornton, however, was a London banker by profession, and as a result only he explains in detail the role that banks play in the monetary system by supporting a system of credit that is "conducive to the interests of trade." He explains that the local commercial community supported the circulation of a banker's liabilities, because they were the ones who benefited from the banker's loans.[47] Thus, the key function of banks was to provide credit to commerce that would not have been available in the absence of

---

[42] Smith, II.2.26 ff.

[43] Henry Thornton, *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*, 1802, p. 92.

[44] Henry Thornton, *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*, 1802, p. 86. *See also* Marc Flandreau and Stefano Ugolini, Where it all began: lending of last resort and the Bank of England during the Overend-Gurney panic of 1866, Norges Bank Working Paper 2011-3 at 22 (2011). Note that some modern scholars have claimed that 19th c. bills were "typically collateralized by tradable goods." Perry Mehrling, Zoltan Pozsar, James Sweeney, and Daniel Neilson, Bagehot was a Shadow Banker 5 (Nov. 2013). Presumably these authors are confusing a "real" bill with a collateralized bill.

[45] Smith at II.2.28.

[46] See Thornton, at 180 ("They know, that if the character of their house should be brought into question, through the fears or even the caprice of any of those strangers into whose hands their circulating paper passes, some distrust may be excited among their customers, the effect of which may be a sudden demand for the payment of large deposits.")

[47] See Thornton, at 173 ("The circumstance which chiefly operated in procuring currency to the new circulating paper, was that participation of the benefit resulting from it which was enjoyed by the customers of the country banker; for he lent among them the capital which was acquired by the issue of his paper, and they became his instruments in sending it into circulation, by accepting it as a ready-money payment in return for bills discounted.")

the circulation of bankers' liabilities.[48] Of course, as lenders, bankers were also monitors of those who borrowed. For Thornton, the key to the economic growth that he was witnessing in Britain was the fact that Britain had bankers in most country towns monitoring the unsecured credit that was being extended within the local community. He writes:

> Through the creation of banks, the appreciation of the credit of numberless persons engaged in commerce has become a science; and to the height to which this science is now carried in Great Britain we are in no small degree indebted for the flourishing state of our internal commerce, for the general reputation of our merchants abroad, and for the preference which in that respect they enjoy over the traders of all other nations. It is certainly the interest, and, I believe, it is also the general practice, of banks to limit not only the loan which anyone trader shall obtain from themselves, but the total amount also, as far as they are able, of the sum which the same person shall borrow in different places; at the same time, reciprocally to communicate intelligence for their mutual assistance; and, above all, to discourage bills of accommodation.[49]

Thornton observes that the existence of this system meant that bankers also received deposits that could be lent out alongside the bankers' own funds, but he makes it clear that this is not a core function of banking, but just another advantage of having a banking system.[50]

Not only did Smith and Thornton have a profound understanding of the role played by banks in the economy, but they both were aware that the Bank of England was a lender of last resort to the banking system. In his discussion of banking Smith remarks on the important role played by the Bank of England in "support[ing] the credit" of the largest banks in England, Germany, and Holland. He emphasizes this point with an anecdote: "in 1763, [the Bank] is said to have advanced for this purpose, in one week, about £1.6 million, a great part of it in bullion."[51] At the time, £1.6 million was approximately equal to a quarter of the Bank's total liabilities.

Thornton is generally viewed as the first theorist who explained in detail the lender of last resort role played by a central bank.[52] He writes: "That a state of distrust causes a slowness in the circulation of guineas, and that at such a time a greater quantity of money will be wanted in order to effect only the same money payments, is a position which scarcely needs to be proved."[53] He continues to establish the difference between a central bank, which "is completely subjected to the interests [of the public]" and a "private house" which "may be in general considered as having in the bank [of England] a sure resource."[54] Thornton defends the 1797 suspension by the Bank of England of payments in gold and argues that, if the Bank of England erred in this time period, it erred "on the side of too much restricting

---

[48] See Thornton, at 173-74.

[49] Thornton, at 175-76.

[50] Thornton, at 174.

[51] Smith at II.2.85. See also John Clapham, The Bank of England: A History, vol. I, at 239 – 40; Charles Wilson, Anglo-Dutch Commerce and Finance in the 18th Century 179, reprint of 1941 ed. (1977). Note that modern scholars are apparently unaware of the role played by the Bank of England in supporting the Dutch banking system through the crisis of 1763. See, e.g., Stephen Quinn & William Roberds, *Responding to a Shadow Banking Crisis: The Lessons of 1763* at 2 (FRB of Atlanta Working Paper No. 2012-8, 2012); Isabel Schnabel & Hyun Song Shin, *Liquidity and Contagion: the Crisis of 1763*, 2 J. Eur. Econ. Ass'n 929, 945 (2004).

[52] Forrest Capie, 200 year of financial crises: lessons learned and forgotten 11 (2012); Thomas Humphrey, Lender of Last Resort: What it is, Whence it came, and Why the Fed isn't it, 30 Cato J. 333, 334 (2010).

[53] Thornton, at 99.

[54] Thornton, at 126-27.

its notes in the late seasons of alarm."[55] In short, Thornton makes it abundantly clear that a central bank which can expand the money supply is needed to support the banking system through a liquidity crisis.

Thornton does not only demonstrate a comprehensive understanding of the role played by the lender of last resort in the banking system; he also claims that the existence of such a reliable source of high-powered money "promote[s] the institution of private banking."[56] In other words, Thornton makes it clear that the growth of the banking system is closely tied to the presence of a central bank that stands ready to provide liquidity to the banking system as a whole and to protect the unsecured credit system from collapse.

At the dawn of modern banking, contemporary monetary theorists drew a clear connection between the system of unsecured credit that was underwritten by the banking system and economic growth. The essential elements of this system of unsecured credit were (i) private banks that, first, were trusted by the members of the local community and that, second, used their own ability to borrow to make it possible for the other members of the economy to borrow by lending to them and monitoring their behavior, and (ii) a central bank that provides liquidity to protect this unsecured credit system from collapsing when the trust on which the system is based is thrown temporarily into doubt.

C. How Risk-Free Assets Were Created by 19th c. Bankers

This system of trusted bankers who underwrote the debt of the broader economy and were supported by the central bank was able to generate assets that were effectively risk-free. Because of the high quality of the assets, the crises that took place in 19th c. Britain were pure liquidity crisis, and, when the Bank of England acted as a lender of last resort, almost all losses on the assets were borne by market participants, not the Bank. There were no central bank bailouts of banks in 19th c. Britain.

The Bank of England, which was a publicly listed company and had no formal obligation to support the banking system, was willing to provide liquidity to the banking system, because it did so by purchasing assets that were virtually risk-free even though they were unsecured. After explaining the costs associated with managing collateral, Thornton concludes that these costs clearly exceeded any benefits from taking collateral, given the quality of the assets the bank purchased: "the bills which the bank discounts, are, generally speaking, so safe, that the security either of goods, or stocks, or land, none of which are received in pledge by the directors, may be considered as nearly superfluous."[57] Seventy years later, Walter Bagehot explained that the lender of last resort could reject dubious assets – because there are so few of them in a "commercial country":

> No advances indeed need be made by which the Bank will ultimately lose. The amount of bad business in commercial countries is an infinitesimally small fraction of the whole business. That in a panic the bank, or banks, holding the ultimate reserve should refuse bad bills or bad securities will not make the panic really worse; the 'unsound' people are a feeble minority, and they are afraid even to look frightened for fear their unsoundness may be detected.[58]

---

[55] Thornton, at 127. Thornton continues the discussion of the lender of last resort on pages 179 ff.

[56] Thornton, at 90. Thornton later explicitly prescribes the lender of last resort role of the central bank as a means of ensuring the stability of the money supply and promoting the banking system. Thornton, at 188.

[57] Thornton, at 186.

[58] Bagehot, Lombard Street, VII.59.

And after careful analysis of the data economic historians have found that these contemporary authors were indeed correct: "The Bank of England operated in an almost perfectly risk-free market, whereby losses were entirely transferred to market participants."[59]

Four properties made it possible for the British banking system to create these risk free assets: the "wrapping" of the debt by every party who sold it, the requirement that even stockholders in a failed joint stock bank protect creditors by contributing to the liquidation of the failed bank, the limitation of the money market to short-term debt, and the presence of a lender of last resort which made it possible for the money market to survive a financial panic without collapsing.

1. Banks Wrapped the Debt that They Sold

As was discussed above, a watershed in financial history occurred in the sixteenth century when the legal innovation of endorsement made it possible to transfer a debt obligation from one creditor to a new claimant. Historians have discussed at length the importance of the fact that every endorser is liable in full for the debt.[60] In modern financial terms we might say that each endorser wraps the debt by providing his own guarantee that it will be paid.

This legal structure was integral to the circulation of 19th century bills in the British money supply. Not only was the original issuer of the bill liable for its payment, but so was any banker who discounted it by endorsing it over to another party. The Bank of England only discounted bills that had been accepted by a merchant bank,[61] which thereby put its own guarantee to pay on the bill. (The closest analogy in modern banking is to a checking account system where all the checks are post-dated and payable only when due: by "accepting" the check the bank promises to make payment on it when it is due – independent of whether there are funds in the check-writer's account on the due date.)

Once a bill had been accepted, it circulated as a liquid money market instrument in London – because it could be discounted at the Bank of England by any of several hundred eligible discounters. In short, the liquidity of assets in 19th c. London derived from the combination of bank guarantees and the availability of the central bank as a backstop.[62]

As a result of this system, every bill discounted by the Bank of England carried with it three independent promises of payment: that of the issuer, that of the acceptor, and that of the discounter.[63] Observe that the legal structure of such guarantees was designed to align every party's interest in the quality of the underlying debt. Because the discounter was fully liable for the debt, debt was likely discounted only if

---

[59] Vincent Bignon, Marc Flandreau, & Stefano Ugolini, Bagehot for beginners: the making of lender-of-last-resort operations in the mid-nineteenth century, 65 Econ. Hist. Rev. 580, 602 (2012).

[60] Herman van der Wee, "The Medieval and Early Modern Origins of European Banking," in Banchi pubblici, banchi privati e monti di pietà nell'Europa preindustriale, 1159, 1162-70 (1990). See also Larry Neal, The Rise of Financial Capitalism: International Capital Markets in the Age of Reason 5-9 (1990).

[61] By the second half of the 19th c. the list of eligible acceptors was much broader. Marc Flandreau and Stefano Ugolini, Where it all began: lending of last resort and the Bank of England during the Overend-Gurney panic of 1866, Norges Bank Working Paper 2011-3 at 19 (2011).

[62] Note, however, that it was not unusual for bills to circulate locally without a London acceptance. See, e.g., T.S. Ashton, The Bill of Exchange and Private Banks in Lancashire, 15 Econ. Hist. Rev. 25 (1945).

[63] Marc Flandreau and Stefano Ugolini, Where it all began: lending of last resort and the Bank of England during the Overend-Gurney panic of 1866, Norges Bank Working Paper 2011-3 at 22 (2011).

the discounter had reason to believe that it was good. Similarly, the acceptor would only accept the bill if the issuer of the bill either had funds (or more accurately goods) deposited with the acceptor or the acceptor was confident that value of the bill would soon be available. Thus, a bank-based monetary system developed in an environment where all of the parties transacting in debt were liable and none had any interest in circulating debt that wasn't of high quality.

In 19th c. Britain there was an active money market in bills, that were tradable because they had been accepted – or wrapped – by merchant banks. In other words, it was the guarantees provided by banks that made secondary markets in debt possible.

2. Bank Owners Were Liable for Unpaid Debts

The fact that issuers and bankers faced unlimited liability – or capital calls if they were stockholders – on their obligations also played an important role in the risk-free nature of 19th c. money market assets. While the unlimited liability of partners in an unincorporated bank is well-understood, many do not realize that the joint-stock banks that proliferated over the course of the 19th c. in Britain were also structured so that the owners would face a capital call in the event that there were unpaid debts in bankruptcy.

Joint stock investors in 19th c. Britain typically paid only a fraction of the par value of the shares. For this reason, the corporation – or in the case of bankruptcy the liquidator – retained the right to call the remaining value of the shares until par was fully paid up. Thus, when Overend, Gurney & Co. failed in 1866, the joint shareholders were required to pay to the liquidators 50% of par, more than their initial investment of 30%, and, as a result, the creditors were finally paid in full.[64]

Given this liability structure, in 19th c. Britain one of the reasons for confidence in the guarantees provided by the banking system was the visible personal wealth of the bankers themselves and the knowledge that this wealth was at stake.[65]

3. The Assets That Were Financed Were Short-Term

The general policy of the Bank of England was to discount only bills with less than three months to run. For this reason a standard reason for rejecting a bill presented for discount was "beyond 95 days."[66] Presumably in a crisis such policies could be relaxed to some degree; it is unlikely, however, that the in the 19th c. the Bank discounted significant amounts of paper that could be considered long-term. Because of this policy, the bills that were actively traded on the 19th c. British money market were those with three months or less to run.

The term of the bills that were discounted undoubtedly played an important role in the fact that they were virtually risk-free. Financial market participants have always understood that fewer unexpected events are likely to occur over the course of three months, than over the course of a year, and certainly than over the

---

[64] Bankers Magazine, Supplement, Overend, Gurney & Co. Trial 18 (January 1870); Margaret Ackrill and Leslie Hannah, Barclays: The Business of Banking, 1690-1996 at 46-7 (2001).

[65] A lyric for Gilbert & Sullivan's *Trial by Jury* makes use of the phrase "rich as the Gurneys," referring to banker-cousins of those Gurneys who were involved in the Overend collapse.

[66] Marc Flandreau and Stefano Ugolini, Where it all began: lending of last resort and the Bank of England during the Overend-Gurney panic of 1866, Norges Bank Working Paper 2011-3 at 8 (2011).

course of five years. Short-term debt has the advantage that its behavior is much more predictable than long-term debt, and this fact contributed to the risk free character of the 19th c. money market.

4. The Central Bank Backstopped the Bank-Based Credit System

Bank guarantees, bank owner liability, and the short-term nature of the debt ensured that any accepted bill circulating in the London money market was intrinsically a high-quality bill. Contemporaries were well aware, however, that the quality of the money supply was not enough to protect it from instability. Both Henry Thornton and Walter Bagehot explain in detail the nature of a liquidity crisis.[67]

The basic problem with a credit-based monetary system is interconnectedness. Although every member of the economy can pay his debts on time as long as those who are indebted to him pay on time, when one large member of the economy fails to pay on time, the repercussions can spread in a chain throughout the economy. If five of the initial defaulter's creditors have to suspend payments too, each setting off a new sequence of failures, the whole economy can be affected.

The financial panics that took place in 19th c. Britain were driven by fear of such a chain of failures. Every bank held liquid assets that could help it weather losses. In a crisis, however, everyone in the economy would simultaneously foresee a need for a larger stock of liquid assets and would tend to draw liquid assets out of the banking system, while the banks themselves would also seek to increase their stores of liquid assets. As a result, the stock of high-powered money that was circulating in the banking system would shrink, and both bankers and tradesmen who needed access to that stock of high-powered money would be at risk of failure. The role of the Bank of England in this situation was to ensure that high-powered money was readily available for those that had accepted bills to discount.

While Bagehot does not emphasize the monetary nature of this process, Thornton explains very clearly that the lender of last resort's job is to prevent the money supply from collapsing.[68] Money markets are not capable of providing a consistent source of liquidity, because adverse events tend to create sudden shifts in the demand for money or liquid assets.[69] The job of the lender of last resort is to supply liquidity in order to meet this additional demand, and as a result to maintain the liquidity of the money market. This action has the effect of preventing failures that are caused only by the fact that an accepted bill has, in the midst of a crisis, become entirely illiquid, even though in normal times it could be discounted anywhere.

Because money markets are subject to liquidity crises, the intrinsic qualities of the assets that circulate on the money market are insufficient to protect the market from collapse. The lender of last resort was the key player that made it possible for the assets that circulated on the 19th c. London money market to be

[67] Thornton, at 180-88. Bagehot at II.40 ff.

[68] Thornton, at 186-87 ("The observations which have now been made sufficiently shew what is the nature of that evil of which we are speaking. It is an evil which ought to be charged not to any fault in the mercantile body, but to the defect of the banking system. It is a privation which the merchants occasionally experience of a considerable part of that circulating medium which custom has rendered· essential to the punctual fulfilment of their engagements. In good times, the country banks furnish this necessary article, which they are enabled to do through the confidence of the people in general; but when an alarm arises, the country banks cease to give it out, the people refusing what they had before received; and the Bank of England, the only body by whose interposition the distress can be relieved, is somewhat unwilling to exercise all the necessary liberality, for the reasons which have been so fully mentioned.")

[69] Thornton, at 180.

risk-free by stepping in to keep the market from collapsing whenever a panic created a sudden imbalance in the supply and demand for money.

An important point can be drawn from the risk free nature of the assets that circulated in the 19th c. money market: that when a financial crisis took place, the crisis was a pure liquidity crisis. The only job of the lender of last resort was to remedy the liquidity problem. As a result, there were no "bailouts" of the banking system in 19th c. Britain. Even if a bank failed, the creditors were paid in full after the liquidation of the bank – and possibly of a portion of the bank owners' assets – was complete.

D. Contrasting This History with the Current Conventional Wisdom

The description given here of the 18th and 19th c. English banking system is very different from the current conventional understanding of this banking system. Most economic historians describe the English banking system in this period as extraordinarily unstable and prone to impose significant losses on bank creditors[70] – although some acknowledge that the English banking system also provided a "highly sophisticated" payments system.[71] Below I discuss first English bank failure rates, then the question of losses to bank creditors, and finally the role played by the assumption that short-term credit cannot play an important role in economic growth.

When economic historians describe the instability of the English banking system in this period, they are almost always citing one statistic that is derived from data collected by James Gilbart in the 1830s: based on this data Larry White found that from 1809 to 1830 the average of annual failure rates per thousand banks for English country banks was 18.[72] There are many reasons that this number should not be used to represent the quality of the early English banking system. First, this statistic is calculated for the weakest part of the banking system and does not include data on the more stable, but less numerous, London banks. Second, almost nobody observes that White chose his data for comparability to Scottish banking and therefore relies on Gilbart as a source and restricts the years of comparison to 1809-1830. There is a much more authoritative source on English country bank failures: L. S. Pressnell poured over the public records of bankruptcies for the purpose of creating accurate counts. Based on Pressnell, the average of annual failure rates of English country banks from 1760 to 1825 is 11.6 per thousand.[73] Furthermore, because Pressnell is unable to exclude from his data those bankers against whom a claim was filed, but

---

[70] Peter Temin & Hans Joachim Voth, Prometheus Shackled 36-37 (2013) ("banks went bankrupt with alarming frequency" and "[b]ankruptcies, of course, brought … massive losses to many depositors."); Charles Calomiris & Stephen Haber, Fragile by Design 103 (2014) ("Bank failure rates in England were almost five times those in Scotland from 1809 to 1830, and while losses to English debt holders were sometimes significant, there were virtually no losses on Scottish deposits or notes."); Richard Sylla, *Comparing the U.K. and U.S. financial systems, 1790 – 1830*, in The Origins and Development of Financial Markets and Institutions 224 (Jeremy Atack and Larry Neal ed. 2011) (finding that English banks had a particularly high failure rate, but making no claims about losses).

[71] Temin & Voth, at 34.

[72] Lawrence White, Free Banking in Britain 43 (1995).

[73] L.S. Pressnell, Country Banking in the Industrial Revolution (1956). Data analysis by author based on Appendix 20 for country bank failures, and Table 1 (at 11) for numbers of country banks, with a starting number of 12 banks in 1760 (see Pressnell at 4 indicating that there were about a dozen banks in 1750). For years where the number of banks is missing, I interpolate linearly between the points where data is available. Using Pressnell's data the failure rate from 1809 to 1825 is 21.8 per thousand.

who were never declared bankrupt (e.g. because the claim was quickly resolved), this estimate is probably somewhat inflated.[74]

Sylla compares the English failure rate to the failure rate he has calculated of 5 per thousand for the U.S. banking system from 1782 to 1837.[75] As argued above, the failure rate calculated based on Pressnell should probably be used instead of the White calculation used by Sylla. Furthermore, Sylla's data is probably biased downwards compared to Pressnell's, because (i) it includes all banks and is not limited to country banks, (ii) Sylla's determination of "true failure rates" is probably careful to exclude solvent banks from his data, and (iii) this data excludes the late 1830s which include some of the worst years for the performance of U.S. banking.[76] The latter is important, because, for example, if the failure rate is calculated for English country banks from 1760 to 1824, it falls to 10.0 per thousand. In short, there is evidence that English banks failed at a higher rate than U.S. banks, but it seems highly likely that the actual difference in failure rates was less than double. Exaggerated claims about the instability of early English country banking are probably unwarranted.

That monetary factors almost certainly played a significant role in the failure rate of English country banks should also be recognized. The average failure rate while the gold standard was in place, from 1760 to 1797, was 6.9 per thousand, whereas the average failure rate during the suspension of gold convertibility and in its immediate aftermath, from 1797 to 1825 was 16.6 per thousand. The period of suspension was characterized by moderate inflation until the war ended in 1815, and then by deflation until the resumption of convertibility in 1821. These events were then followed by a return to the discipline of the gold standard. It seems likely that bankers found that these changes in monetary policy made bank management difficult and that this played a role in the increase in failure rates that is observed over this period.

In short, there was a problem of instability in the banking system in England in the early years of the 19th century; it seems likely, however, that the stabilization of the monetary environment played as much a role in reducing this instability in later years, as did the adoption of joint stock banking. Furthermore, as Pressnell notes, "financial instability was not introduced by the country banks," but had been a problem long before they were established,[77] and as I wrote in 2003, "If one views the English banking system in the eighteenth century as beset by instability, one must also recognize that it was the first instance where the instability intrinsic to banking was managed so as to protect the banking system as a whole."[78] The role of the lender of last resort is discussed in detail in Section IV.

Several authors assume that the fact that there were many bank failures in England meant that there were significant losses to creditors from such failures.[79] I am unaware of any evidence substantiating this claim. It is important to understand that a bank can be put into bankruptcy whenever the bank is illiquid and has

---

[74] Pressnell at 445.

[75] Sylla 2011, supra note 70, at 224

[76] Richard Sylla, *Early American Banking: The Significance of the Corporate Form*, 14 Bus. & Econ. Hist. 105, 118 (1985).

[77] At 448.

[78] Carolyn Seydel, On the Monetary Role of Banks 77 (dissertation, 2003).

[79] Temin & Voth claim that there were "massive losses to depositors," but do not cite a source for English losses, *supra* note 70, at 37.

failed to pay a debt owed to a creditor when it is due. The fact that a bank was declared bankrupt and liquidated, thus, has no bearing on the question of whether creditors were paid in full after liquidation of the bank. To make the determination that there were losses to creditors it is necessary to look at the outcome of the bankruptcy proceedings and I am unaware of any studies of this nature for English country banks.

Furthermore, I am unaware of claims by contemporaries that losses were put to bank creditors. Although Calomiris and Haber state that "losses to English debt-holders were sometimes significant" and cite White (1985), they appear to be relying on a passage that quotes William Graham (1911) and mentions "heavy losses and bank failures."[80] Not only is it unclear whether "losses" refer to commercial or to bank losses, but when one looks at the original quote, one finds that the quote is not a reference to English banks: Graham is writing about a global credit crisis that started in 1836 and hit Ireland, the U.S., Belgium and France, in addition to England. In fact, the passage in Graham discusses in some detail the significant losses of the U.S. banking system, and the fact that British creditors bore more than 10% of those losses. Overall, it is clear that bank failures in England were viewed as a problem by contemporaries due to the disruption they caused to the money supply and to economic activity, but it is far from clear that there were significant losses that were born by bank creditors and not by the bankers themselves – who faced unlimited liability through much of the period.

Indeed, the losses incurred by the U.S. banking system after the crisis of 1837 appear to have instigated reform of the U.S. system of limited liability banking. The U.S. moved towards a system of double liability for bank shareholders as early as the 1840s and this system was adopted in the National Bank Act in 1863.[81] Double liability was eliminated only in the years after the FDIC was created in 1933. This is a clear indicator that in the 19th century U.S. legislators and depositors found that a simple limited liability regime resulted in a banking system that imposed an unacceptably large measure of losses on creditors.

In a book titled *Prometheus Shackled* Peter Temin and Hans-Joachim Voth argue that "English [banking] regulation before 1850, … almost surely got the trade-off [between financial stability and the growth-enhancing effects of financial intermediation] wrong, producing too little in terms of the growth-enhancing effects of finance without creating a particularly stable system."[82] Temin and Voth rely, however, on data from banks that catered to the wealthy, and do not collect data on the discount banks that were serving the commercial community. They appear to assume that the short-term credit that was provided in abundance by the "highly sophisticated" payments system was unimportant and could not play a role in English growth. They write that the Bank of England "focused almost exclusively on lending to the government,"[83] and thus treat the Bank's well-recognized role as the anchor of the discount banking system as something so trivial it can be ignored. They claim that "usury laws made it hard to lend to any but the most privileged groups," discounting entirely the fact that the average daily circulation of short-term bills in 1815 was £119 million, which was approximately equal to the best "guesstimates" of

[80] Calomiris & Haber, *supra* note 70, at 103; White, *supra* note 72, at 44; WILLIAM GRAHAM, THE ONE POUND NOTE IN THE HISTORY OF BANKING IN GREAT BRITAIN 202 (1911).

[81] N.Y. Const. of 1846, art. VIII § 7 (repealed 1935); Jonathan Macey & Geoffrey Miller, *Double Liability and Bank Shareholders*, 27 Wake Forest L. Rev. 31, 36-38 (1992); Joseph M. Leonard, Note, *Superadded Liability of Bank Stockholders,* 14 Temp. L.Q. 522 (1940).

[82] Temin and Voth, supra note 70, at 180.

[83] Id.

the total money supply at the time.[84] They find that wartime borrowing crowded out lending to the wealthy, but fail to observe the remarkable fact that the discount market – and commercial credit – functioned reasonably well throughout the war, although as noted above there was a significant increase in banking instability due the monetary effects of war finance. In short, the conclusion that the Financial Revolution played little role in English economic growth depends heavily on the assumption that the discount market and the short-term credit that it provided to commerce could not play a role in economic growth. One goal of this paper is to dispel that assumption.

In the next section I summarize my model of discount banking, and argue that in order to understand the process of growth, economists will have to analyze the institutions that support reputation-based debt. My theoretic work indicates that the late 18th c. English banking system may have been an engine for growth based on safe, short-term credit.

III. Comparing the 19th c. British Money Market to a Model of "Information Insensitive" Debt

In this section of the paper, I outline a theoretic framework that is consistent with the preceding discussion of the 19th c. British money market and compare it to the leading explanation that has been established subsequent to the 2008 financial crisis of the economic function of banks: Gary Gorton's theory of banks as issuers of "information insensitive" assets. The first subsection draws heavily from my theoretic work on the macroeconomic role of banks, where banks are modeled as resolving liquidity constraints by issuing unsecured, circulating debt and lending to the rest of the members of the economy. The second subsection discusses the Gorton and Ordonez model of information insensitive assets in detail, and shows that this approach is founded on the assumption that unsecured debt, such as that which circulated in 19th c. Britain, cannot be issued. The value of information insensitive assets is that they make it possible to borrow against valueless collateral as a substitute for the unsecured debt that is assumed away.

A. The Elements of a Model of the 19th c. British Money Market

The British money market was based on unsecured debt, repeated transactions between the same individuals, and long-term relationships. In short, it was built on what is commonly called "reputation" in economic models together with a limited legal right to the assets of the debtor and all guarantors of the debt in the event of a failure to pay. As Thornton observed, the institutional structure that had developed by the 19th century ensured that this unsecured debt was so safe that the costs of collateralizing it exceeded the benefits.

So what were the institutions in the 19th c. financial system that supported, not only the issue of high-quality unsecured debt, but also the expansion of the money supply by the circulation of the debt? The same institutions that were just discussed above: credit wrapping by every party who circulated the debt, the effective liability of bank shareowners, and the short-term nature of the debt. These same institutions also played an important role in generating liquidity crises.

[84] Pressnell at 172; Sylla, *supra* note 70, at 218; Forrest Capie, *Money and economic development in 18th century England*, in Exceptionalism and Industrialization 222-23 (Leandro Prados de la Escosura ed. 2004).

These three aspects of the 19th c. British money market, reputation, the circulation of unsecured debt that expands the money supply, and the liquidity problems associated with this circulating debt, will be discussed in turn below.

1. Reputation-Based Debt

The principle mechanism that is used to support the issue of unsecured debt in economic models is reputation: agents who default are unable to access debt markets in the future either temporarily or permanently. This mechanism is particularly effective in models of money where the circulation of debt increases the supply of money as well as financing the particular transaction that leads to the creation of the debt. For example, in a search model of money when bankers are defined to be those for whom any default will be made public, a banker earns rents from this special characteristic, and, if the banker is sufficiently patient, reputation will be enough to support repayment of the banker's debt.[85] And in a model where agents use credit to overcome liquidity constraints and are patient, a trusted banking system in which information about defaulting debtors is shared can make it possible for everyone in the economy to borrow.[86]

Reputation works as an enforcement mechanism, first, because the history of repayment can be verified, and, second, because the loss of access to credit in the future is a penalty that makes repayment today incentive compatible. It does not, however, provide a failsafe mechanism for those agents, in particular, for whom the incentive structure is not correctly calibrated to induce repayment. For example, reputation will clearly fail for those agents who learn before borrowing that they will die at the end of the day and therefore place no value on future access to credit.

On the other hand, where the incentive structure is correctly calibrated, it has a significant advantage over most other forms of credit: the borrower is incentivized to produce information about his ability to repay debt and to borrow only as much as he can repay.[87] When the costs of determining the value of the production opportunity are lower for the borrower than for the lender, as for example in environments with shirking, then this incentive structure will reduce the costs of information production.

For reasons that remain unclear, reputation-based lending and the institutions that support it have not been an important area of study for economists. The fact that the 19th c. British money market is an environment where unsecured debt was so safe that the costs of collateralizing it were viewed as clearly exceeding the benefits of doing so, is an indicator that this agenda merits further study. After all, Henry Thornton, the foremost monetary theorist of his time, attributed British economic growth to the fact that the banking system had turned the "appreciation of credit" into a highly developed "science." This appears to be a fairly direct claim that the development of institutions that support reputation-based lending had an important effect on Britain's economic performance.

---

[85] Ricardo Cavalcanti & Neil Wallace, *A Model of Private Bank-Note Issue*, 2 Rev. Econ. Dynamics 104 (1999).

[86] Carolyn Sissoko, *An Idealized View of Financial Intermediation*, Economics: The Open-Access, Open-Assessment E-Journal, 2007-5 (2007).

[87] The self-limiting nature of reputation-based credit is discussed in Sissoko, *supra* note 86.

2. Circulation of Credit

For Henry Thornton banking – and the expansion of the money supply through the circulation of credit instruments – is the link between reputation-based credit and economic performance. Thus, it is important that institutions developed not only to support reputation-based lending, but also circulating credit.

The key institutions have been identified above: everyone who circulated the debt wrapped it, bank shareowners were effectively liable for the debts they wrapped, and the debt itself was always short-term.

As noted above, reputation-based lending requires knowledge of the borrower's repayment history and an incentive structure that induces repayment. In the 19th c. the local banker typically had this information about the borrower, but others to whom the debt was transferred did not. Credit wrapping and bank-owner liability ensured that the bankers' incentives were to circulate debt that could be supported by reputation-based lending and that did not have the hallmarks of impending default. Furthermore, these two institutions also played an important role in the circulation of debt, since they enabled the buyer of the debt to focus only on whether or not reputation-based incentives would be sufficient to induce the direct seller of the debt to repay it, and the buyer did not need to make an inquiry into the incentives faced by the issuer of the debt.

Finally, when debt is short-term it is subject to much less deterioration of information than in the case of long-term debt. Thus, the short-term nature in the market meant that the situation of every borrower could be re-evaluated frequently and those who wrapped the debt did not have to worry that after the passage of several years' time, a decision to wrap debt that was well-founded when it was taken, would turn out to be costly.

3. Coordination-Liquidity Crises

Although credit wrapping, bank owner liability, and short-term debt supported reputation-based lending and the circulation of credit by ensuring the intrinsic quality of the circulating paper, they also tied the economy together with an interconnected web of debt that had to be rolled over in order for economic performance to be maintained. As a result, crises that were described as driven by liquidity problems above can also be viewed as coordination problems. When everyone was confident that the credit based monetary system was (temporarily) stable, debt was easily rolled over and the economy could hum along; but when doubts arose about the availability of credit, and as a result people feared that a reduction in credit would reduce economic performance and trigger defaults, potentially taking down the credit system itself, the doubts themselves could cause a reduction in credit.

Financial crises were called "panics" in 19th c. Britain, because there was an element of irrationality to them. The fundamental quality of the debt was such that creditors faced almost no risk of loss. Thus, panics were driven by the fact that a financial system built on short-term reputation-based lending can only function as long as people continue to expect it to function. As game theory predicts, beliefs matter.

Diamond and Dybvig (1983) showed that even when production is riskless, a financial system can have two equilibria, one of which is efficient, while in the other the financial system collapses. In this environment a lender of last resort can eliminate the bad equilibrium by eliminating the benefits of

participating in a panic.[88] In short, Henry Thornton's understanding of the role played by the Bank of England is consistent with modern economic theory: by expanding the money supply to prevent a panic from causing unnecessary failures, the central bank eliminates the benefit of being one of the first to recognize that the financial system is in trouble – and so the panic abates.

B. Compared to a Model of "Information Insensitive" Debt

An important competing strain of the literature views banks, not as intermediaries who use their own public reputations for repaying debt to extend reputation-based credit throughout the economy and circulate it, but as issuers of "information insensitive" assets. In this approach, crises do not represent coordination failures, but instead are triggered by the fact that the assets that circulated were always risky – it just happened that it was socially optimal for this risk to be ignored prior to the crisis. Thus a crisis occurs when agents become sensitive to the risk that was always present in the assets. Although the logic behind this approach is not immediately intuitive, Gorton and Ordonez have modeled this view of financial crises formally and here I will compare this model to the modeling elements described above.[89]

First, let me give a brief outline explaining why the authors find that it is optimal to ignore the riskiness of the collateralized loans in this paper and instead to treat them as "informationally insensitive." In Gorton-Ordonez half of the agents have capital and half of the agents have a risky, but positive-net-present-value production opportunity, so the agents with the production opportunity need to borrow to make use of the capital. As will be discussed in detail below, the authors assume that reputation-based lending is impossible, and focus on a second-best equilibrium that is attainable, because it so happens that every agent with this production opportunity also has a second production possibility via ownership of an asset that may produce consumption goods every period (e.g. a fruit tree) or may be an unproductive dud. Unlike the first production opportunity, this asset can collateralize a loan and is used to get financing for the first production opportunity. Complicating matters is the fact that determining whether the asset is a dud or not is costly, and the information degrades over time. (That is with some probability the asset is replaced by a new one.) The key to the authors' results is, thus, that this is a model where it is socially beneficial for agents with bad collateral to be able to borrow.

Gorton-Ordonez finds that even in a world where there is a social planner (who is also subject to the no-reputation-based-lending constraint), the costs of determining whether an asset is a dud means that, where the probability of a dud is sufficiently small, the social planner will choose not to engage in costly information gathering – even if this decision leads to rare crises. Furthermore, a high cost of information production is a desirable characteristic for collateral,[90] because the high cost makes it easier for agents with bad collateral to borrow.

1. Reputation-Based Debt is the First-Best Solution, But is Assumed Away

First, as is always the case in models of collateralized lending, if reputation-based debt could be supported, it has the potential to be a means of attaining the first-best solution, instead of restricting the economy to the second-best outcomes that can be achieved using collateral.

---

[88] Douglas Diamond & Philip Dybvig, *Bank Runs, Deposit Insurance, and Liquidity*, 91 J. Pol. Econ. 401, 417 (1983).
[89] Gary Gorton & Guillermo Ordonez, *Collateral Crises*, 104 Am. Econ. Rev. 343 (2014).
[90] Gorton & Ordonez, *supra* note 89, at 345.

As is common in the literature, in order to address the issue that is of interest to the authors, unsecured debt based on long-term relationships is simply assumed away. In Gorton-Ordonez, the agents are short-lived and matched randomly in each period, so there are no long-term relationships. Furthermore, in order to rule out the possibility of debt based on a contingent contract, the authors assume that the output of the first production opportunity is unverifiable for a lender.

The implication of these assumptions is clear, even in papers that model collateralized lending, the first best is unsecured debt, if a mechanism can be found to sustain it.

2. Debt is Based on Collateral and it is the Collateral that Circulates

In Gorton-Ordonez collateral, not reputation, makes lending possible. They find that it may be socially beneficial to lend to those with bad collateral, because it is the only way to fund the debtor's production opportunity. Furthermore in this model the debtor will repay the debt from the proceeds of this production opportunity even when the collateral posted is worth nothing and the debtor would default on unsecured debt.[91] Debt repayment takes place because the model is carefully structured so that there is no opportunity for the borrower to learn whether the collateral is good or bad before the debt must be repaid.[92]

In short, the welfare implications of the analysis – i.e. the fact that there are social benefits to "information insensitive" assets – depend not only on the use of collateral to fund an unrelated opportunity that cannot be financed on unsecured lending markets, but also on structural assumptions that ensure that even when the collateral is bad the debtor repays the debt. Both of these criteria may imply that the analysis is relevant only to a very narrow set of real-world situations.

In this paper the collateral itself, not the debt, circulates, making it possible for one generation after another of agents to borrow. The welfare benefits of the circulation of "information insensitive" collateral derive from the fact that agents who own bad collateral and would otherwise be credit constrained are able to borrow to finance their production opportunity.

3. Financial Crisis Takes Place When Those with Valueless Collateral Can't Borrow

A financial crisis takes place in the Gorton-Ordonez environment when an adverse shock triggers the expenditure of resources on learning the difference between good and bad collateral – and as a result all the agents with bad collateral face a binding credit constraint. More precisely, the dynamics of the economy are as follows: A credit boom takes place as information about assets deteriorates over time making it impossible to distinguish those with good from those with bad collateral and enabling the latter to borrow. When information deteriorates to the degree that information production is a "credible threat," however, such information production is averted by constraining borrowing and production – there is a credit crunch. Finally, when a shock takes place that results in information production, those with bad collateral – inefficiently – lose access to credit in a financial crisis.[93]

---

[91] Gorton & Ordonez, *supra* note 89, at 351 eq'n 3.
[92] Gorton & Ordonez, *supra* note 89, at 357-58.
[93] Gorton & Ordonez, *supra* note 89, at 368.

Gorton and Ordonez set out to show why assets that are generally of high quality, but also difficult to value individually, are desirable for use as collateral in a financial system. Their analysis focuses on the benefits of not having precise information about the value of collateral. For this reason, a financial crisis in their environment takes place when accurate information is revealed. While the Gorton-Ordonez view is well-argued and clearly presented, the environments in which it can be applied are, I would argue, very rare: it requires credit constraints that can only be addressed by collateralized lending, not reputation-based lending, and it requires debtors who do not repay unsecured debt, but do repay debt secured by valueless collateral.

The Gorton-Ordonez knowledge-revelation-causes-crisis approach is very different from the theoretic literature on coordination-based financial crises that builds on Diamond and Dybvig, and that is one of the elements of the analysis of the 19th c. British monetary system outlined above. The coordination-based crisis literature, however, addresses one of the key criteria motivating the Gorton-Ordonez model: "explaining a financial crisis requires the modeling discipline of fixing the shock size and showing how that shock can sometimes have no effect and sometimes lead to a crisis."[94] As Gorton and Ordonez observe, models of coordination problems such as Diamond and Dybvig typically have this property.[95] Indeed, in a paper modeling the 19th c. British monetary system along the lines outlined above, I wrote "It is possible that a single [bank] default will cause the whole system of intermediated credit to collapse – or that it will have no effect whatsoever. From a historical perspective, this range of possibilities may be appropriate when discussing financial markets."[96]

The lesson that I draw from this comparison is that reputation-based models of circulating credit are more likely to explain the historical record than models of collateralized lending. As I have argued repeatedly, if the relationship between banking and growth is to be explained and if financial crises are to be avoided the economics profession needs to start focusing its attention on the institutions that support reputation-based lending and how they can be modeled.

IV. What Does a Lender of Last Resort Do? A Historical Perspective

Part II of this paper discussed the banking revolution that took place in late 18th c. Britain and the role that banks play in the process of modern economic growth. Notably, the lender of last resort played a key role in this revolution by making it possible for the banking system to survive a financial crisis.

A review of the role played by the Bank of England in the 19th c. money market indicates that the core duty of a lender of last resort is the management of expectations about which firms will be allowed to fail, so that when a failure occurs, it generates only a liquidity crisis that can be addressed by expanding the money supply. Effectively, the primary duty of a lender of last resort is the careful management of the moral hazard that is created by its existence. This view is consistent with Bagehot's discussion of the lender of last resort – after one takes into account the historical context in which he was writing. In fact, the Bank of England's management of the failure of Overend, Gurney & Co. – by effectuating a policy that had been announced eight years earlier – has long been recognized as a master stroke that, some have argued, resulted in 100 years of stability for the British banking system. Because modern analyses, when

[94] Gorton & Ordonez, *supra* note 89, at 368.
[95] Gorton & Ordonez, *supra* note 89, at 348.
[96] Carolyn Sissoko, *An Idealized View of Financial Intermediation*, Economics: The Open-Access, Open-Assessment E-Journal, 2007-5 at 22 (2007).

they address the lender of last resort at all, have focused on the fact that it is a lender and ignored the duties created by the fact that it is also a last resort, they did not serve central banks in the 2007-2008 crisis well; as a result, expectations were poorly managed and expansion of the money supply was far from sufficient to address the resulting crisis.

A. A Lender of Last Resort Manages Moral Hazard

The term "lender of last resort" has its origins in Francis Baring's *Observations on the Establishment of the Bank of England* published in 1797. He referred to the Bank of England as the "dernier resort" or court of last appeal. The analogy is clear: just as a convicted man has no recourse after the court of last appeal has made its decision, so a bank has no recourse if the central bank decides that it is not worthy of credit. In short, the very concept of a "lender of last resort" embodies the idea that it is the central bank's job to determine which banks are sound and which banks are not – because liquidity is offered only to sound banks. And the central bank's determination that a financial institution is unsound has the same finality as a last court of appeal's upholding of a lower court's death sentence.

That the Bank of England was recognized as the final arbiter of credit in the British economy of the late 18th and 19th centuries is clear from the writings of contemporary authors. For example, in 1772 the Bank of England stopped discounting the bills of the Ayr Bank (including accepted bills), because the number of bills the Ayr Bank was circulating was large enough that it was almost certain that the bills were not all "real" -- that is, created only through the process of actual commercial trade. At the time a letter was published in the *London Chronicle* stating:

> You will find that the only cause of such bills as are good at bottom being refused by private bankers in London, is because the Bank of England will not discount them, and on that account such bankers cannot turn them into cash till due, be their necessity ever so great. For this and other obvious reasons, you will find it impossible to carry on your business as a banking company independent of the Bank of England, that being the great source of the British funds, and credit without whose countenance and occasional aid, no banker, nor merchant even in London can do business with safety and profit. [97]

The author of this letter makes it clear that a bank was solvent because the Bank of England stood behind it, and was insolvent if the Bank did not. In short, after the failure of the Ayr Bank, if not before, contemporary bankers understood that solvency was not an exogenous state, but for each bank depended upon the on-going support of the Bank of England.

In 1802, thirty years after the Ayr Bank collapse, Henry Thornton explained that one of the fundamental roles played by the Bank of England was to limit the amount of credit available to both London banks and country banks.

> While the transactions of the surrounding traders are thus subject to the view of the country banks, those of the country banks themselves come under the eye of their respective correspondents, the London bankers; and, in some measure, likewise, of the Bank of England. The Bank of England restricts, according to its discretion, the credit given to the London banker. Thus a system of checks is established, which, though certainly very imperfect, answers many important purposes, and, in particular, opposes many impediments to wild speculation.[98]

Thornton adds:

---

[97] London Chronicle, Sept. 15-17, 1772.

[98] Thornton, at 176.

> There seems to be a medium at which a public bank should aim in granting aid to inferior establishments, and which it must often find very difficult to be observed. The relief should neither be so prompt and liberal as to exempt those who misconduct their business from all the natural consequences of their fault, nor so scanty and slow as deeply to involve the general interests. These interests, nevertheless, are sure to be pleaded by every distressed person whose affairs are large, however indifferent or even ruinous may be their state.[99]

Thus, Thornton focuses attention on the key policy question faced by the Bank of England: When to withdraw its support from a bank or bill broker that is undermining the quality of origination practices in the market, despite the possibility that the decision could have an adverse effect on the broader market. In modern terminology, we would say that the Bank's task is to balance its duty to support the economy through a liquidity crisis with the moral hazard created by the knowledge that it is likely to do so.

Because the 19th British financial crises were true liquidity crises, the banks from which the Bank of England withdrew support could have continued in business certainly over the short-term and possibly for years in the absence of the denial of liquidity by the Bank of England.[100] For this reason, in 19th c. Britain the lender of last resort's determination that a bank was not worthy of its support was the direct cause of the bank's failure. In short, the term "lender of last resort" refers to the central bank's role of protecting the financial system by denying liquidity to those firms that don't meet the standards of the central bank.

Modern research confirms that the Bank of England actively and continuously monitored the individual banks and merchants that made use of its credit facilities. Flandreau and Ugolini find that the Bank of England limited moral hazard "by *not* lending 'anonymously'," but instead by carefully tracking both its exposure to each individual acceptor and discounter, and the degree to which each acceptor and discounter was extending – or overextending – credit to others.[101] This monitoring was accompanied by the rarely-used, but ever-present, threat of refusing liquidity to the acceptor or discounter by refusing to discount its paper.[102] Flandreau and Ugolini do not appear to realize, however, that similar policies were in place a full century before the time period that they study.

In early 19th c. Britain, the Bank of England was already recognized as the foundation upon which the banking system was built. It was recognized that for a partial reserve bank "solvency" is a state of affairs that exists only as long as the bank has access to central bank support and that a key function of the central bank was the use of the threat of withdrawal of such credit to maintain the quality of the money market.

B. Placing Bagehot's Lender of Last Resort in Context

Thus, in 19th c. Britain the lender of last resort was viewed as the final arbiter of whether any given member of the economy was sound enough to have access to credit.[103] The corollary of this policy was

---

[99] Thornton, at 188.

[100] Indeed, Bagehot remarks that Overend, Gurney & Co. might well have been able to continue in business if the original owners had managed to swallow the losses quietly. Bagehot, *supra* note 58, at ¶ X.11.

[101] Flandreau and Ugolini, at 23 (emphasis in original).

[102] Id.

[103] It is interesting to note that in 1910, Hartley Withers describes the Bank of England as the "final arbiter" which determines whether an acceptance house has been conducting its business to acceptable standards. Hartley Withers, The English Banking System 56 in National Monetary Commission Rep. Sen. 61st congress, Doc. 492.

that discounters who were deemed sound enough to receive credit from the Bank could draw on that line of credit as long as they had accepted bills to present – and in a crisis Bank lending skyrocketed.

Because the emphasis in the early 19th c. was on the "last resort" nature of central bank lending and on the management of moral hazard, the fact that approved entities would be supported by the Bank through a crisis was never a publicly-stated policy. After all, the Bank always reserved the right to choose not to lend to anyone. On the other hand, in the minutes of internal meetings of the directors, the Bank recognized a duty to use discount policy to support the country's commercial credit as early as 1810.[104]

Walter Bagehot in *Lombard Street* takes the Bank to task for its failure to acknowledge publicly that it was a lender of last resort. His criticism of the Bank's lending in crises after 1830 is, however, extremely modest,[105] and his principal concern is that support of the banking system is possible only if the Bank maintains a sufficient reserve of gold.[106] He explains that it is only very recently that the Bank has learned that it can both lend and maintain its reserves if it raises interest rates when it lends into a panic.[107] Thus, Bagehot's critique of the Bank's lender of last resort activities in 1825, 1847, and 1857 is based less on inadequate lending, than on a failure to maintain the gold reserve while lending.[108] In short, while Bagehot objects strongly to public statements denying that the Bank is a lender of last resort which has a duty to lend in a crisis,[109] he does not confuse such claims with an actual failure to lend in crises. Instead he argues that the best way to both preserve the gold reserve and lend freely in a crisis is to have a public policy of lending in crises.[110]

Thus, the context in which Bagehot framed the affirmative duty of the lender of last resort – to "lend freely [in a panic]. . . at very high rates . . . whenever the security is good"[111] – is one where the negative duty of the lender of last resort – to exclude from access to the Bank's lending facilities any institution whose conduct was undermining the quality of the money market – was generally accepted, and indeed differed from the behavior of any other bank only in the magnitude of its consequences. In fact, the

---

[104] Ian Duffy, *The Discount Policy of the Bank of England during the Suspension of Cash Payments, 1797-1821*, 35 Econ. Hist. Rev. 67, 76 (1982). *See also* Clapham, v. 2 at 15.

[105] Bagehot writes "though the Bank of England certainly do make great advances in time of panic, yet as they do not do so on any distinct principle, they naturally do it hesitatingly, reluctantly, and with misgiving," and continues to note that when "it was believed" that the Bank hesitated to lend on Consols, the panic grew worse. Bagehot, *supra* note 58, at II.60. He later argues, once again after noting that "It is certain that in all of these panics the Bank has made very large advances indeed," that the Bank should have a policy of lending against corporate bonds in addition to discounting bills and lending against government paper. VII.71, 73.

[106] Bagehot, *supra* note 58, at II.37.

[107] Bagehot, *supra* note 58, at II.37. To be more precise, Bagehot distinguishes between internal and external demands on the reserve. When the demand for money is purely domestic (that is, an internal drain), Bank of England lending may not deplete the reserve. II.45. However, when the demand is foreign (that is, an external drain), lending does deplete the reserve and thus demand must be quelled in part by raising the Bank Rate. II.36. Bagehot points out, however, that as a rule internal and external drains occur together; therefore it is necessary to raise Bank Rate, even as the increase in the domestic demand for money is met by lending readily. II.47.

[108] Bagehot, *supra* note 58, at VII.32.

[109] Bagehot, *supra* note 58, at VII.45 – 57.

[110] Bagehot, *supra* note 58, at VII.75 ("The public is never sure what policy will be adopted at the most important moment . . . The best palliative to a panic is a confidence in the adequate amount of the Bank reserve, and in the efficient use of that reserve. And until we have on this point a clear understanding with the Bank of England, both our liability to crises and our terror at crises will always be greater than they would otherwise be.").

[111] Bagehot, *supra* note 58, at II.41, 48.

qualification in Bagehot's prescription "whenever the security is good" refers to the negative duty of the Bank. In the 19th c. a "good" bill was determined by the quality of the acceptor and of the discounter. Thus, Bagehot understood his prescription to be entirely consistent with the Bank's rejection of paper on which one of the required guarantees was given by a bank that didn't meet the Bank's criteria. In fact, Bagehot's references to the failure of Overend, Gurney & Co., discussed below, demonstrate that he did not perceive a conflict between acting as a lender of last resort and denying liquidity to firms that are so badly run that they cannot give good security.

C. The Failure of Overend, Gurney & Co.

Overend, Gurney & Co. was the largest of the bill-brokers, which were a type of money market lender that maintained a lower capital ratio than traditional banks and therefore put extreme stress on the central bank in times of crisis.[112] The bill-brokers were similar to modern shadow banks, because banks used them to reduce their day-to-day needs for capital, and thus lower their costs. Like modern shadow banks, they rivaled and even exceeded traditional banks in size. Overend had ten times the assets of the largest traditional banks.[113]

After the bill-brokers borrowed more than the banks in the 1857 crisis, the Bank of England put in place in 1858 a policy of not standing ready to support the bill brokers in a crisis, but of lending to them on a "special" basis only.[114]

In 1866, the Bank of England refused to discount Overend's bills, and the huge firm collapsed, shaking the 19th c. banking system to its core. By providing liquidity freely throughout the crisis – including to other bill-brokers – the Bank allowed the rest of the financial system to weather the crisis.

Bagehot makes it clear that he approved of the Bank of England's decision not to support Overend, despite the fact that its failure caused a massive liquidity crisis, because of Overend's "bad business."[115] Indeed, Bagehot holds Overend out as an example of what a banker should not do.[116]

While Bagehot approved of the Bank's policy toward Overend, others have argued that Overend's failure was a watershed. According to Forrest Capie, the decision to let the largest of the shadow banks fail was "an important step on the road towards a sound policy towards financial crises," and it was this decision that led to "over 100 years of financial stability" in Britain.[117]

The Overend failure also offers an illustration of how carefully the Bank of England managed the moral hazard of the lender of last resort's role. By announcing that bill-brokers should "keep their own reserve,

---

[112] Bagehot, *supra* note 58, at XI.21 ff. *See also* Capie, *supra* note 52, at 16.
[113] Capie, *supra* note 52, at 16.
[114] Bagehot, *supra* note 58, at XI.30; Flandreau and Ugolini, *supra* note 63, at 11 n. 31.
[115] Bagehot, *supra* note 58, at VII.37, VIII.12, X.10.
[116] Bagehot, *supra* note 58, at VIII.12, X.10.
[117] Capie, *supra* note 52, at 16. Flandreau and Ugolini argue that the fact that the Bank lent freely to the remaining bill-brokers after allowing by far the largest of them to fail is evidence that the 1858 policy was simply rhetoric. Flandreau and Ugolini, at 14. However, if one views the Bank's purpose as being one of saving the money market, not destroying it, then one can easily conclude that by allowing Overend to fail, the Bank's goal was achieved and the subsequent aid to other bill-brokers was entirely consistent with that goal.

and not to be dependent on the Bank of England" in the year after the 1857 crisis,[118] the Bank carefully laid the groundwork for the failure of Overend eight years later.

D. 2007-08: The Crisis of Extraordinary Liquidity Support

When one understands that the term lender of "last resort" is a direct reference to the central bank's duty to manage the moral hazard of lending by limiting access to central bank liquidity to those institutions that are well-managed and will not undermine the quality of the money market, one recognizes that the financial crisis of 2007-08 was the culmination of a disastrous failure to manage expectations about access to central bank support.

The fault for this error, however, lies as much in the modern understanding of what a lender of last resort is, as in the conduct of the central bankers themselves. Because Bagehot wrote the seminal book on the lender of last resort, and focused on the "lender" part of the term, modern scholars have followed this lead. However, as Bagehot's treatment of the Overend failure makes clear, he simply assumed that it was obvious that a lender of last resort also manages the moral hazard of such lending by acting as a final tribunal, withdrawing credit from questionable institutions, and precipitating their failure. Thus, modern scholars have erred by failing to recognize that the words "last resort" refer to the management of the moral hazard of central bank lending by making the determination of which institutions are worthy of central bank support.

Oddly enough, some modern scholars have gone so far as to claim that the Bank of England never considered the moral hazard of its lending activities. For example, prominent shadow banking researchers have written: "the central banks of Bagehot's time . . . employed their balance sheets to stem the downturn . . . [but] did so without much prior theory about why it would work, and with hardly any thought about possible implications for more normal times."[119] This assertion stands in stark contrast to the fact, that from the mid-18th c. on the Bank of England played an important role in managing the quality of the bills that comprised the British money market, by monitoring the bills that passed through its hands and cutting off credit to institutions such as the Ayr Bank that were overissuing.

These researchers also view government bailouts to protect bank solvency as a necessary corollary to lender of last resort activities "in any real world crisis."[120] They are not alone. The claim that financial crises always require bailouts has become a commonplace.[121] Indeed the Gorton and Ordonez theory of "informationally insensitive" assets, that was critiqued above, is an effort to design an environment where bailouts would be efficient.

In short, modern scholars are largely ignorant of the fact that modern banking was born in an environment where bankers didn't receive bailouts, but instead saw their personal assets liquidated to honor the guarantees they made to creditors and where a lender of last resort was needed in order to make it incentive compatible for bankers to take on the significant personal risks of banking.

---

[118] Bagehot, *supra* note 58, at XI.30.

[119] Mehrling et al., *supra* note 8, at 1.

[120] Mehrling et al., *supra* note 8, at 14-15 ("In any real world crisis, of course, there are both liquidity and solvency elements at play, so liquidity backstop is insufficient. Just so, in the US crisis, there was the Treasury standing in the wings to provide capital as needed (e.g. TARP).").

[121] *See supra* note 6.

Is it any wonder that in such a confused intellectual environment, where lenders of last resort aren't expected to manage moral hazard, bank solvency is treated as an exogenous state, and financial crises are always accompanied by bailouts, central bankers mismanaged the financial system. For decades central bankers have replaced the warnings that remind the largest banks that their existence depends on central bank support with promises that liquidity will be forthcoming even when, indeed especially when, a large bank is mismanaged. This is the "too big to fail" problem that has been a matter of concern at least since the failure of Continental Illinois National Bank and Trust Company in 1984.

As a result of "too big to fail," in the financial world large banks are often expected to meet a weaker test for solvency than that faced by other firms: if a bank is cash flow positive and can become solvent in balance sheet terms with a few years of support from the central bank and bank regulators, then many argue that it should be deemed solvent and meritorious of central bank support.[122] Both credit rating agencies and bank auditors have been known to adopt this approach.[123] Critics have dubbed this "extend and pretend," but often acknowledge that given sufficient support, banks typically can earn their way to solvency.[124]

Central bank disregard of moral hazard was institutionalized by 1998 when Federal Reserve Chairman Alan Greenspan announced: "The management of systemic risk is properly the job of the central banks. Individual banks should not be required to hold capital against the possibility of overall financial breakdown. Indeed, central banks, by their existence, appropriately offer a form of catastrophe insurance to banks against such events."[125]

Just a decade later Greenspan's promise of catastrophe insurance was tested. Instead of allowing Bear Stearns, an investment bank, to be put through bankruptcy, the Federal Reserve in a bailout of Bear Stearns' creditors took on the risk of $29 billion of its assets. After the bailout of Bear Stearns, the Reserve Primary Fund more than doubled its exposure to commercial paper issued by Lehman Bros.,[126] even though Lehman's weakness relative to the remaining investment banks was widely recognized.[127] When Lehman Bros. was allowed to fail, this exposure caused the fund to "break-the-buck," which destabilized money markets entirely, and required not only heroic efforts on the part of the Federal Reserve and Treasury Department to prevent a financial collapse, but the passage by Congress of a bill authorizing a bailout.

By contrast, in 19th c. Britain, the Bank of England appeared to have a much better understanding of moral hazard and of the consequences both of lending and of the decision not to lend. The Bank of England supported the shadow banks through the crisis of 1857. After the crisis was over, the Bank put

---

[122] James Suroweicki, What does insolvency mean for a bank?, New Yorker Balance Sheet Blog, Jan. 20, 2009; John Hempton, Voodoo maths and dead banks, Bronte Capital Blog, Jan. 19, 2009; John Hempton, Bank solvency and the "Geithner Plan", Bronte Capital Blog, Feb. 16, 2009.

[123] Bank auditors in the U.K. based the passing grades they gave to the banks on the expectation of government support. Huw Jones, Lawmakers attack auditors over bank statements, Reuters, Nov. 23, 2010. William Dudley, Ending Too Big to Fail, speech, Nov. 8, 2013 ("the major rating agencies add an uplift to their credit ratings for the largest banks due to the prospect of government support")

[124] John Hussman, Extend and Pretend, Hussman Funds Weekly Market Comment, Apr. 12, 2010.

[125] Speech, Feb. 26, 1998.

[126] Bob Ivry, Mark Pittman and Christine Harper, Sleep-At-Night-Money Lost in Lehman Lesson Missing $63 Billion, Bloomberg, Sept. 8, 2009.

[127] Joe Bel Bruno, *Wall Street crisis of confidence stoked by Lehman*, Associated Press, June 10, 2008.

forward a clear policy of not supporting the shadow banks in the future. This not only gave the banking system time to adjust to the new policy, but indeed put a careful banker in the position of having to take the new policy into account. When Overend, Gurney & Co., the second-largest financial institution in England, failed, only a few small banks were so exposed to it that they too failed. Most of the banking industry had managed its exposure to the company.

The management of moral hazard is as important a function of the lender of last resort as the act of lending. This fact was obvious to early modern bankers, because their business was extending credit to borrowers, who, if left unsupervised, would have a tendency to borrow excessively. It should be equally obvious that the necessary consequence of using access to bank credit to keep poorly run firms alive will over time undermine and destabilize the money markets. Indeed, Bagehot warned explicitly that government support of a "bad bank" can have extremely adverse effects on the financial system, because such support "is the surest mode of preventing the establishment of a future good bank."[128]

Given these very basic facts, it seems likely that need for bailouts that is common to so many recent crises reflects the failure on the part of the central banks to manage the moral hazard of acting as a last resort lender, and the failure to protect the money supply by refusing liquidity support to those firms which are so poorly managed that they are introducing bad assets into the money supply.

It is clear that the current approach of using public funds to recapitalize insolvent banks was never entertained by either Thornton or Bagehot, who were both writing in an era in which bank owners bore the full costs of bad lending decisions. Furthermore, the theoretic foundations for the use of public funds to recapitalize insolvent banks are unclear. The logic of bank bailouts appears to progress from the fact that banks play an important role in economic activity and growth to the fact that nobody wants to see a decline in economic activity to the conclusion that the banks that we have must be protected from the consequences of their own decisions. In short, in modern markets large banks are often supported by central bank lenders who act in concert with government treasury departments and do not recognize that the qualifier "of last resort" was designed to impose a duty to evaluate and manage the moral hazard of such central bank activities. Instead in modern markets there is a strong tendency to argue that, since no partial reserve bank is solvent without access to central bank support, the test of bank solvency – at least as applied to large banks – should be whether the bank can become solvent given central bank support.

A true lender of *last resort* plays a key role in the incentive structure of an unsecured credit-based monetary system by withdrawing support from banks that "overissue" by lowering their standards for originating assets. Thus, a lender of last resort does not only avert coordination-liquidity crises, but also, and equally important, conditions the money market so that only high-quality assets back the money supply. It is the latter function that makes it possible for the central bank to provide only liquidity support to the economy and to avoid engaging in a bailout when it lends as a lender of last resort.

V. Shadow Banking and the Instability of the Collateralized Money Market

In the absence of a lender of last resort money markets have a long history not just of instability, but of instability that leads to the complete collapse of the private banking system. Examples of money markets that played an important role in international trade before they collapsed include those of Venice in the

[128] Bagehot, Lombard Street, at IV.4 ("The cardinal maxim is, that any aid to a present bad Bank is the surest mode of preventing the establishment of a future good Bank.")

14th, 15th, and 16th centuries and of Amsterdam in the 17th and 18th centuries.[129] British bankers were familiar with this history – and indeed witnessed the troubles faced by the Dutch banks with whom the larger British banks did business – so they were under no illusions as to the ability of the "market" to regulate itself. Thus, they accepted the Bank of England as the final arbiter of acceptable banking practices – even though there is evidence, especially in the early years of last resort lending, that the Bank's decisions were influenced by discriminatory prejudice.

Modern central banks, instead of recognizing that, as lender of last resort, they have the duty of setting the tone that determines the structure and stability of the money market, appear to believe that money markets can regulate themselves – despite the evidence of almost a millennium of history of money markets that collapsed in Europe.

The modern collateralized money market has grown up in this environment where central bankers do not understand that these markets cannot function over the long-term unless the moral hazard in them is actively managed. The failure to credibly commit to allowing institutions to fail has created a money market where low-quality assets can circulate and the actions that are needed to protect the money market from collapse are increasingly aggressive. In this section of the paper I respond to complacency about this evolution of the market and to proposals that the Federal Reserve's 2008 support of the dealer banks be institutionalized in a "dealer of last resort" policy.

The first subsection contrasts the 19th c. money markets discussed above with modern collateralized money markets. The second subsection relates the rise of wholesale funding and repo markets to the extraordinary liquidity support provided by the Federal Reserve, and questions whether these money market assets can play the same role in the economy as more traditional unsecured assets. The third subsection explains how these modern money market instruments are destabilizing. The fourth subsection discusses the dealer of last resort and why it can only support the collateralized money market by "becoming" the market for a prolonged period of time and the adverse effect such central bank behavior is likely to have on the realization of market prices. The final subsection explains that the functional differences between commercial banks and dealer banks justify supporting only commercial banks through crises.

A. Key Differences Between Modern Money Markets and the 19th c. British Money Market

The principal instruments used in modern money markets are commercial paper, asset-backed commercial paper and repurchase agreements, where ABCP and repos are collateralized money market instruments. Just like 19th c. bills these instruments generally are supported by bank guarantees. Unlike 19th c. markets, bank owners are not subject to capital calls and therefore not personally liable for any assets that go bad. Nor is the debt funded on modern money markets limited to short-term debt. The implications of each of these issues will be discussed in turn.

1. Banks Wrap Debt in Both Markets

As was discussed in Part I above, the distinction between "market-based" and "bank-based" lending is not particularly meaningful with respect to money markets, because most modern money market instruments carry a bank funding guarantee of some sort: the rating agencies usually require that non-financial

[129] Jan de Vries & Ad van der Woude, The First Modern Economy 154-156 (1997).

commercial paper and asset-backed commercial paper be supported either by a bank liquidity facility or by a maturity-matched repo with a bank; and tri-party repo is used by investment banks to borrow against the assets on their balance sheets and is also protected to a lesser degree by clearing bank liquidity guarantees.

In 19th c. Britain, it was also the case that money market instruments were liquid because they had been "accepted" – or in other words carried a (merchant) bank guarantee – and were therefore eligible for discount at the Bank of England. Modern non-financial commercial paper is almost perfectly analogous: it has a rating high enough to be sold to money market funds only when it carries both the borrower's guarantee and that of a bank. In addition, just as in Britain the bank that has guaranteed the commercial paper cannot borrow against it at the Federal Reserve's discount window – at least under normal circumstances.[130] Despite bank guarantees, ABCP and repo can also be distinguished from 19th c. money market instruments.

The concept behind the "securitization" of the assets that underlie an ABCP issue was to remove the assets from a bank's balance sheet via a "true" sale for accounting purposes. In some sense the very reason for the existence of ABCP conduits was to eliminate the exposure of the seller to the risk of the debt – and thus to avoid taking on liability for the debt at the time of the sale. While it is true that commercial paper issued against securitizations required a bank guarantee that was different only in form from a seller's guarantee, most of the assets issued against securitized assets as collateral were not money market instruments, but longer term instruments. Thus, to the degree that securitized assets were funded with medium and long term debt, banks could succeed in eliminating their exposure to the assets they sold.

The use of securitization to protect the banks from guaranteeing the quality of the assets that they issued can be contrasted with the practice in 19th c. money markets, where debt was transferred by endorsement and every seller was liable for the full value of the debt. In fact, the 17th c. establishment of the legal principles underlying endorsement has been designated a financial revolution, precisely because liability was retained even as the debt was transferred, enabling debt to circulate and augment the money supply.[131]

In short, the problem that incentives are not aligned when debt is sold and the seller's obligation to the buyer is extinguished at the time of sale rather than at the time of payment is as old as credit itself. The legal innovation of endorsement was designed to address this incentive problem, and as was explained in Part II, this innovation played a key role in the development of both modern banking and fiat money. Our early modern ancestors would have considered it obvious that allowing the originator of an asset to sell it without a guarantee would lead to the sale of low quality assets.

As was discussed above, even those debt-based financial systems where every creditor is sure to be made whole are subject to panics, or in theoretic terms subject to fear of a transition to a bad equilibrium where credit ceases to circulate. A lender of last resort can, however, allay the panic by expanding the money supply and thereby guaranteeing that the panic will not lead to massive illiquidity and the eventual

[130] See Greg Ip, *New York Fed Takes Steps to Bolster Credit Market*, Wall St. J., Aug. 27, 2007.

[131] Larry Neal, Rise of Financial Capitalism 5 (1990); Herman van der Wee, *Monetary, Credit and Banking Systems*, Cambridge Economic History of Europe vol. V at 322-32 (1977). In fact, the law of endorsement (or negotiable paper) is still in use today. See U.C.C. Art. 3.

liquidation of the bankers' personal assets. If, however, low-quality assets are circulating in the money market, the problem is not one of illiquidity, but of insolvency, and there is no reason to believe that the actions of a lender of last resort can assuage a panic – after all the low quality of the assets implies that the creditors must in the end take losses. In short, the legal principle of endorsement that was established in 17th c. Europe made it possible for first Dutch and then British money markets to be composed strictly of high-quality assets. Only in recent decades have bankers experimented with money market instruments, such as ABCP, that attempt to evade these protections. A 19th c. British banker would probably find it unsurprising that in such a market, where the legal mechanisms that align economic incentives are circumvented, defaults would end up destabilizing the money market.

In repo markets the "wrapping" of debt takes a different form: instead of guaranteeing the debt of, for example, a prime brokerage client, investment banks typically enter into back-to-back transactions. This means that when the prime brokerage client repos a bond with the investment bank, the investment bank immediately enters into a repo with the same terms on the tri-party repo market. Here, funding from the tri-party repo market is used to finance the prime brokerage clients' trade, but only the investment bank has an obligation to the tri-party lenders.

Thus, a debt guarantee is similar to a back-to-back transaction because both can be used by banks to make funding available to lower quality borrowers whom the bank is willing to support. The two also differ: where debt is guaranteed the lender has recourse against both the original borrower and the bank-guarantor; by contrast, in a back-to-back transaction the lender has recourse only against the bank-intermediary and cannot make a claim against the original borrower.

When both types of transactions are collateralized, the protection offered to the lender is incrementally weaker in the case of a back-to-back transaction, since recourse against the original borrower is an added protection for the lender who purchases guaranteed debt. There is another important property of back-to-back transactions: only the bank-intermediary knows which transactions are back to back and thereby intermediating non-financial demand for assets and which transactions form part of the bank's proprietary trades, where a bank is strategically taking on the risk of a position. By contrast, when debt is guaranteed, it is always clear that there is an underlying transaction that the bank-intermediary is helping to finance.

2. Bank Owners are not Liable for Bad Debt in Modern Markets

Another important factor aligning incentives in 19th c. British money markets was the fact that bank shareowners were effectively liable for unpaid bank debts due to the fact that shares were not fully paid up and the unpaid balance was subject to a capital call. It is worth noting that this was not a uniquely British practice. In the U.S. up until the Great Depression a system with comparable effects on economic incentives was in place: commercial bank shareowners faced double liability and therefore in a liquidation or receivership they could be called upon to pay the par value of the shares they owned a second time.[132] This regime was effective enough that even over the desperate years of bank failures in the U.S. from 1930 to 1934, less than one-twelfth of one percent of deposits in national banks were lost annually[133] --

[132] Jonathan Macey & Geoffrey Miller, *Double Liability and Bank Shareholders*, 27 Wake Forest L. Rev. 31 (1992). Macey and Miller observe that this regime had the beneficial effect of inducing shareholders to liquidate troubled banks voluntarily before they built up losses. From 1863 to 1912 more than four times as many banks liquidated voluntarily as involuntarily. *Id*. at 57-58.
[133] Jonathan Macey & Geoffrey Miller, *supra* note 132, at 59.

despite the fact that in, for example, 1931 6% of all national banks failed.[134] Observe also that the most important modern banks are investment-commercial bank hybrids, and in both Britain and the U.S. investment banks were unlimited liability partnerships up until recent decades. Goldman Sachs, the pre-eminent U.S. investment bank, for instance, only incorporated in 1999.

Although the U.S. like Britain has a history of bank shareowners facing capital calls, 21st century U.S. banks are corporations that are protected from shareowner liability by the same corporate law that protects every other corporation.[135] Once again, a simple analysis of economic incentives explains in part the poor quality of the assets underlying the money market leading up to the 2007-08 crisis: virtually everyone acknowledges that bank underwriting standards during this period can be described, at best, as "sloppy,"[136] and there are lingering suspicions that had regulators aggressively pursued evidence of fraud, convictions might have been forthcoming.[137]

3. Long-Term Debt is Funded in Modern Markets Creating Market Risk

In 19th c. Britain the Bank of England stood ready to discount accepted bills that had 95 days or less to run.[138] These, then, were the most liquid private sector money market instruments. While banks certainly invested in and presumably sometimes traded longer bills and securities, the core 19th c. money market assets were short-term private debt and government bonds.

Shadow banking, by contrast, has been defined as "money market funding of capital market lending,"[139] where capital markets are distinguished from money markets by the fact that they fund debt with a term in excess of one year. Unlike 19th c. Britain almost all of the assets funded on modern markets using ABCP or repos are capital market instruments.

Because long-term debt is being funded on short-term markets, when the short-term debt matures it must be rolled over or an alternate means of funding the long-term debt must be found. This maturity mismatch creates two forms of liquidity risk: funding risk and market risk. Funding risk is the risk that the borrower will not be considered creditworthy when the debt matures and is unable to retire or to roll over the debt. Market risk is the risk that the value of the long-term debt that is being used as collateral falls, and as a result, the lender is no longer fully secured and the borrower no longer has enough collateral to borrow on ABCP or repo markets.

Observe that funding risk exists when either short-term or long-term assets are being financed on the money market. The traditional means of dealing with funding risk was to require that every seller of the debt wrap it. Bank liquidity guarantees play a similar role in ABCP markets – as did the intraday credit

---

[134] David Wheelock, *Regulation, Market Structure, and the Bank Failures of the Great Depression*, Fed. Res. Bank of St. Louis Rev. 27, 31 (Mar-Apr 1995).

[135] Note, however, that modern commercial banks are not subject to the general corporate law governing liquidations, reorganizations and the distribution of remaining assets to shareowners, but to a regime in which regulators like the FDIC play a much greater role.

[136] Mehrling et al., *supra* note 8, at 13 n. 3.

[137] Jed Rakoff, *The Financial Crisis: Why Have No High-Level Executives Been Prosecuted?*, N.Y. Rev. of Books, Jan. 9, 2014.

[138] *See supra* note 66 and accompanying text.

[139] Mehrling et al., *supra* note 8, at 2.

provided by the tri-party clearing banks to the repo market up until recent months.[140] Furthermore, when an investment bank enters into a back-to-back transaction with a prime brokerage client on one side and the tri-party repo market on the other, the bank effectively substitutes the bank's credit for that of the client on the tri-party repo market.

It is worth emphasizing here a point made in Part I: because the finance of longer-term assets requires that these short-term instruments be rolled over, funding risk is always a concern in the so-called "market-based" short-term credit system, and this almost certainly means that this "market-based" credit system cannot exist except when it is backstopped by the banking system. Thus, what is commonly known as the "market based" short-term credit system – including most of the shadow banking system – should properly be understood to lie within the "bank-based" credit system.

Unlike funding risk, market risk is very important only in modern collateralized money markets. Because 19th c. money market loans were only due at maturity and were not secured, market risk was not a factor that could itself cause default. (Of course, a collapse in the value of inventory could still cause a business to fail and then default on its unsecured debt – but these were knock-on effects, not direct effects.) By contrast, when the market value of the collateral underlying a repo falls, the lender risks holding a loan that is not fully secured. Whereas some collateralized markets such a mortgage markets place this market risk on the lender, in repo markets, market risk is borne by the borrower, who faces a "margin call" and is required to post additional collateral within one or at most two days.[141] A borrower who does not have additional collateral available, may have to sell assets to get enough cash to pay off the difference between the value of the collateral and the debt. (See the example in Part I.C.) If a borrower fails to meet a margin call, the repo lender has the right to sell the collateral and keep an unsecured claim against the borrower for any remaining value of the loan. Similarly if the value of the collateral underlying ABCP falls too low, the lenders have the right to liquidate the collateral.

In short, modern money markets face market risk in addition to the funding risk common to all debt markets, because long-term assets are used as collateral to support short-term loans and the loans can be withdrawn due to inadequate collateral. The fact that an additional risk needs to be managed in modern markets contributes to the comparative lack of safety of such markets.

B. Extraordinary Liquidity Support and the Rise of Wholesale Funding and Repo Markets

In Part IV, I demonstrated that the management of moral hazard was a core function of a traditional lender of last resort, and that modern scholars and policy-makers have failed to appreciate this. As a result the duty of a lender of last resort has morphed in modern times to encompass the provision of liquidity support to any large bank that will be solvent, given such liquidity support.

Modern banks monetize this expansive central bank liquidity guarantee, not only by borrowing directly on financial markets, but also by using collateral to transform credit risk into liquidity risk.

---

[140] Observe that the finance of agency mortgage-backed securities on short-term markets also followed this model: Fannie Mae and Freddie Mac guaranteed all of their MBS issues. Paul Davies, Ousted Fannie Mae CFO Tells His Side of the Story, American Banker, Nov. 21, 2013.

[141] Int'l Capital Market Ass'n European Repo Council, Repo Margining Best Practices 2012 at 9 (2012).

1. Wholesale Funding Markets

First, observe that in this environment where large banks have access to extraordinary liquidity support, they are also able to borrow directly on financial markets, or in other words they have access to wholesale funding. Over $500 billion is currently being raised in the form of unsecured financial commercial paper. Because the credit rating agencies have a stated policy of increasing credit ratings on the basis of likely government support,[142] and because a bank can only issue commercial paper without a third-party liquidity guarantee if it has the highest credit rating, the extraordinary liquidity support offered to the largest banks almost certainly plays an important role in their ability to borrow on wholesale markets.

[142] William Dudley, Ending Too Big to Fail, speech, Nov. 8, 2013.

2. Repo Markets Transform the Lender's Credit Risk into the Borrower's Funding Risk

19th c. money markets were unsecured and, thus, focused on managing credit risk. This is illustrated both by the fact that every bill discounted by the Bank of England was guaranteed to be paid in full by at least three different parties, the issuer, the acceptor, and the discounter, and by the fact that the Bank had negligible losses on its discount portfolio, even after a crisis.[143] Managing credit risk also required management of the moral hazard of lending, and in 19th c. Britain a basic criterion for access to the discount window was that the bank itself was a good credit risk.

In modern markets, our largest banks have access to the discount window even when they are not good credit risks. Because the Federal Reserve offers extraordinary liquidity support to the largest banks, it should come as no surprise that modern markets transform credit risk into liquidity risk. A mechanism by which this takes place is the collateralization of lending.

This can be seen in the ABCP market in 2007 when credit concerns about the underlying collateral prevented conduits from rolling over their commercial paper and as a result bank liquidity facilities were called to provide funding for the commercial paper. In the repo market, collateral provides additional security to the lender, and the credit risk that lenders traditionally face is replaced by market risk or the risk that when the collateral falls in value the borrower does not post additional collateral and lender is not able to sell it fast enough to cover the value of the loan.

From the borrower's point of view, on the other hand, because collateral reduces credit risk, creditworthiness becomes less important and borrowing can be relatively inexpensive based on the quality of the collateral. The cost of repo borrowing is, however, that the borrower now must worry not only about having the resources to pay the debt at maturity, but also about maintaining sufficient collateral to back the loan throughout the life of the loan. For the borrower, market risk is experienced as funding risk – the borrower faces the risk of having to find liquid funds or post additional collateral in the event that the market price of the asset falls.

Note something odd about this system. Presumably a collateralized borrower is in need of liquid funds – that is, after all why most people borrow. On the other hand, the structure of repo is such that the borrower is at risk of significant losses if market risk is realized and the borrower, who has already exhibited a need for liquid funds, does not have access either to liquid funds or to collateral. In short, the only people for whom repo borrowing is likely to make sense as a matter of regular practice are those who cannot borrow unsecured, those who are collateral-rich and are seeking to monetize their collateral, and those who expect to have access to outside support in the event that they are liquidity constrained.[144]

Some researchers have argued that the growth of the collateralized money market is a demand-side phenomenon.[145] Zoltan Pozsar has claimed that the largest lenders in modern markets, asset managers such as mutual and pension funds, are reluctant to extend unsecured credit to the banks in the form of

[143] Vincent Bignon, Marc Flandreau, & Stefano Ugolini, Bagehot for beginners: the making of lender-of-last-resort operations in the mid-nineteenth century, 65 Econ. Hist. Rev. 580, 602 (2012).
[144] The connection between the latter group and the "dealer of last resort" proposal will be addressed below.
[145] Mehrling et al., *supra* note 8, at 12.

uninsured deposits and prefer to lend via repos or asset backed commercial paper.[146] This point should not, however, be overemphasized, because banks are able to raise significant funds, unsecured, by issuing commercial paper. Perry Mehrling has argued that the growth of the collateralized money market is due to the modern asset management practice of using derivatives markets to take on the risks of investing while holding invested funds in monetary assets.

While demand may have played a role in the development of the collateralized money market, the advantages of the market for borrowers presumably also played a role in its growth. For investment banks the market is convenient for financing inventories, and possessed of the useful property that in normal times the market is not very sensitive to the credit quality of the borrower.[147] Thus, as was discussed in Part I, the collateralized money market plays an important role in the finance of investment banks and this fact is likely a significant factor in the development of the market. Subsequent to the support for investment banks and collateral prices that was extended to the repo market in 2008, it is also possible that there is now an expectation of liquidity support that makes the market even more attractive.[148]

3. Can Collateralized Money Markets Support Economic Growth?

These shifts in the structure of the money market have consequences: instability is discussed below. Here, I address the implications of these changes in the money market for economic growth.

As was discussed in Part II, early monetary theorists such as Henry Thornton believed that banks contributed to economic growth because they allowed the money supply to expand based on the needs of the economy and because they employed the "science" of credit to issue high-quality unsecured credit abundantly. Because lenders on collateralized money markets avoid credit risk and focus their attention on market risk, techniques for issuing unsecured non-bank money market instruments may fall into disuse and, as a result, the ability to borrow unsecured may be more limited than the principles of managing credit risk would require. Indeed, arguably unsecured credit is already less available than it was in centuries past, when businessmen in international centers of trade could borrow unsecured at annual rates of 3-5%.[149]

If unsecured credit falls into disuse, it is not clear that the money supply will be able to grow with the needs of the economy. Whereas traditional money markets supported economic growth through the issue of high-quality unsecured debt, modern collateralized money markets – to the degree that they fund private sector assets at all – fund instruments that trade actively and have meaningful market prices, although they may be risky assets such as equities. Modern money markets extend credit to corporations

---

[146] Zoltan Pozsar, Institutional Cash Pools and the Triffin Dilemma of the U.S. Banking System 10-11 (IMF Working Paper No. 11/190, Aug. 2011).

[147] Fitch Ratings, Repo Emerges from the "Shadow" 3 (Feb. 3, 2012).

[148] *See, e.g.* Letter from Int'l Capital Market Ass'n European Repo Council in response to FSB Shadow Banking Workstream – Interim Report on Securities Lending and Repos, Annex, (May 22, 2012) ("The question is how to mitigate such systemic liquidity risk [from a fire sale of collateral assets]. We believe that systemic risks require systemic responses. In this case, the authorities can be expected to intervene as lenders of last resort to ensure the liquidity of the system as a whole. For their part, market users should be expected to remain creditworthy and to have liquidity buffers sufficient to sustain themselves until official intervention restores sufficient liquidity to obviate the need for fire sales.")

[149] For Amsterdam, see Violet Barbour, Capitalism in Amsterdam in the 17th Century 85 (1963); Herbert Bloom, The Economic Activities of the Jews of Amsterdam in the 17th and 18th Centuries 192 (1937). Usury laws in England prevented the rates on bills from rising above 5% throughout the 18th c.

and other established entities, but are not designed to finance the individual tradesmen and small businesses that were financed by traditional banks and money markets.

To the degree that modern money markets are no longer able to expand the money supply to meet the needs of individuals and small businesses, but instead are constrained to rely on collateral issued by governments and large companies, they may fail to meet the needs of the economy and to support economic growth. In short, it is not clear that collateralized money market instruments can play the same role in expanding the money supply and in economic growth as that played by unsecured money market instruments.

C. How Wholesale Funding and Repo Reduce the Stability of the Money Market

Many view this transformation of the money market from one in which credit risk was carefully managed to one where market and liquidity risk substitute for credit risk as a benign, if not beneficial, development.[150] There are, however, many concerns that this development should raise.

Most importantly, collateralization is destabilizing because it drains liquidity when liquidity is most needed. This is addressed in subsection three, after discussing the instability of commercial paper-based wholesale funding and the effects of collateralization on credit quality.

1. The Absence of a Non-Bank Guarantor Makes Wholesale Funding Inherently Unstable

The dangers of bank reliance on wholesale funding were made clear in the 2007-08 crisis. When Lehman Brothers failed, the collapse in value of its commercial paper caused a money market fund to "break the buck," and triggered a mass exodus from non-government-backed commercial paper and extraordinary danger for financial markets. One way to greatly reduce the risks of such events is to require that financial institutions only raise money on commercial paper markets against paper that is also guaranteed by a non-financial issuer. After all, if Lehman Brothers had backup guarantees on its paper, it is unlikely that the paper could have fallen in value sufficiently to cause the Reserve Fund to break the buck.

In short, the danger that can be cause by a large bank borrowing only on its own credit has been recognized for centuries. In 19th c. Britain when a bank turned to such direct funding – as opposed to the sale of real economy assets that were wrapped by the bank – it was viewed as a sign that the bank was in difficulty. And it was widely recognized that banks that were on the edge of failure usually ended up relying heavily on such funding.[151]

2. Collateralization May Both Support and Aggravate a Decline in Bank Credit Quality

Collateralized funding markets are no solution to the instability of unsecured wholesale funding markets. The transition to collateralized money markets may both reflect declining credit quality among financial institutions that makes it difficult for a growing number of them to borrow on an unsecured basis, and at the same time enable these financial institutions to continue operating despite the decline in their credit quality. The fact that this change took place alongside the transformation of the investment banking industry from unlimited liability partnerships to limited liability corporations is probably an indicator that

[150] Mehrling et al., *supra* note 8.

[151] *See, e.g.,* W.T.C. King, History of the London Discount Market 247-50 (1972) (Overend, Gurney & Co.'s reliance on accommodation paper before it failed).

declining credit quality is an important driving force behind this change. After all, centuries of evidence demonstrate that the partners in the largest investment banks were generally very careful not to blow their firms up, whereas over recent years we have seen that the managers of investment banking corporations are, if not less careful, at least less subject to discipline by their colleagues.

A second concern is that the movement to collateralized short-term lending aggravates declining credit quality among financial institutions. Research has shown that repo lending terms are principally determined by the quality of the collateral posted and do not tend to reflect incremental changes in the credit quality of the borrower.[152] For this reason, it is possible that the movement towards collateralized borrowing makes borrowers less concerned about whether or not they are viewed as good quality borrowers.

3. Collateralization Drains Liquidity When It Is Most Needed, Destabilizing the Money Market

Most importantly, however, collateralization drains liquidity when it is most needed. Repo markets and the collateralized derivatives markets that are closely related to them are designed to protect the lender from credit risk by imposing liquidity risk on the borrower. Thus, it should come as no surprise that, in a crisis, these forms of lending increase the need for liquidity in a way that unsecured lending does not.

The demand for liquidity arises, because a fall in the market value of collateral leads to an immediate demand for more collateral – or margin call – from the borrower, who must either have liquid funds or excess collateral on hand in order to meet the demand with ease. If the borrower is liquidity constrained, the borrower will have to reduce his loan to one that the collateral can support – this may require the sale of unencumbered assets to pay off part of the loan. A "liquidity spiral" can take place when a large financial institution finds that the sale of assets drives the prices of the existing collateral down and, as a result, the institution faces yet another margin call. When a borrower fails to meet a margin call, it is generally in the lender's interests to sell the collateral as quickly as possible in order to avoid the possibility that the collateral's value deteriorates further and less of the loan's value is covered. The lender's sale can then have an adverse effect on the value of all assets similar to the collateral.

Sales that are forced by margin calls whether executed by the borrower or the lender, by pushing prices down, can reduce the value of the collateral posted by many different parties and have the effect of generating a large number of new margin calls. In short, when lending is collateralized a small decline in prices can result in a margin call on a large market participant. If this participant is illiquid and forced to sell assets, the sale itself can trigger new margin calls – and sales – and the whole collateralized market can end up experiencing a liquidity spiral. For this reason, it is clear that collateralized lending drains liquidity when it is most needed – that is, when a large financial institution is failing.

Contrast this environment with the 19th c. British money market. In unsecured money markets, a lender can demand payment only at maturity. Panics take place when lenders choose not to roll debt over, but instead seek payment of matured debts from borrowers. It was in this environment that the lender of last resort's guarantee that borrowers had a place to sell the unsecured assets on their balance sheets at a high, but not extortionate, discount rate was sufficient to quell the panic. Note the dynamics that take place in this action: the supply of high-powered money is increased by the central bank, transferred first to the borrowers who discount their assets, and then to the lenders, who now hold the most liquid of assets and

[152] Fitch Ratings, Repo Emerges from the "Shadow" 3 (Feb. 3, 2012).

can use them to meet any demands that are made on them. Observe that the high quality of the unsecured assets that circulated in the money market were what enabled this system to work: the Bank of England checked that the bills that were discounted had the right guarantees and were of the right maturity and didn't need to inquire further into their value. Observe, also, that because transfers took the form of unsecured, wrapped sales, lenders could only demand payment at maturity. As a result, the unsecured nature of the market limited the liquidity demands that could be made in a panic, and once lenders were confident that money supply would expand to meet the demand for liquidity there was no reason for the panic to continue.[153]

The demands for liquidity that can be made in collateralized money markets are much more sudden. Not only must collateral calls in the tri-party repo market be met, as a rule, within one day, but all borrowers who are using the collateral that has fallen in price – even those whose debt does not mature for months – will face a call. Furthermore, the fact that the collateral on repo markets is funded on a leveraged basis means that small changes in the market prices of assets can result in the need to sell off a large fraction of assets. In the example discussed in Part I, for the fully leveraged borrower a 2% decline in the price of collateral subject to a 5% haircut can force a sale of 42% of the assets posted as collateral. Such sales can potentially trigger a liquidity spiral if prices fall further and many others are forced to sell assets as well.

Observe that, in addition, the problem of leverage is not solved by requiring that only high-quality assets be posted as collateral. Treasury bonds are the highest quality collateral, but the duration of a 30 year bond means that its price can fall dramatically in response to a small interest rate increase.

To summarize, modern collateralized money markets are likely to face more sudden demands for liquidity than traditional unsecured money markets because (i) margin calls are made on a daily basis, not when the debt matures; (ii) the leverage inherent in collateralized borrowing can force borrowers to sell collateral in order to pay off loans aggravating the price decline; and (iii) the reliance of the market upon collateral means that when price falls all lenders make margin calls simultaneously. Unlike traditional unsecured money markets where it took many individual decisions by independent lenders to create a panic, liquidity crises in collateralized markets are an almost mechanical function of declines in the price of collateral – that are exacerbated by the market's reaction to the decline in price.

D. What Should a Lender of Last Resort Do?

This subpart evaluates proposals to deal with the price fluctuations that are created by the collateralized money by expanding the duties of the lender of last resort. I explain that, even though these policies are often framed as seeking to support "markets," their effect is to expand access to central bank liquidity facilities to the dealer banks. I also explain why the central bank is unlikely to be able to stabilize the collateralized money market without extraordinary intervention in the market in every crisis, and why the central bank will be incentivized to distort asset prices upwards in this regime. Finally, I discuss the difference between dealer banks and commercial banks and why asset markets are likely to be adversely affected by the provision of lender-of-last-resort-like liquidity to dealers.

[153] Indeed, the announcement that support would be extended to the market was often enough to stop the panic.

1. Distinguishing Crises on Collateralized Money Markets from Traditional Crises

In 2008, the Federal Reserve stepped in to prevent the forced sales that would have been triggered in the collateralized money markets when Bear Stearns and Lehman Brothers failed by providing loans to dealer banks against all collateral that had been acceptable on the tri-party repo market. This action can be distinguished from the traditional lender of last resort action because:

(i) The lending took the form, not of wrapped sales, but of collateralized loans, where the lender has the right to demand additional collateral on a daily basis.

(ii) The central bank did not lend only against high-quality assets, but on many days accepted tens of billions of dollars of equities, junk and unrated securities from each of a variety of primary dealer borrowers.[154]

(iii) In collateralized money markets the temporary provision of cash is not sufficient to quell a crisis.

To explain this latter point, first, let us consider a typical 19th c. British lender of last resort operation and then compare it to the 2008 action. The Bank of England purchased short-term unsecured debt that carried a guarantee of payment from the issuer, the acceptor, and the discounter. These parties were all insulated from liquidity demands, because none of them faced any obligation to make a payment on this debt until it matured. (It was, however, short-term debt, so they were insulated only for a matter of months.) In short, because of the nature of the money market instruments there were structural limitations on the amount of liquidity that could be demanded at any given time.

The fact that money market assets were of extremely high quality also contributed to the success of lender of last resort operations. Liquidity panics were literally panics – as long as there was no switch to an equilibrium where the credit system collapsed, losses were minimal. Only those bankers that had mismanaged their lending, or whose debt was not supported by the Bank of England, would experience losses. Their creditors, and in particular the non-bank business community could expect to receive payment on all the bank-guaranteed debt that they held.

When growing distrust in the credit system raises the possibility of a switch to the bad equilibrium, the monetary injections of the lender of last resort support the good equilibrium and reverses the growth of distrust by showing that there is a mechanism to keep the switch to the bad equilibrium from happening over the near horizon. This trick is, however, likely to work only if the creditors are actually going to be paid. After all, if the creditors don't get paid, the good credit equilibrium is no longer rational. What lender of last resort actions did in 19th c. Britain was avert the liquidation of bank owners' assets to pay creditors, making continued offering of banking services together with the bankers' liability for bad debts incentive compatible.

Now consider the case of modern money markets which fund long-term assets. In order to fund long-term assets on the money markets, the assets must be used as the collateral that backs short-term liabilities (since a sale, whether wrapped or unwrapped, can't turn a long-term asset into a short-term asset). Long-

[154] Data on the Primary Dealer Credit Facility and the Term Securities Lending Facility are available here http://www.federalreserve.gov/newsevents/reform_pdcf.htm and here http://www.federalreserve.gov/newsevents/reform_tslf.htm. See, e.g. October 10, 2008.

term assets are, however, also subject to significant price volatility when either interest rates or expectations about the future change.

Consider the case where an unexpected rise in interest rates leads to a fall in the value of Treasuries. Those who have posted long Treasuries as collateral own collateral that is worth less, and there will be a flurry of margin calls. Because many borrowers on these markets do not have the ability to borrow unsecured,[155] a decline in the value of collateral can have the direct effect of limiting access to credit. Some of the borrowers will not have managed their interest rate risk well and will be forced to sell. Dealers recognizing that the sales are forced – and uncertain of how many more such sales loom on the horizon – are unlikely to give these sellers good prices. Due to the structure of the collateralized money market it seems almost certain that the collateral price decline will overshoot.

The credit crunch will become a liquidity crisis whenever margin calls trigger enough asset sales to drive prices down even farther, generating a liquidity spiral. Such asset sales can take place either because a single large borrower is forced to sell, or many small borrowers are simultaneously forced sell. Note, however, that the fact that a margin call necessarily follows a price decline on these markets means that the coordination of simultaneous small borrower sales is built into the market structure.

In short, it is far from clear that the stability of a money market collateralized with long-term assets and remargined daily is consistent with price realization in such a market. Even the safest long-term debt loses value when interest rates rise, and that price change can easily trigger a liquidity spiral.

Furthermore, it is far from clear that a lender of last resort can solve this problem. When interest rates rise or expectations about the future fall, the value of existing long-term collateral falls even though the debt that it is backing does not. Thus in order for previous levels of credit to continue to be available to financial institutions on the collateralized money market, either the central bank must continue to provide cash over a much longer horizon than the traditional lender of last resort – potentially for the full life of the long-term assets, or it may be necessary to support asset prices without regard to fundamental value.

In short, modern collateralized markets are designed to restrict credit and trigger forced sales whenever collateral prices fall and, as a result, they are almost certainly less stable sources of funding for financial institutions than deposits. Not only are they structurally unstable, but they also have both the danger of the lender "runs" that are common in unsecured money markets, and the additional danger of inter-dealer and borrower "runs," as was discussed in detail in Part I. Overall, it is distinctly odd that regulators have permitted commercial bank lending to be disintermediated by investment banks borrowing on collateralized money markets that are far more unstable than deposits.

2. What Does a Dealer or Market Maker of Last Resort Do?

Three policy options are generally considered when discussing dealer or market maker of last resort activities:

(i) Purchase wrapped securities. A lender of last resort can purchase securities subject to the guarantee of the seller to make good on any debt that defaults. In 19th c. Britain the Bank of England's discount of commercial bills took this form.

[155] Bank of America Corp., Citigroup, and Morgan Stanley, for example, currently have second-tier short-term credit ratings.

(ii) Purchase securities outright. Recent proposals that the lender of last resort should become the "dealer of last resort" have sometimes included recommendations that the central bank support bank balance sheets by taking on the risk of some assets without the seller's guarantee.[156] The obvious dangers of this policy means that the recommendation is often restricted to the purchase of "prime" securities.[157] Of course, when implementing monetary policy – but not when acting as lender of last resort – central banks have long purchased government and quasi-government assets.

(iii) Lend against the collateral of posted securities. When the securities are long-term, central banks have not traditionally purchased them, but instead made collateralized loans against them. In the 19th c. the Bank of England's "advances" took this form, but were only available to a very select group of borrowers.[158] Two important policy questions arise in the case of collateralized loans: whether or not to remargin the loan and how frequently to remargin the loan.[159]

Because most of the policy proposals for a dealer or market maker of last resort focus on the last form of support for the market, collateralized lending, my discussion here will focus on it too.[160]

a. A Dealer of Last Resort Expands Who Gets Access to the Central Bank

The first thing to observe about the dealer of last resort policy proposals is that they are usually framed as changing the type of collateral that is accepted for emergency lending by the central bank. This is misleading, however. In the U.S. the Federal Reserve has been able to accept as discount window collateral virtually any asset since the 1930s – as long as it was collateral for a loan to a commercial bank.[161] The problem in 2008 was not the nature of the collateral that could be used, but the fact that investment banks didn't have access to the central bank. Thus, the key innovation of the "dealer of last resort" proposal is not the type of collateral that can be used for a loan, but the fact that dealers – or investment banks – are able to borrow from the central bank using that collateral.

A dealer of last resort, however, is not generally presented as supporting the dealer system through a crisis, but as supporting the liquidity of asset markets. For example, when motivating his "market maker of last resort" proposal, Willem Buiter writes:

> Market illiquidity is a form of market failure. Liquidity can be provided privately, by banks and other economic agents holding large amounts of inherently liquid assets (like central bank reserves or TBs). That would, however, be socially and privately inefficient. Maturity transformation and liquidity transformation are essential functions of financial intermediaries. Private financial entities should hold (or have access to, through credit lines, swaps etc.) enough liquidity to manage their business during normal times, that is,

[156] *See, e.g.,* Willem Buiter, *Central Banks and Financial Crises* 31 (2008); Mehrling et al., *supra* note 8, at 9.
[157] Mehrling et al., *supra* note 8, at 9.
[158] Flandreau & Ugolini, *supra* note 44, at 7, 31.
[159] Note that Willem Buiter, the originator of the term market maker of last resort, does not list these issues as ones that need to be addressed when the central bank establishes its market maker of last resort facility. Buiter, *supra* note 156, at 32. Presumably he assumes daily margining of repos, as takes place in private sector markets.
[160] Buiter observes that the reason he employs the term "market maker of last resort" and not "buyer of last resort" is that he expects most of the action to take the form of repurchase agreements not outright purchases. Buiter, *supra* note 156, at 31 n. 9. *See also* Mehrling et al., *supra* note 8, at 9.
[161] Critics complained that "any cat and dog" could be used as collateral at the Fed, after the Federal Reserve Act was amended in 1932. David McKinley, The Discount Rate and Rediscount Policy 97, quoted in David Small and James Clause, The Scope of Monetary Policy Actions Authorized Under the Federal Reserve Act 10 n.22 (FEDS Research Paper 2004-40, 2004).

> when markets are liquid and orderly. They should not be expected to hoard enough liquid assets (or arrange liquid stand-by funding) during normal times to be able to survive on their own during abnormal times, when markets are disorderly and illiquid. That is what central banks are for.
>
> Central banks can create any amount of domestic currency liquidity at little or no notice and at effectively zero marginal cost. It would be inefficient to privatise and decentralise the provision of emergency liquidity when there is an abundant source of free liquidity readily available.[162]

Similarly, Mehrling et al. write that they use the term "dealer of last resort" "in order to draw attention to the modern importance of market liquidity, and hence the importance of placing bounds on price fluctuation."[163]

While the purpose of these proposals is clearly to support the liquidity of certain asset markets, since the ability of commercial banks to deal in these markets is limited by regulation, traditional lender of last resort transactions with commercial banks cannot have the desired effect on the markets in question. As a result, the effect of the proposals is to expand access to central bank liquidity to the securities dealers who are in a position to use central bank funding to limit price fluctuations in asset markets.

b. Market Liquidity is Not a Substitute for Funding Liquidity, But Strictly Adds to Liquidity Risk

When motivating their "dealer of last resort" proposals, both Buiter and Mehrling et al. appear to assume that market liquidity in modern markets substitutes for funding liquidity in traditional markets.[164] Thus they effectively argue that we need a dealer of last resort, because the kind of liquidity we rely on in modern markets is different from the kind of liquidity that the lender of last resort was designed to address.

This, however, is a mistaken understanding of the relationship between market and funding liquidity, as was explained in Parts V.A.3 and V.B.2. First, modern money market instruments rely on both funding liquidity – provided by bank guarantees – and market liquidity. There is no "market-based" money market, but only a "bank-guaranteed" money market.[165] That is, market liquidity isn't substituting for funding liquidity, but is an additional liquidity demand that is created by the structure of modern money markets. Second, in modern markets reliance on market liquidity generates an additional source of funding liquidity risk for borrowers that only exists in collateralized markets that are remargined regularly.

In short, the "dealer of last resort" proposals appear to built on a misconception about the nature of modern money markets.

---

[162] Buiter, *supra* note 156, at 30-31. Buiter's approach is very similar to that of Greenspan, *supra* note 125.

[163] Mehrling et al., *supra* note 8, at 9. They also write that the central bank should accept as collateral non-prime securities in order "to put a floor on their price in times of crisis." *Id.*

[164] Mehrling et al., *supra* note 8, at 7 ("Bagehot's world was centrally about funding liquidity, whereas our world is centrally about market liquidity."); Buiter, *supra* note 156, at 28 ("It may seem that this commoditisation and marketisation of financial relationships that are the essence of the [transaction-oriented-model] would solve the banks' liquidity problem and would make even bank runs non-threatening. If the bank's assets can be sold in liquid markets, the cost of a deposit run or a 'strike' by other creditors need not be a fatal blow.").

[165] See Part I.D.

c. The Problem of Pricing Illiquid Assets

The clear difficulty created by the "dealer of last resort" proposals is that of pricing the transactions. The central bank is asked to intervene because markets are "disorderly and illiquid," but this immediately implies that market prices are not a good way to value the assets that are being supported. Furthermore, because modern money markets don't have the 19th c. safeguards that ensured that only high-quality assets circulated, price determination can be very complicated. In particular, flaws in modern origination processes together with a paucity of personal guarantees make it possible for low quality assets to end up backing the money supply, and to require support from a "dealer of last resort."

There are several possibilities for dealing with this pricing challenge: the central bank prices the assets, the borrowing bank prices the assets, a reverse price auction is held, or a third party prices the assets.

In theory, the central bank can price the assets, but, in practice, no one expects the central bank to have the skill-set to do this. Accurate pricing would require a review of the bank-specific origination and servicing practices that will determine the true quality of each asset, and would require the central bank – in the midst of a crisis – to review the underwriting of the loans created by the banking system. In fact, however, because this is just a matter of pricing collateral and the banks remain liable for the full value of the loan, whether or not it is sufficiently collateralized, it's possible that precise pricing is not that important – especially if the central bank has no intention of letting the dealers to which it is lending fail.

As long as the point of the policy is to protect the dealer banks from failure, then whether or not the loans are fully collateralized is not of primary importance. In this case, the point of the policy may be to prevent the dealers from having to recognize the true value of their assets on their balance sheets – to avoid balance sheet insolvency. If the borrowing bank is allowed to price the assets that it repos, the central bank is making it possible for the borrower to carry the asset at the valuation at which the borrower is willing to buy the asset back in the future. By allowing troubled dealers to avoid recognizing losses on their balance sheets, the central bank supports prices on asset markets that would be adversely affected by the dealers' forced sales.

Buiter argues that pricing can take place in reverse auctions. How such auctions would be designed is, however, a problem that he argues can be solved by auction theorists, but does not address in detail.[166] Perhaps this problem is more difficult than he imagines. The one program that was designed to support a very illiquid asset market while at the same time finding a market price for such assets, the Public Private Investment Program financed a total of $30 billion, and was generally deemed ineffective.[167]

In 2008, the Federal Reserve delegated the task of valuing the collateral that it accepted from the dealer banks to the tri-party clearing banks, which regularly perform that service for other lenders on the tri-party clearing market. Because more than 38% of the borrowings on the Primary Dealer Credit Facility (PDCF) from September 15, 2008 on were collateralized with below investment grade or unrated securities and loans,[168] many of which almost certainly could not be traded actively during the period in which the PDCF was active, it is not clear how the clearing banks went about valuing the worst of the

[166] Buiter, *supra* note 156, at 31-32.

[167] http://www.treasury.gov/initiatives/financial-stability/TARP-Programs/credit-market-programs/ppip/Pages/default.aspx; Cardiff Garcia, Performing PPIFs, FT Alphaville Blog, Oct. 22, 2010.

[168] Data on the Primary Dealer Credit Facility is available here http://www.federalreserve.gov/newsevents/reform_pdcf.htm.

collateral. On the other hand, to the degree that the goal of the policy was to support the dealer banks, perhaps it doesn't matter how these assets were valued.

3. Can a Dealer of Last Resort Stabilize the Collateralized Money Market?

The stated purpose of the "dealer of last resort" intervention is to prevent the market illiquidity that is caused when forced sales are triggered in collateralized money markets after asset prices fall and margin is called. The "market illiquidity" criterion presumably means that if just few small hedge funds are sold out, the dealer of last resort will not intervene, whereas if one large market participant or too many small participants are facing forced sales the dealer of last resort will intervene. The second subsection below evaluates what would be needed for the "dealer of last resort" proposal to successfully address this problem. It is preceded by a subsection that explains why the fundamental differences between dealers and banks mean that the analogy between the "dealer of last resort" and the lender of last resort is very misleading. The third subsection examines the effect on market prices of socializing the dealers' risk. The fourth subsection observes that the "dealer of last resort" cannot address instability created by fundamental changes in the value of assets. The final subsection explains why trading swaps cannot eliminate the funding problems created by the collateralization of the money market.

a. The Central Bank May Have to "Become the Market" in Every Crisis

The function of dealers in supporting market liquidity is very different from the function of banks in honoring deposits and funding guarantees; as a result the nature of a central bank backstop, and indeed the degree of reliance on the central bank is likely to be very different for a dealer of last resort as compared to a lender of last resort. Whereas banks have an obligation to pay their deposits and their guarantees in full, dealers stand ready only to buy and sell assets without any price guarantees whatsoever. Thus, the primary work of a dealer is that of smoothing the market's movement to a new price, and not of setting a floor on asset prices.

Some proponents of "dealer of last resort" activities do not appear to recognize this distinction. Mehrling et al., for example, appear to assume that the dealer of last resort functions just like a lender of last resort, stepping in very briefly to support markets until the dealer system recovers. They write:

> what is clearly needed is some entity that is willing and able to use its own balance sheet to provide the necessary funding. … what we need is a dealer system that offers market liquidity by offering to buy whatever the market is selling. Only in crisis time does the central bank backstop become the market; in normal times, the central bank backstop merely operates to support the market.[169]

This description appears to assume that in normal times dealer systems will prevent market fluctuations due to fire sales by small market participants. In fact, when the dealer system is functioning correctly fire sales can – and arguably should – have price effects. This takes place because the dealers make decisions based on order flow and a large order should cause prices to fall.[170] This was explained by Jack Treynor in his classic model of dealer pricing. End investors are the ones who might be seen as setting a floor on assets prices, because they are ones who may stand ready to buy at the price that they perceive to be fundamental value and then to hold the position over the long term. The traditional role of the dealer is

[169] Mehrling et al., *supra* note 8, at 9.

[170] "Short range [is] what's happening this morning – and long range [is] what's going to happen this afternoon." is how John Whitehead, a retired Goldman Sachs' CEO, described the thinking process of Gus Levy, who developed Goldman Sachs' market making ability in the 1960s and '70s. Charles Ellis, The Partnership 177 (2008).

not to support market prices. Instead, Treynor writes: "the dealer has very limited capital with which to absorb an adverse move in the value of the asset. Furthermore, the dealer's spread is too modest to compensate him for getting bagged."[171]

Thus, Mehrling et al. are mistaken. A traditional dealer does not "use its own balance sheet to provide the necessary funding" except over very short time horizons. And both the terms "dealer of last resort" and "market maker of last resort" are misnomers, because no one who is proposing central bank intervention into markets is actually proposing that that the central bank act as a market maker, entering the market with the expectation of laying off all risk, if not on the same day, at least in the same week.

It is true, however, that large-scale proprietary trading and the management of balance sheet exposure to related risks has become a core function of dealers in recent decades. (This change has taken place not coincidentally alongside the conversion of dealers from partnerships with unlimited liability to limited liability corporations – where management is taking risks with shareowners' money.) Despite this evolution, it seems likely that even dealers engaged in large-scale proprietary trading will behave to some degree like traditional dealers and let a liquidity spiral run before stepping in to buy in significant quantities – and then they may well allow the price to continue falling as they build up a significant stake in the assets. After all, if sellers want to sell at unreasonably low prices, this will only add to the dealers' proprietary trading profits in the end. Indeed, this is what we witnessed in 2008 when Bear Stearns and Lehman were failing: for the most part, the dealers instead of using their balance sheets to support prices on the market sought to avoid being caught holding assets that are falling in value.

Thus, if the purpose of the dealer of last resort is to prevent fluctuations caused by market illiquidity – or the disappearance of end investors willing to buy into the market – then the "dealer of last resort" is going to have to step into the shoes, not of the traditional dealers, but of the proprietary traders and end investors. And the burden of protecting markets from plummeting prices for assets will fall heavily on the dealer of last resort or central bank.

Indeed this burden did fall heavily on the Federal Reserve in 2008, which repeatedly accepted as collateral tens of billions of dollars of assets that were not trading from the investment banks.[172] At the start of October 2008, the two facilities that allowed the investment banks to convert private sector assets into cash and Treasuries accounted for 60% of the massive expansion of the Federal Reserve's balance sheet as compared to the previous year.[173] In short, because providing a price floor for assets is not the economic function of a dealer, a central bank that acts as a so-called "dealer" of last resort must be prepared to purchase assets on this scale – and effectively to become the market – in every crisis.

The Fed in 2008 cannot, however, be accurately designated the prop trader of last resort, because all of the asset risk that the Fed took on was in the form of repos. Thus, there was always a dealer bank that was obliged to make good on the loan, even if the value of the collateral fell below the value of the loan. If the Fed's goal was to support the dealer system by preventing the remaining dealer banks from failing (just as a traditional lender of last resort protects the banking system), the collateral was, in some sense, irrelevant.

---

[171] Jack Treynor, The Economics of the Dealer Function, 43 Fin. Analysts J. 27, 27 (1987).

[172] *See supra* note 154.

[173] Federal Reserve Board of Governors, H.4.1 Release: Factors affecting reserve balances, Oct. 2, 2008.

The Fed's repo facilities for dealer banks were in use for a full eight months after Lehman Brothers' failure.[174] This is fairly strong evidence that, even if impending fire sales of assets were the symptom that induced the Fed to act initially, the underlying problem was the inability of dealer banks to fund themselves either unsecured or on repo markets. Because of the inherent instability of repo market funding, a central bank that adopts a "dealer" of last resort policy can anticipate offering such prolonged support to the dealer banks on a regular basis.

b. Who Gets Access to the Central Bank?

Market illiquidity on collateralized money markets can be caused by the forced sales of a single large borrower or many small borrowers, all of whom received margin calls due to the same price movement. The emergency facilities that supported the repo market in 2008 were specifically directed to the primary dealers, or in other words, the largest dealer banks. This policy was clearly tailored to avert forced sales by the largest borrowers on the repo markets.

It is unlikely, however, that the same policy could be successful in averting market illiquidity created by the simultaneous forced sales of many smaller borrowers. As I have just explained, the dealers do not perceive their role in markets to be the support of asset prices. Thus, even if the central bank eliminates the risk for the dealer banks of supporting asset prices, it does not eliminate the profits created by buying at a low price an asset that can later be sold at a high price. In short, even with "dealer of last resort" support, the dealer banks have every incentive to let a liquidity spiral run, before choosing to step in to support the price of the asset, and, if market illiquidity is caused by the forced sales of many small borrowers, central bank lending to the primary dealers is unlikely to forestall the price decline.

Because dealers do not play the same role in asset markets that banks play in loan markets, in order for the "dealer of last resort" to address a problem of market illiquidity created by the simultaneous forced sales of many smaller borrowers, the central bank would have to be willing to enter into repo transactions directly with these smaller borrowers. Determining which borrowers are eligible for such treatment and averting the danger of borrowers who default is likely to be very difficult. If these problems are insuperable, as they may well be, the "dealer of last resort" policy will fail to address some instances of market illiquidity created by collateralized lending markets.

c. Will Socializing Dealers' Risk Distort Market Prices?

As long as the "dealer of last resort" policy restricts access to central bank credit to a select group of firms, it will undoubtedly be very effective in protecting asset prices from instability due to fire sales by these select firms. But the proposed policy would also introduce a very troubling asymmetry into our markets. Who has access to central bank's discount window? Retail investors clearly do not, whereas "dealers," however they end up being defined, do. The privileged dealers effectively have access to an unlimited balance sheet and can employ leverage without worrying about being forced into a fire sale – and no longer face the traditional constraints that govern dealers' activities.[175]

This result may actually be the goal of supporters of the dealer of last resort policy. Buiter writes "[Private financial entities] should not be expected to hoard enough liquid assets (or arrange liquid stand-by

[174] Data on the Primary Dealer Credit Facility is available here http://www.federalreserve.gov/newsevents/reform_pdcf.htm.
[175] Treynor, *supra* note 171, at 27.

funding) during normal times to be able to survive on their own during abnormal times, when markets are disorderly and illiquid."[176] Presumably he wants private financial entities to be able to behave as though they have unlimited balance sheets and can hold illiquid assets until they mature.

Even though this policy is framed as being available to all private financial entities, in practice, it seems highly unlikely that the dealer of last resort facilities will be open to all private financial entities -- if only because it would be an invitation for fraud. As a result, those without the privilege of dealer of last resort access will be limited by their capital position in the degree to which they can increase their profits using leverage.[177]

In short, as a practical matter, this policy is likely to have the effect of protecting the privileged dealers from losses due to market risk, while other market participants do not receive similar protection. By reducing the costs of leverage to the privileged dealers it is also likely to increase their use of leverage. If the central bank does not monitor the behavior of the privileged dealers vigilantly, it could end up making financial markets more risky, by making the largest financial market participants believe that it is "safe" for them to take on more risk.

These dangers are offset in part by the fact that when the dealer of last resort lends to these firms in a crisis, it will forestall an immediate collapse and the economic repercussions of such a collapse. And this fact almost certainly justifies the Federal Reserve's actions in 2008. It is less clear that this fact is enough to justify embracing the "dealer of last resort" proposal as a standing policy. After all it is possible that a standing policy will result in enough increased risk taking by the privileged firms that crises requiring dealer of last resort intervention will occur with frequency.

d. Collateralized Money Markets Can be Destabilized by Fundamental Asset Price Changes

Forced sales and the consequent price movements that may be mitigated by central bank lending to dealer banks are not the only problem created by collateralized money markets. These markets also need stabilization when the fundamental value of the collateral used in these markets falls; for example, when interest rates rise or expectations about the future fall. When this happens, it is possible that the collateralized money market as a whole will not have the capacity to provide credit consistent with previous levels – even in the absence of fire sales and illiquidity. In the worst scenarios, the fundamental decline in the value of collateral can have an effect on the money market comparable to the failure and disappearance of banks that took place during the Great Depression.

To the degree that the purpose of the "dealer" of last resort is to alleviate market illiquidity, problems created by fundamental changes in assets prices would appear to be outside the remit of the central bank. Collateralized money markets, however, mean that the analysis must be more complex than this. The central bank's primary duty is to prevent sudden declines in the money supply, as was recognized by Henry Thornton more than two centuries ago.

When money markets are collateralized, a decline in asset prices can map directly into a decline in credit availability on the money market. Such a decline is likely to have a real economic effects as

---

[176] Buiter, *supra* note 156, at 30-31.

[177] For the role that capital constraints typically play in risk-taking, see Andrei Schleifer & Robert Vishny, The Limits of Arbitrage, 52 J. Fin. 35 (1997).

intermediaries who used to allocate credit to the real economy will be credit constrained. To mitigate the effects of this credit crunch, the central bank is likely to feel obliged to provide support to the market through its dealer of last resort facilities, even when there is no reason to believe that the low asset prices are caused by market illiquidity. In short, a dealer of last resort facility may well be used to distort market prices upwards, because this is the only way for the central bank to improve credit availability on collateralized money markets.

e. Why Swaps Don't Mitigate Instability on the Collateralized Money Market

Individual repo borrowers can offset, or "hedge," the dangers of an unexpected increase in interest rates or in the likelihood of default by entering into an interest rate swap or a credit default swap. An interest rate swap can be used to protect against the possibility that interest rates rise and the value of the debt that is used as collateral falls. A credit default swap can be used to offset a fall in bond value due to an increase in the credit risk of the bond issuer.

In theory, combining a bond investment with interest rate and credit default swaps is a way of manufacturing a "safe" asset. In practice, while the swaps will often offset a substantial portion of the interest rate or credit risk, payments due on these swaps are not perfectly correlated with any particular bond investment, and as a result some of these hedges turn out to be much less effective than they were initially expected to be.

These, however, are issues faced by individual investors who wish to use swaps to hedge the risk that their collateral falls in value. The broader question for the collateralized money market as a whole is whether these instruments can reduce the risk of instability on the market.

Swaps are derivatives and exposures on swaps are usually collateralized, especially since the implementation of regulatory reforms that require central clearing of the vast majority of these derivatives contracts. Collateralization has the same effect in derivatives markets that it has in repo markets: it protects the party to whom money is owed at the expense of exposing the party who owes money to the risk of repeated demands for additional collateral.

What happens, then, when a repo borrower hedges the risk of an asset she has repo'd with a swap. From the point of view of the borrower it matters whether the repo is with the same counterparty as the swap. If the counterparties are the same – and the hedge works – when the repo'd asset falls in value, the swap will rise in value, and the borrower will be required to post only the net collateral. In short, if the counterparties are the same, combining the repo with the swap has the effect of reducing the size of the margin call when the repo'd asset falls in value.

If the counterparties are not the same, however, the repo borrower will face a margin call for the full decline in value of the repo'd asset from the repo counterparty, and at the same the repo borrower or her representative, will issue a margin call to receive collateral from the swap counterparty. Whether or not the swap counterparty meets the margin call, the repo borrower still has the obligation to post collateral on the repo. In short, when the counterparties are not the same, swaps do not reduce the collateral posting requirements created by the repo system. On the contrary, in many ways they increase the need for collateral posting.

Because of recent rules requiring that swaps be traded in clearinghouses – with the admirable objective of ensuring that regulators have access to information about these markets – repo borrowers are more likely than ever to face different counterparties for their repos and their swaps. Unsurprisingly, services related to collateral posting have become one of the more important growth industries for the dealer banks.[178]

Consider now the effects of swaps on the collateralized money market as a whole. The only circumstance under which swaps reduce collateral posting requirements is when both the swap and the repo hedged by the swap are transacted with the same counterparty, for example, a prime broker. This transaction may also have the effect of transferring the collateral posting obligation to a third party: if the prime broker lays off its risk on the swap by entering into a back-to-back transaction with a third party, then the repo borrower is protected by the swap from having to post collateral, but that protection comes at the cost of transferring the obligation to someone else. And, when the repo counterparty and the swap counterparty are different, the swap may have the effect of increasing collateral posting obligations by creating a situation where both the repo borrower and her swap counterparty are required to post collateral. In the latter case, if everything goes smoothly the repo borrower can use the collateral received and post that against the repo. One can easily imagine circumstances, however, where everything does not go smoothly and collateral posting requirements become very onerous.

Overall the only situation in which collateral posting requirements for the market as a whole are reduced by swap contracts is when a repo borrower enters into a swap with the same counterparty and the risk of the swap is not laid off on a third party. In all other circumstances the swap serves only to transfer – or to increase – the collateral posting obligations in the market.

4. A Better Solution to Instability: Money Markets Based on Unsecured Debt

The underlying problem with the collateralized money market is that the use of collateral to protect the lender relies too heavily on the presence of both market liquidity and funding liquidity.[179] While proposals have been made for a "dealer of last resort" to address these liquidity problems by giving dealer banks access to the central bank's lending facilities, these proposals assume incorrectly that liquidity will be transferred from the dealers to the market as a whole in the same way that it is transferred by banks to the market as whole. Because central bank liquidity cannot be expected to flow through the dealers to the forced sellers, to address market illiquidity this policy would have to allow for very broad access to the central bank.

In practice in 2008, the Federal Reserve did not "put a floor" on asset prices, but instead supported the largest dealer banks through the crisis using repos against the collateral that was at risk of a fire sale. This action required that the Federal Reserve accept as collateral vast amounts of below-investment-grade assets for a period spanning eight months.

---

[178] Morgan Stanley & Oliver Wyman, Wholesale and Investment Banking Outlook 31 (2013) (Exhibit 42 estimates $5 – 8 billion in revenue from these services over the next five years.).

[179] Note that it is true that in commodities derivatives markets exposures have been collateralized for more than a century and this system has functioned very well. It is likely, however, that the reason this collateralization was successful was that it accounted for only a very small fraction of U.S. financial markets. That is, when the collateral is a small fraction of the total market or when the borrowers are small both individually and in aggregate, the forced sales intrinsic to collateralized lending have little effect on price. In short, the commodities exchange experience provides no evidence that the general movement towards collateralization of money market exposures can possibly be successful.

Because forced sales are integral to the way that the collateralized money market functions, the Federal Reserve should anticipate that, if it adopts a "dealer of last resort" policy, it will be required to engage in such substantial support of the dealer system in every crisis. Furthermore, because even fundamental movements in long-term asset prices can cause a credit crunch and instability on the collateralized money market, the Federal Reserve will face significant pressure to support the collateralized money market by supporting asset prices at a level above their fundamental value.

Because the instability of collateralized money markets and its effects on the real economy can only be addressed if the central bank is willing to become the market, to eliminate entirely the traditional balance sheet risk faced by the largest dealer banks, and to distort the market, some other solution must be found. The historical record indicates that unsecured money markets can be very successful and can function as the underpinning for flourishing economic growth. An advantage of unsecured money markets is that, by eliminating the use of collateral, market forces will help to reduce the ability of poorly managed firms to access the money markets and less of a burden will be placed on regulators to determine which firms are troubled.

The reliance of collateralized money markets on market and funding liquidity could also be reduced, however, without entirely giving up on them. For example, margin calls could be permitted only on a monthly or a quarterly basis. This would delay the reaction to price changes and give the market time to adjust to a change in fundamental value with fewer fire sales. It would also allow many borrowers time to search for a buyer for the assets that had fallen in value.

Even this policy, however, cannot mitigate the problem that the stock of collateral can decline when interest rates rise or expectations about the future worsen. For this reason, even if the collateralized money market is preserved, we need a functional unsecured money market that can act as a buffer when lending is contracting on the collateralized one.

We also need a central bank that recognizes that the stability of the money market depends fundamentally on its willingness to act as a monitor of moral hazard. Partial reserve banks are only able to operate because there is a central bank that supports them through a liquidity crisis. For this reason, every bank should be fearful of falling out of the good graces of the central bank. While the central bank's power of life and death over the banking system should not be used abusively, it needs to present enough of a threat to prevent banks – especially large banks – from engaging in activities that undermine the quality of the money markets.

5. Why a Lender of Last Resort is Good Policy and a Dealer of Last Resort is Not

The reason why the lender of last resort is good policy, whereas the dealer of last resort is not, lies in the distinction between commercial and the dealer banks.

Commercial banks fund real economy activity, not only by making loans, but also by guaranteeing the debt of real economy borrowers, making it easier for them to access credit markets. Commercial banks have not traditionally had to collateralize their guarantees, and because they don't need to post collateral, they are able to extend short term credit to borrowers on an unsecured basis. In modern markets, this credit often takes the form of a credit line accessible through a business checking account. Thus the two key characteristics of commercial bank lending are that the instruments used, loans and guarantees, ensure that commercial bank funding is directly related to real economic activity, and that unsecured credit is

available to qualified borrowers.[180] In short, traditional commercial banks are experts in bearing and managing credit risk.

The value of commercial banking lies in making it possible to have money markets based on unsecured credit that is then extended broadly across the business community and makes modern economic growth possible. The banking system merits the protection of a lender of last resort, because it provides such broad benefits to the community at large, and there is good reason to believe that without a lender of last resort a banking system can be expected to collapse entirely, since it is based on interlocking chains of credit and is subject to liquidity crises.

By contrast, there is a long history of dealer systems that support trade on financial markets and are not at risk of collapse in the absence of a dealer of last resort. Whereas commercial banks rely on their own creditworthiness to make credit more generally available, dealer banks traditionally don't take deposits and don't guarantee the debt of others, but instead are true intermediaries. They are experts in the markets in which they trade, and are therefore willing to take on asset risk temporarily because they have good information about their ability to lay off that risk in the immediate future.

Traditionally dealer banks avoid bearing risk for more than very short periods of time. When they lend to brokerage clients, they demand that the loans be collateralized, and that they have the right to liquidate that collateral if a margin call isn't met. When they buy an asset, they do so with a plan to sell it. When they take on risk in a derivative contract, they enter into a back-to-back trade to lay of the risk of that contract. In short, the traditional role of a dealer bank is to manage risk, by doing their best to bear as little of it as possible. When dealer banks play this role – as they have for centuries in Anglo-American markets – market prices can accurately reflect the forces of demand and supply, because the dealers avoid bearing risk, but simply intermediate markets.

Dealer banks have a very ambivalent relationship with the real economy. If investors want to put money into a security that is issued by a business which has a prospectus that states outright that it never intends to do business and to make money (and if it is legal to sell this security), the investment bank will intermediate these purchases. A commercial bank, by contrast, will lose money if it is so poorly managed that it guarantees this company's debt.

This description is, of course, a general one. It has always been the case that a few dealer banks try to use their information about the market to predict where it will go, and earn proprietary trading profits. Financial history is littered with dealer bankruptcies caused by prop trading and a failure to correctly predict price movements. As up until recent decades dealer banks were partnerships with unlimited liability for their debts, these failures were often viewed as cautionary tales that served to encourage very careful risk management practices among the dealer banks – and close monitoring of every partner's activities.

More recently dealer banks have become limited liability corporations and have started to take on proprietary trading risk: they use their information about markets to take positions on where those

[180] The modern "originate to distribute" model of banking attenuates the connection between commercial bank lending and the real economy and makes commercial bank more like investment banks. Because commercial banks continue to issue debt guarantees and to hold significant numbers of loans on their balance sheets, however, the traditional model of commercial banking is still important today.

markets are headed. Because dealer banks traditionally lay off risk in back-to-back trades, from the point of view of the public a dealer bank's proprietary purchases are indistinguishable from transactions where it is acting as an intermediary. The fact that the bank's positions do not offset each other is only likely to be evidenced by the bank's internal records of its exposures. When this fact is combined with the fact that many financial contracts are just agreements to make future payments depending on the realization of future prices – that is, they don't finance any economic activity – we recognize that dealer bank activities are not always closely connected to the finance of real economic activity. Sometimes they are simply predictions about the future psychology of the markets.

Contrast the very imprecise connection with the real economy that is exhibited by dealer banks, with the tight connection that exists between traditional commercial banks and the real economy. The business of commercial banking is that of lending directly and of providing guarantees of payment that facilitate borrowing. When the commercial bank bears risk it is very likely to be real economy risk; by contrast when a modern dealer bank bears risk, it may just be the risk that a predicted event does not take place.

In short, the reason that dealer banks have not traditionally had access to the lender of last resort is that they do not play the same role in the real economy that banks do. They don't bear real economy risk, but only facilitate the allocation of that risk to others. When a commercial bank fails, the economy's ability to bear real economy risk declines. When a large dealer bank fails, the economy's mechanism for allocating risk may be temporarily dislocated, but it's much less likely that the economy's ability to bear risk has been reduced. In short, dealer banks are more expendable than commercial banks.

Furthermore, there is little or no evidence that collateralized money market instruments play the same role as uncollateralized money market instruments in economic growth, and thus little or no evidence that collateralized money markets are necessary to the community at large. In fact, the growth of collateralized money markets may undermine traditional unsecured money markets, where a financial institution's ability to borrow depends on its credit quality, and thereby undermine the market forces that promote high credit quality in the financial industry. For this reason, the collateralized money markets may be destabilizing the financial industry. While the temporary support of these markets in 2008 was well justified, much more evidence of the value of collateralized money markets to the process of economic growth needs to be presented before dealer banks are given privileges similar to those of commercial banks.

VI. Conclusion

Modern repo and derivatives collateral markets grew up over the past decade or two as a consequence of radical legal reform of the bankruptcy code and of derivatives regulation.[181] The crises of 2007-08 should be understood as a warning signal that these markets are dangerously unstable and put the entire financial system at risk. While proposals to use the central bank to stabilize these markets are well-meaning, they do not reflect a macroeconomic understanding of the role played by the banking system in the economy, and thus fail to take into account the costs of replacing the traditional bank-based unsecured money market with a dealer-based collateralized money market.

[181] *See supra* note 1. Of course, central bank use of repo for monetary policy is of much older vintage, but also bears little or no relationship to modern repo markets.

A key theme in the monetary literature of 19th c. England is the importance of the circulation of uncollateralized debt to the ability of the money supply to grow with the needs of the economy. For this reason, it should be a matter of serious concern that many modern authors do not appear to understand that the collateralized money market that we have today is an experiment that differs significantly from the uncollateralized money markets that existed from the late 18th century through about 1980.[182] Given the different economic roles of commercial banks and dealer banks, it is unlikely that the collateralized money market has the same macroeconomic properties as the uncollateralized money market. While it is true (as the crises of 2007-08 demonstrated) that a collapse in the money supply has adverse macroeconomic consequences whether or not that money supply was collateralized, many questions about the collateralized money market remain: Is it possible to stabilize the collateralized money market? Is economic growth adversely affected by the presence of a collateralized money market? Will the new monetary system undermine the traditional system which was the foundation upon which the growth of the 19th and 20th centuries was based?

[182] For example, Mehrling et al., *supra* note 8, at , claim inaccurately that the bills of Bagehot's time were "collateralized."