Financial Markets, Institutions and Liquidity
Franklin Allen and Elena Carletti*

1. Introduction

One important reason for the global impact of the 2007–2009 financial crisis was massive illiquidity in combination with an extreme exposure of many financial institutions to liquidity needs and market conditions. As a consequence, many financial instruments could not be traded anymore; investors ran on a variety of financial institutions, particularly in wholesale markets; financial institutions and non-financial firms started to sell assets at fire sale prices to raise cash; and central banks all over the world injected huge amounts of liquidity into financial systems.

But what is liquidity and why is it so important for financial institutions and firms to command enough liquidity? In this paper we consider liquidity from a number of different perspectives and consider how the academic literature has evolved in light of developments in funding markets and their liquidity during the crisis that started in 2007. We start in Section 2 with a discussion of funding liquidity for banks and financial institutions. We review the literature before the crisis on deposits as a funding mechanism for banks, which was traditionally what the literature on financial stability focused on. This was because most historical financial crises involved the withdrawal of money from deposit accounts and the hoarding of this cash. We next consider interbank markets. There is a large literature on liquidity in these markets and how they operate. There are also many papers that consider their role in propagating contagion. The literature before the crisis is surveyed and then recent contributions made since the crisis are discussed. We argue that deposit insurance to a large extent solved the problem of runs by retail investors. In fact, deposits were one of the most stable forms of finance during the crisis. However, going forward this may not be the case as the credibility of deposit insurance depends on the extent of guarantees and the fiscal position of the government. In contrast, interbank markets were heavily affected by the crisis and in many cases froze. It seems that asymmetric information played an important role and a number of theories based on asymmetric information have been developed to explain market freezes.

Section 3 considers funding markets for firms. Important theories of liquidity management were developed by Holmstrom and Tirole (1998) and Gorton and Huang (2004). Funding for firms was not as disrupted during the crisis as funding for banks and other financial institutions. However, one of the important implications of these theories is that there may be a shortage of safe liquid assets for firms to manage their liquidity needs and that governments may need to intervene to supply such assets. It seems that there was in fact a shortage of these assets. Instead of government securities, the market supplied securitised mortgages. Although most tranches were rated AAA by the credit rating agencies, these securities were not in fact safe. When house prices began to fall, this became clear and this triggered the crisis.

* We thank participants for helpful comments.
Section 4 considers the funding markets for real estate. Despite the importance of real estate in triggering previous financial crises, relatively little attention has been paid to these funding markets in the literature compared with those for financial institutions and firms. However, as indicated above, securitised mortgages played a very important role in triggering the crisis.

The disruption of so many funding markets led to intervention in many markets by central banks (Section 5). Although the literature has for many years considered the role of central banks as lenders of last resort (LOLR), other kinds of intervention were not anticipated or analysed. The academic literature has only just begun to consider the different kinds of interventions that central banks should undertake when funding markets break down. The other aspect of central bank policy that has not fully been analysed is the effect of low interest rates and quantitative easing on asset prices. The effect of withdrawing these measures and returning to a normal interest rate regime is also not well understood. Finally, Section 6 concludes.

The paper draws on discussions of the literature in Allen, Babus and Carletti (2009), Allen et al (2011) and Allen and Carletti (2013).

2. The Funding of Financial Institutions

2.1 Views before the crisis

2.1.1 Bank deposits

The most important type of financial institution is the bank. Historically banks were mostly funded by deposits. Much of the traditional literature on financial crises has focused on bank runs where depositors withdraw their money and hoard cash. There are two theories to explain the origins of banking crises. One line of argument maintains that they are undesirable events caused by random deposit withdrawals unrelated to changes in the real economy. In the influential work of Bryant (1980) and Diamond and Dybvig (1983) bank runs are self-fulfilling prophecies. In these models, agents have uncertain needs for consumption in an environment in which long-term investments are costly to liquidate. If depositors believe that other depositors will withdraw then all agents find it rational to redeem their claims and a panic occurs. Another equilibrium exists where everybody believes no panic will occur and agents withdraw their funds according to their consumption needs. In this case, their demand can be met without costly liquidation of assets.

The second set of theories of banking crises posit that they are a natural outgrowth of the business cycle. An economic downturn will reduce the value of bank assets, raising the possibility that banks are unable to meet their commitments. If depositors receive information about an impending downturn in the cycle, they will anticipate financial difficulties in the banking sector and try to withdraw their funds, as in Jacklin and Bhattacharya (1988). This attempt will precipitate the crisis. According to this interpretation, crises are not random events but a response of depositors to the arrival of sufficiently negative information on the unfolding economic circumstances. This view is consistent with the evidence in Gorton (1988) that in the United States in the late 19th and early 20th centuries, a leading economic indicator based on the liabilities of failed businesses could accurately predict the occurrence of banking crises.
One strand of the business cycle explanation of crises emphasises the role of information asymmetry in triggering a banking crisis. In this view, a panic is a form of monitoring. Chari and Jagannathan (1988) focus on a signal extraction problem where some depositors withdraw money for consumption purposes while others withdraw money because they know that the bank is about to fail. In this environment, depositors may also withdraw because they cannot distinguish whether there are long lines at banks because of consumption needs or because informed depositors are getting out early. Chari and Jagannathan show crises occur not only when the outlook is poor but also when liquidity needs are high despite there being no new information on future returns.

Building on the empirical work of Gorton (1988), Allen and Gale (1998) develop a model that is consistent with the business cycle view of the origins of banking crises. They assume that depositors can observe a leading economic indicator that provides public information about future bank asset returns. If these returns are expected to be high then depositors are quite willing to keep their funds in the bank. However, if returns are expected to be sufficiently low, they will withdraw their money and there is a crisis.

Calomiris and Kahn (1991) show that the threat of bank liquidation disciplines the banker when he can fraudulently divert resources ex post. The first-come, first-served constraint provides an incentive for costly information acquisition by depositors. Calomiris and Kahn regard bank runs as always beneficial since they prevent fraud and allow the salvage of some of the bank’s value. Diamond and Rajan (2001) develop a model in which banks have special skills to ensure that loans are repaid. By issuing demand deposits with a first-come, first-served feature, banks can pre-commit to recoup their loans. This allows long-term projects to be funded and depositors to consume when they have liquidity needs. However, this arrangement leads to the possibility of a liquidity shortage in which banks curtail credit when there is a real shock.

While the multiple equilibria theory of bank runs explains how panics may occur, it is silent on which of the two equilibria will be selected. Depositors’ beliefs are self-fulfilling and are coordinated by ‘sunspots’. Sunspots are convenient pedagogically but they do not have much predictive power. Since there is no real account of what triggers a crisis, it is difficult to use the theory for any policy analysis. The business cycle theory also has panic runs as well as fundamental runs. Again there is no natural way to choose the equilibrium.

A selection mechanism that applies to this type of coordination game is introduced in Carlsson and van Damme (1993). The authors analyse incomplete information games where the actual pay-off structure is randomly drawn from a given class of games and where each player makes a noisy observation of the game to be played. Such games are called global games. In a global games setting, the lack of common knowledge about the underlying pay-off structure selects the risk dominant equilibrium as the unique equilibrium of the game. Morris and Shin (1998) successfully apply this approach to coordination games in the context of currency crises, when there is uncertainty about economic fundamentals. Rochet and Vives (2004) and Goldstein and Pauzner (2005) have used global games to study banking crises. Global games represent a nice combination of the multiple equilibria and business cycle approaches to panics. The equilibrium selected is unique and there are both fundamental and panic runs. An important contribution by Chen, Goldstein and Jiang (2007) establishes the empirical applicability of the global games approach. The authors develop a global games model of mutual fund withdrawals, where
strategic complementarities among investors generate fragility in financial markets. Using a detailed dataset, they find that, consistent with their model, funds with illiquid assets exhibit stronger sensitivity of outflows to bad past performance than funds with liquid assets.

Deposit funding plays a significant role in the theoretical literature on financial crises because historically crises have involved the withdrawal of money from banks and the hoarding of cash. However, the widespread introduction of deposit insurance meant that deposits became one of the more stable forms of funding during the recent crisis, as we shall see. We turn next to another form of funding that was important in the theoretical literature and the breakdown of which played an important role in the crisis.

There is a large empirical literature on banking crises. Friedman and Schwartz (1963) have written a comprehensive monetary history of the United States from 1867–1960. Friedman and Schwartz argue that the crises were panic-based, as evidenced by the absence of downturns in the relevant macroeconomic time series prior to the crises. This contrasts with Gorton’s (1988) evidence that banking crises in the National Banking Era were predictable, which suggests banking crises are related to the business cycle. Calomiris and Gorton (1991) provide a wider range of evidence that crises are grounded in fundamentals. Wicker (1980, 1996) shows that, despite the absence of collapses in US national macroeconomic time series, in the first two of the four crises identified by Friedman and Schwartz in the early 1930s there were large regional shocks. The authors attribute the crises to these shocks. Calomiris and Mason (2003) undertake a detailed econometric study of these four crises using a broad range of data and conclude that the first three crises were fundamentals-based while the fourth was panic-based.

2.1.2 Interbank markets

In recent decades interbank markets have come to play an increasingly significant role in the funding of banks. Ideally, these markets should ensure an efficient liquidity transfer between surplus and needy banks. They are the focus of central banks' implementation of monetary policy and a smooth functioning of interbank markets is essential for maintaining the stability of the overall financial system. Despite this key role and the potentially significant effect their functioning has on the whole economy, there was not a large literature studying interbank markets prior to the crisis.

Bhattacharya and Gale (1987) is the pioneering theoretical study in this area. They analyse a setting in which individual banks face privately observed liquidity shocks due to a random proportion of depositors wishing to make early withdrawals. In addition, each bank has private information about the liquid fraction of its portfolio. Since the liquidity shocks are imperfectly correlated across intermediaries, banks co insure each other through an interbank market. Bhattacharya and Gale show that, even in the absence of an aggregate liquidity shock for the intermediary sector as a whole, banks are induced to under-invest in liquid assets and free-ride on the common pool of liquidity because of the lower return that liquid assets yield. A central bank can mitigate this problem by (even imperfectly) monitoring banks’ asset choices. However, the authors argue that one would not expect to achieve the first-best, as in such an asymmetric information setting it seems unrealistic to assume that a central bank can elicit perfect knowledge of the quality of the assets across all banks’ portfolios.
The Bhattacharya-Gale model provides a foundation for the analysis of the functioning of financial markets and financial intermediaries, optimal liquidity provision and financial fragility. One line of research that was developed aims to provide a deeper and more general understanding of the interaction of financial markets and financial intermediaries. In the Bhattacharya-Gale setup, the characterisation of interbank markets is quite rudimentary and, in addition, interbank markets are not part of an optimal arrangement. Thus, it is important to have a framework with a role for both financial intermediaries and for markets, modelled from first principles. Allen and Gale (2004a, 2004b, 2007), among others such as Diamond (1997) and Fecht (2004), develop such an approach. They argue that in modern financial systems financial markets and financial intermediaries are complementary. As in Diamond and Dybvig (1983), intermediaries provide an insurance function to consumers against their individual liquidity shocks. However, individual investors cannot trade directly in the full range of markets since it is too costly for them due to information and transaction costs. This is the reason why markets also play an important role in this environment. Markets allow financial intermediaries (and hence their depositors) to share risk. Intermediaries such as banks and mutual funds can invest in financial markets. They provide risk-sharing services by packaging existing claims on behalf of investors who do not have access to markets and, of course, are trading these claims on markets. Such a general equilibrium framework allows a normative analysis of liquidity provision by the financial system.

Consumers deposit funds into banks which provide liquidity insurance such that depositors can withdraw whenever they have liquidity needs. Banks accumulate the funds and lend them to firms to fund long-term investments. There are two types of uncertainty concerning liquidity needs which makes liquidity management on the part of banks quite difficult. The first is that each bank is exposed to idiosyncratic liquidity risk. At any given date its customers’ liquidity needs may be higher or lower than expected. The second type of uncertainty is the aggregate liquidity risk which banks have to face. In some periods system-wide liquidity demand is high while in others it is low, thereby exposing all banks to the same shock at the same time.

What Allen and Gale analyse in this framework is the ability of banks to hedge themselves against these liquidity shocks. They show in Allen and Gale (2004b) that this crucially depends on the completeness of financial markets. If markets are complete, in the sense that for each aggregate state an Arrow security can be traded, then the financial system provides liquidity efficiently as it ensures that banks’ liquidity shocks are hedged. In particular, they show that in an environment with complete markets and in which intermediaries can offer complete contingent contracts, the resulting allocation is incentive-efficient. With complete contracts, the consequences of default will be anticipated and therefore included in the contract, so default and financial crises do not occur. If intermediaries can only offer incomplete contracts – a case in point is where banks only offer deposit contracts – default can improve welfare by improving the contingency of contracts. Thus, financial crises do occur in such a model, but are not necessarily a source of market failure. Hence, even in this case with incomplete contracts, the financial system provides optimal liquidity and risk sharing if markets for aggregate risks are complete. A set of complete and perfect financial markets, which includes interbank markets, is necessary for the efficient functioning of the financial system. However, missing markets may provide a role for government intervention. If markets are incomplete, then there may be too much or too little liquidity and government regulation may be welfare-improving.
Allen and Gale (2004a) explore in further detail interesting ramifications of this framework. By using a simplified version of the general equilibrium model introduced in Allen and Gale (2004b), they investigate the role of liquidity in determining asset prices. The incompleteness of markets leads to inefficient provision of liquidity by the financial system. This can generate cash-in-the-market pricing, which implies that the prices of long-term safe assets can fall below their fundamental value. This leads to financial fragility, which means that even small shocks can have large price effects and this can lead to financial crises.

An important contribution to the literature on interbank markets is that of Freixas and Holthausen (2004). They analyse the scope for international interbank market integration when cross-border information about banks is less precise than home-country information. The timing of consumption needs generates liquidity shocks for the banks, both at the individual and at the aggregate level. Banks can cope with these shocks by investing in a storage technology or can use the interbank market for channelling liquidity. The authors look at secured repo and unsecured interbank lending markets, since both allow banks to cope with liquidity shocks, and they consider under what conditions segmented or integrated international interbank markets exist. They show that a segmented interbank market is always an equilibrium, while the emergence of an integrated international market depends on the quality of cross-border information. Only if cross-border information is sufficiently precise is integration of markets possible.

2.1.3 Contagion and interbank markets

Another important aspect of the operation of interbank markets is the possibility for financial contagion that they introduce. Financial contagion refers to the process by which a shock in one part of the financial system spreads to other parts through a series of ‘interlinkages’. The literature on contagion takes two approaches: examining direct linkages and indirect balance-sheet linkages. In looking for contagious effects via direct linkages, research by Allen and Gale (2000) studies how the banking system responds to contagion when banks are connected under different network structures. Banks perfectly insure against liquidity shocks by exchanging interbank deposits. The connections created by swapping deposits, however, expose the system to contagion. The authors show that incomplete networks are more prone to contagion than complete structures. Better-connected networks are more resilient since the losses incurred on one bank’s portfolio are shared with more banks through interbank agreements. To show this, they take the case of an incomplete network where the failure of a bank may trigger the failure of the entire banking system. They prove that, for the same set of parameters, if banks are connected in a complete structure, then the system is more resilient with regard to contagious effects.

The research that followed, although using stylised models, captured well the network externalities created from individual bank risk. Freixas, Parigi and Rochet (2000) consider the case of banks that face liquidity shocks due to uncertainty about where consumers will withdraw funds. In their model, the connections between banks are realised through interbank credit lines that enable these institutions to hedge regional liquidity shocks. As in Allen and Gale’s study, more interbank connections enhance the resilience of the system to the insolvency of a particular bank. One drawback is that this weakens the incentives to close inefficient banks. Moreover, the authors find that the stability of the banking system depends crucially on how many depositors choose to consume at the location of a bank that functions at a money centre.
Concerned with the optimal financial network, Leitner (2005) constructs a model where the success of one agent’s investment in a project depends on the investments of other agents to which she is linked. Since endowments are randomly distributed across agents, any given agent may not have enough cash to make the necessary investment. In this case, agents may be willing to bail out other agents to prevent the collapse of the whole network. Leitner examines the design of optimal financial networks that minimise the trade-off between risk sharing and the potential for collapse. In a related paper, Kahn and Santos (2008) investigate whether banks choose the optimal degree of mutual insurance against liquidity shocks. They show that when there is a shortage of exogenously supplied liquidity, which can be supplemented by bank liquidity creation, banks generally fail to find the correct degree of interdependence. In aggregate, they become too risky.

Dasgupta (2004) also explores how linkages between banks, represented by crossholdings of deposits, can be a source of contagious breakdowns. The study examines how depositors who receive a private signal about banks’ fundamentals may wish to withdraw their deposits if they believe that enough other depositors will do the same. To eliminate the multiplicity of equilibria the author uses the concept of global games. Dasgupta isolates a unique equilibrium, depending on the value of the fundamentals. In the same spirit, Brusco and Castiglionesi (2007) show that there is a positive probability of bankruptcy and propagation of crises across regions when banks keep interbank deposits and that they may engage in excessive risk taking if they are insufficiently capitalised.

Parallel to this literature, other researchers apply network techniques developed in mathematics and theoretical physics to study contagion. For instance, Eisenberg and Noe (2001) investigate default by firms that are part of a single clearing mechanism. First, the authors show the existence of a clearing payment vector that defines the level of connections between firms. Next, they develop an algorithm that allows them to evaluate the effects that small shocks have on the system. This algorithm produces a natural measure of systemic risk based on how many waves of defaults are required to induce a given firm in the system to fail. Similarly, Afonso and Shin (2008) use lattice-theoretic methods to study liquidity and systemic risk in high-value payment systems, such as those used for the settlement of accounts receivable and payable among industrial firms and interbank payment systems. Gai and Kapadia (2010) develop a model of contagion in financial networks using techniques similar to those in the epidemiological literature on the spread of disease in networks to assess the fragility of the financial system. As with Allen and Gale (2000), they find that greater connectivity reduces the likelihood of widespread default. However, shocks may have a significantly larger impact on the financial system when they occur.

The effects of informational contagion are investigated in a paper by Acharya and Yorulmazer (2008b). They show that banks engage in herding behaviour in order to minimise the information spillover from bad information about other banks on their own financing costs. In an otherwise standard banking model with two banks that have access to risky loans and get their funding from risk-averse depositors, the return on each bank’s loans is determined by a common systematic component, say an industry effect, and an idiosyncratic component. The exposure of each bank’s loan return to the systematic and idiosyncratic factors is common knowledge. The ex post performance of each bank’s loan portfolio is also observed by the public. However, what is not observed by the economic agents is the exact realisation of the systematic and idiosyncratic..
factors. The banks can choose their lending structure, in particular whether they want to lend to similar industries, which implies a high level of interbank correlation. However, lending to similar industries has a negative impact on the lending margin because competition in the loan market increases.

Given this structure, Acharya and Yorulmazer (2008b) show that the cost of borrowing for a bank increases when depositors can observe bad news about another bank, since such news conveys bad information about the common factor. In particular, the increase in a bank’s borrowing costs relative to a situation where one can observe good news about other banks is greater when bank loans have less commonality. Therefore, banks have incentives to herd and undertake similar investment in order to minimise the expected costs of borrowing. Only the erosion of lending margins that arises if banks lend into similar industries counteracts this herding force. However, as long as the competitive effect is not too large, banks will herd even if this leads to productive inefficiency due to underinvestment in profitable projects in other industries of the economy.

Since herding behaviour and contagion effects lead to a higher number of systemic banking crises, an interesting issue is how one should design the resolution of bank failures when many banks fail. This question is taken up by Acharya and Yorulmazer (2008a). They argue that during systemic banking crises, a private sector resolution for bank failures does no good since it leads to allocative inefficiency. The reason is ‘cash-in-the-market’ pricing for bank assets in liquidation. The acquisition of bank assets in liquidation is no problem if only a few banks fail since then these banks can be acquired by the surviving banks. However, in the case of a larger banking crisis with a high number of bank failures, the surviving banks typically have limited liquidity. This implies that they can acquire all the assets of the failed banks only at fire sale prices; there are too many banks to liquidate. However, the resulting ‘cash-in-the-market’ pricing attracts liquidity-endowed investors from outside the banking sector. Since these investors are quite often not the most efficient users of these assets, the wrong agents end up purchasing failed banks’ assets.

Acharya and Yorulmazer (2008a) also analyse which type of regulatory intervention might be optimal for avoiding allocative inefficiencies. An ex post optimal bailout policy is typically suboptimal from an ex ante viewpoint since it induces banks to herd, as in Acharya and Yorulmazer (2008b). Anticipating such a bailout policy, banks have incentives to lend funds to similar industries or to bet on common risks in order to increase the likelihood of being bailed out. In turn, such behaviour increases the probability of experiencing systemic crises. However, there is a solution to this time-inconsistency problem. The government should provide liquidity by subsidising surviving banks for the purchase of failed banks. This liquidity provision policy gives banks incentives to differentiate ex ante, yet can be designed to be ex post equivalent to the bailout policy.

Allen and Carletti (2006) also rely on ‘cash-in-the-market’ pricing in their analysis of contagion effects. They are particularly interested in answering the question of whether the introduction of certain financial innovations, such as credit risk transfer instruments, creates inter linkages between different financial sectors that increase the danger of contagion. They focus on the structure of liquidity shocks that hit the banking sector in order to determine whether contagion results. When banks are not hit by idiosyncratic liquidity shocks, and therefore face a uniform demand for liquidity, they keep a sufficient amount of the short-term asset and do not need to raise additional liquidity in the market. In this case credit risk transfer is beneficial as it improves risk
sharing across different sectors in the finance industry. However, when banks do face idiosyncratic shocks, a negative spillover effect can arise because of credit risk transfer.

There was some interesting empirical research on the functioning of interbank markets before the crisis. One of the first studies was conducted by Furfine (2001). By examining the pricing of interbank lending agreements, the paper investigates whether banks really monitor other banks. The main empirical finding is that the interest rate charged on federal funds transactions reflects, at least in part, the credit risk of the borrowing institution in the sense that borrowing institutions with higher profitability, higher capital ratios, and fewer problem loans pay lower rates. Furthermore, the size and relative importance of the trading institution has a negative impact on the interest rate charged for overnight borrowing. Both results suggest that banks can distinguish credit risk among peers and price it accordingly, therefore effectively monitoring other banks.

In a study conducted by Fecht, Nyborg and Rocholl (2011) which is based on 78 consecutive repo auctions by the Eurosystem between June 2000 and December 2001, the authors confirm the result that the price of liquidity systematically depends on bank characteristics such as size and liquidity position, and on market conditions. They find that a greater imbalance in liquidity positions across banks is associated with a rise in the price of liquidity. In particular, banks that are in need of liquidity are affected by the need to pay a higher price in times when there are overall greater imbalances in liquidity positions. This suggests that, even in relatively normal times in the interbank market, liquidity squeezes occur and needy banks are most affected by the potential for a squeeze. Imperfections in the market for liquidity appear to be an enduring and pertinent feature of modern financial systems.

Furfine (2002) considers the functioning of the interbank market in the second half of 1998 when Russia effectively defaulted on its sovereign debt and the hedge fund Long Term Capital Management needed to be rescued. Furfine shows that in this case the interbank market performed well and was not seriously disrupted by other market developments.

These empirical studies of liquidity in interbank markets suggest that the interbank market worked well, even in circumstances of stress. Unfortunately, this conclusion turned out to be incorrect in terms of its relevance for the current crisis.

There was substantial interest in looking for evidence of contagious failures of financial institutions before the crisis. Most of these papers use balance sheet information to estimate bilateral credit relationships. Based on these estimates, the stability of the banking system is tested by simulating the breakdown of a single bank. Upper and Worms (2004) use this methodology. They estimate a matrix of bilateral credit relationships for the German banking system and then simulate the failure of a single bank. They find broad scope for contagion. However, the prevailing financial safety net in Germany, that is, institutional guarantees for savings banks and cooperative banks, reduces the danger of contagion considerably. But even so, the failure of a single bank could lead to a breakdown of up to 15 per cent in terms of bank assets.

There are many other empirical studies that assess the danger of contagion in interbank markets (see Upper (2011) for a survey). For most countries, the simulations suggest that contagious defaults are unlikely, but at least in some countries they cannot be fully ruled out. If contagion does take place, then the simulations typically indicate a breakdown of a substantial part of the banking system.
With hindsight, the result that contagion was unlikely should have been treated with caution. First, since most of the studies focused on simulating the failure of a single bank for idiosyncratic reasons, this was not the scenario most relevant for supervisors. The analysis of the effects of shocks that affect several banks simultaneously is more relevant for understanding systemic risk and should have been studied more. Exceptions were the studies by Elsinger, Lehar and Summer (2006a, 2006b). The second problem was that the simulations conducted typically did not model the price effects of bank failures. However, as argued by Cifuentes, Ferrucci and Shin (2005) and others, these price effects are the main transmission mechanism for contagion.

2.2 Views after the crisis

The crisis led to an evolution in the financial stability literature. With a few exceptions, such as the run on Northern Rock, withdrawals from banks and the hoarding of cash were not significant problems. The widespread use of guarantees such as deposit insurance by governments prevented such runs. In fact, deposits turned out to be one of the most stable forms of finance.

Interbank markets, however, did not function well during the crisis. This turbulence in interbank markets has led to some very interesting recent contributions. These include Acharya, Gromb and Yorulmazer (2012), Freixas and Jorge (2008), Heider, Hoerova and Holthausen (2009), Diamond and Rajan (2011) and Acharya, Gale and Yorulmazer (2011).

Acharya et al (2012) model the interbank markets as being characterised by moral hazard, asymmetric information, and monopoly power in times of crisis. They show that in such a situation a bank with surplus liquidity has bargaining power vis-à-vis deficit banks which need liquidity to keep funding projects. Surplus banks may strategically provide insufficient lending in the interbank market in order to induce inefficient sales of bank-specific assets by the needy banks, which results in an inefficient allocation of resources. The role of the central bank is to provide an outside option to the deficit bank for acquiring the needed liquidity.

Freixas and Jorge (2008) examine how financial imperfections in the interbank market affect the monetary policy transmission mechanism. In their model, firms face liquidity shocks and rely on bank credit to raise external finance. Through this channel, firms’ shocks result in a demand for credit and a liquidity shock for the banks. As a buffer against liquidity shocks, banks hold assets and liquid securities. Since banks hold different amounts of securities and face different liquidity shocks, there is a role for an interbank market to trade reserves. However, asymmetric information in the interbank market induces an equilibrium with quantity rationing in the bank loan market since the interbank market is unable to efficiently channel liquidity to solvent but illiquid banks. As a consequence, monetary transmission might have a strong effect because tightening monetary policy forces banks with less liquidity to cut down on their lending. In addition, liquidity reserves condition the banks’ reaction to monetary policy.

In a similar vein, Heider et al (2009) analyse the functioning of interbank markets. They build a model in the spirit of Diamond and Dybvig (1983). As banks face individual liquidity shocks, there is a role for an interbank market in which banks with surplus liquidity can lend to those with a liquidity shortage. An interbank loan may not be repaid, however, because the long-term investment is risky, thus giving rise to counterparty risk. Asymmetric information about counterparty risk can elevate interbank market spreads and, in extreme situations, lead to a total breakdown of the
interbank market. In the case of such severe adverse selection problems, either all the lenders in the market prefer to hoard liquidity despite high interest rates, or all the borrowers drop out because they find the interest rates too high.

Diamond and Rajan (2011) relate the seizing up of term credit to an overhang of illiquid securities. When banks have a significant quantity of assets with a limited set of potential buyers, shocks in future liquidity demands may trigger sales at fire sale prices. The prospect of a future fire sale of the bank’s assets depresses their current value. In these conditions, banks prefer to hold onto the illiquid assets and risk a fire sale and insolvency than sell the asset and ensure their own stability in the future. This reflects that the states in which the depressed asset value recovers are precisely the states in which the bank survives. In turn, this creates high expected returns to holding cash or liquid securities across the financial system and an aversion to locking up money in term loans.

Acharya et al (2011) show that freezes in markets for rollover debt, such as asset-backed commercial paper, depend on how information about the quality of the asset is revealed. When there is a constant probability that ‘bad news’ is revealed each period, the value of the assets is high in the absence of bad news. By contrast, when there is a constant probability that ‘good news’ is revealed each period, the value of the assets is low in the absence of good news. In the latter scenario, the debt capacity of the assets is below the fundamental value and is decreasing in the liquidation cost and frequency of rollovers. In the limit, as the number of rollovers becomes unbounded, debt capacity goes to zero even for an arbitrarily small default risk.

A closely related literature focuses on the operation of markets when agents have liquidity shocks and trade in markets. The difference is that there are no banks or other intermediaries. A number of interesting results are derived. Huang and Wang (2009, 2010) show that purely idiosyncratic and non-fundamental shocks can cause market crashes if capital flow is costly. Agents trade to smooth out idiosyncratic shocks to their wealth. Since there is no aggregate uncertainty, their trades will be perfectly synchronised and matched, and there will be no need for liquidity if market presence is costless. In this case, the market-clearing price always reflects the fundamental value of the asset, and idiosyncratic shocks generate trading but have no impact on prices. In contrast, when market presence is costly, the need for liquidity arises endogenously and idiosyncratic shocks can affect prices via two channels: first trading becomes infrequent, which makes traders more risk averse; and second, the gains from trading for potential sellers are always larger than the gains from trading for potential buyers. The asymmetry in their appetite to trade leads to order imbalances in the form of excess supply, and the price has to decrease in response.

Two studies isolate illiquidity risk from other confounding effects. Morris and Shin (2009) define ‘illiquidity risk’ as the probability of a default due to a run when the institution would otherwise have been solvent. They show this differs from ‘asset insolvency risk’, which is the conditional probability of default due to a deterioration in asset quality in the absence of a run by short-term creditors, and ‘total credit risk’, which is the unconditional probability of default, due to either a (short-term) creditor run or (long-run) asset insolvency.

Brunnermeier and Pedersen (2009) distinguish between market liquidity and funding liquidity. Market liquidity reflects how difficult it is to raise money by selling an asset, instead of borrowing against it. Traders provide market liquidity, and their ability to do so depends on their availability of funding. Conversely, traders’ funding, i.e. their capital and margin requirements, depends on
the assets’ market liquidity. They show that, under certain conditions, margins are destabilising and market liquidity and funding liquidity are mutually reinforcing, leading to liquidity spirals.

Another explanation for market freezes relies on asymmetric information. Bolton, Santos and Scheinkman (2011) provide a theory of liquidity provision with asymmetric information in which there is an adverse selection problem due to the superior information that intermediaries have about the assets they hold. When intermediaries sell assets they must do so at a discount that becomes greater the longer they hold an asset. If an intermediary is hit by a liquidity shock, the problem it faces is whether to sell its assets now at a discount or to try to ride out the crisis. In doing so, the intermediary runs the risk of having to sell at a greater discount if the crisis lasts longer than expected. In the immediate trading equilibrium, intermediaries sell assets immediately to ensure they have enough liquidity. In the delayed trading equilibrium, intermediaries try to ride out the crisis and only sell if they are forced to. For some parameter values only the immediate trading equilibrium exists, while for others both do, and in this case the delayed trading equilibrium is Pareto superior.

One important empirical question that arises out of all these theoretical studies is how the liquidity of the financial system, and in particular that of banks, should be measured? Berger and Bouwman (2009) have suggested a method for measuring liquidity created by the banking system and have applied this to the case of the United States. They start by classifying all bank liabilities together with off-balance sheet items as liquid, semi-liquid and illiquid. By assessing weights for these three categories, they calculate the amount of liquidity created by the banking system. They show that liquidity increased every year between 1993 and 2003, almost doubling during the period. By applying this measure, similar results are found by Rauch et al (2010) for the German savings bank sector.

In a subsequent paper, Berger and Bouwman (2013) use their measure of liquidity to examine the relationship between liquidity and crises by focusing on the sample period from 1984 until 2008. They find that banking crises were preceded by abnormally high liquidity creation. Hence, it is important to understand how liquidity is created and how this relates to crises.

3. The Funding of Firms

3.1 Views before the crisis

Liquidity problems not only affect financial institutions but also have a significant impact on firms. They can lead to firms cutting back their investments and operations and as a result there can be significant macroeconomic impacts. An important early contribution that analyses firms’ liquidity risk management is Holmström and Tirole (1998). The authors consider the demand for and supply of liquid assets based on the insight from modern corporate finance that, in general, some part of a firm’s income stream cannot be pledged to investors. Several reasons can be given for why part of the income from investment would be non-pledgeable; for instance, it might be due to the missing participation of investors in certain markets, or due to imperfect information which reduces the potential set of financial claims. However, one key implication of non-pledgeability is that firms as well as consumers can rely on liquidating only part of their wealth whenever they need funds. Hence, to cope with negative financial shocks, they must hoard liquid
assets or contract in other ways for the provision of liquidity. Accordingly, they are willing to pay a premium for financial services. Since the supply of liquid assets is constrained by the fact that only part of the return from productive assets can be pledged, non-pledgeability also reduces the amount of wealth in the economy which investors could use for the future financing of firms. In sum, non-pledgeability gives rise to a demand for stores of value across periods, as well as across future states of nature.

Based on this framework, Holmström and Tirole investigate three ways by which the private sector can meet future liquidity needs on its own: by issuing new claims; by obtaining a credit line from a financial intermediary; and by holding claims on other firms. They show that the private sector provides enough liquidity if the corporate sector is a net borrower and there is no aggregate shock. However, even without aggregate uncertainty, financial intermediaries that hoard financial assets and allocate liquidity to firms in need are required to support such a second-best plan. They have to coordinate the use of scarce liquidity by creating a sufficiently rich set of contingent claims, a function which a stock and bond market in general cannot fulfil. The reason is that, because of limited commitment possibilities, there is an insufficiently rich set of financial instruments and therefore liquidity is wasted. If all firms are hit by the same macroeconomic shock, or more generally if there is sufficient aggregate uncertainty, the private sector cannot provide enough liquidity on its own.

Holmström and Tirole show that in such a situation financial instruments that are originated outside the private sector can improve productive efficiency by facilitating access to liquidity and lowering its cost. They argue that a government can provide these instruments by issuing bonds that commit future consumer income. Since the government has the power to enforce tax payments, it can commit funds on behalf of consumers who are unable to commit their future endowments, unless these endowments are backed by marketable assets. Through its taxation power, a government can issue securities at a premium, thereby creating liquidity for the corporate sector. The authors show that the government should loosen liquidity when the economy is hit by a high negative aggregate liquidity shock, and should tighten liquidity when the shock is low.

The setup in Holmström and Tirole (1998) has proved to be extremely useful for analysing a broad range of questions. For instance, Holmström and Tirole (2011) study how such a theory can explain the pricing of assets, the role of liquidity management, and real investment. They also consider how this theory relates to some classic issues in macroeconomics and international finance, such as the question why international markets cannot meet a country’s liquidity needs even though there are more than enough financial instruments for saving and insurance, and that of why inefficiencies arise when firms do not coordinate either the use or the acquisition of international liquidity.

Gorton and Huang (2004) also consider the situation where firms face liquidity risk. However, they use a framework that is quite different from that of Holmström and Tirole (1998) and also apply a different notion of liquidity. In their general equilibrium model, the focus is on the transferability of distressed projects, which is the motivation for immediate selling and borrowing. In Holmström and Tirole (1998), and many other papers such as Bolton et al (2011), and Diamond and Rajan (2001, 2006), a liquidity shock is typically modelled as a sudden need for funds. Here, by contrast, a shock means that long-term projects have a low value, i.e. are distressed, which may induce owners of these projects to engage in value-destroying activities. In other words, there is moral hazard. Such
a negative shock to the value of assets in the hands of an entrepreneur creates a potential need to sell the control rights to the project. Only a recapitalisation of such a project by selling it in a ‘liquidation’ market to agents with enough available liquid assets can lead to a situation where the new owners will not engage in value-destroying activities. In anticipation of these opportunities in the market for liquidation, entrepreneurs choose at the first date whether to be buyers or sellers in the secondary market. As buyers, they invest at the first date in short-term projects, with the only purpose of obtaining liquid assets to possibly buy a distressed long-term project at a later date. As sellers, they invest in a risky long-term project that is socially more efficient, but illiquid. In that way, the supply of liquidity is endogenised.

Gorton and Huang argue, however, that the private provision of liquidity is typically costly for a society. It would be better if, instead of projects being sold in the liquidation market, original lenders were willing to forgive the debt of borrowers with low-value projects. But, debt forgiveness is not always in the interest of lenders. If it is not in their interest, liquidity will, even though socially inefficient, be hoarded because this is the only way to recapitalise projects. Therefore, they suggest one should analyse whether the government could improve welfare by supplying liquidity. Similarly to Holmström and Tirole (1998), the authors show that the government can use its taxation power on the owners of illiquid high-value projects to recapitalise owners of low-value projects. The state can issue securities backed by tax revenue collected from these entrepreneurs at a later time, to subsidise low-value projects. By extending this analysis to consider possible systemic risk in the banking system, the authors investigate whether bank bailouts are efficient. Building on the same argument, they provide conditions under which the government can improve welfare by bailing out banks. This is the case if, when banks suffer a negative shock to their capital, they are subject to moral hazard by not being interested in recapitalising their borrowers’ projects. Anticipating this, too much liquidity will be privately supplied which reduces welfare.

The papers by Holmström and Tirole (1998) and Gorton and Huang (2004) complement each other. While the first analyses why it is important to have a stock of government-supplied liquidity in the economy and why this liquidity should be actively managed, Gorton and Huang in addition offer a rationale for bank bailouts by the government. They show that this type of intervention can improve efficiency whenever there is the potential problem of banks engaging in moral hazard.

A related paper is Kiyotaki and Moore (2005) who examine the interaction between liquidity, asset prices, and aggregate economic activity. They investigate in a general equilibrium framework the role of liquid assets for resource allocation. Liquid assets, which they also call monetary assets, are defined as any asset class that can be readily sold in the market and, therefore, serves as a medium of exchange since it circulates among many agents as a means of short-term saving. This allows Kiyotaki and Moore to analyse the conditions under which the circulation of liquid assets is essential for the smooth running of the economy. The terms liquid assets and monetary assets will be used interchangeably since all assets are real and not denominated in cash.

Output is produced from two types of assets, capital and land, with the usual assumption that the capital stock can be built up through investment, whereas the supply of land is fixed. They, however, depart from the usual representative agent framework by assuming that at each point in time only a fraction of agents has access to productive investment opportunities, even though agents are equally likely to find investment opportunities in the future. They also assume
incomplete markets in the sense that the agents have no possibilities to buy *ex ante* insurance contingent on the arrival of investment opportunities. Furthermore, the capital stock is an asset with limited liquidity since, at the time of investment, agents can sell only a fraction of their capital stock. This type of financial friction is in its impact similar to the non-pledgeability assumption in Holmström and Tirole (1998) or the limited commitment assumption in Diamond and Rajan (2001). Therefore, investing agents may face binding liquidity constraints. Only land is a liquid or monetary asset in the sense that agents can pledge the full value of their land holdings at the time of investment.

They show that in such an economy the circulation of monetary assets is essential for resource allocation since only this allows a smooth transfer of purchasing power from those without an investment opportunity to those with such an opportunity. Second, this circulation of liquid assets is even more important if each agent rarely has an investment opportunity, if investing agents can only pledge a small fraction of their capital, and if the income share of land is small relative to capital.

The model delivers an explanation of the liquidity premium that liquid assets typically command. If investing agents anticipate liquidity constraints at the time of investment, they have strong incentives to hold the liquid assets in their portfolios, even if the return on the liquid assets is lower than the time preference rate. Furthermore, the model generates an interesting feedback mechanism between asset prices and aggregate economic activity. As in the standard framework, higher future expected dividends translate into higher asset prices. However, here, higher asset prices also lead to higher liquidity, which eases the transfer of resources from savers to investors, and encourages aggregate investment and production. Thus, the model can explain the interaction between asset prices, liquidity, and economic activity. In particular, it delivers an account of the large fluctuations in asset prices and aggregate output that are typically observable in financial crisis, and in ‘normal times’ though to a smaller extent. However, since fiat money has no particular role in the framework, it is, at least in the specific setting developed in Kiyotaki and Moore (2005), not possible to analyse the impact of monetary shocks or monetary instruments. Kiyotaki and Moore (2008) provide an interesting extension of the model, where they study how aggregate production and asset prices fluctuate with shocks to productivity and liquidity. Thereby, they also examine the role of government policy, which basically through open market operations, changes the mix of assets held by the private sector.

An interesting example of inefficiencies created by the working of the financial system is analysed in the paper by Lorenzoni (2008) which raises the question of why inefficient credit booms can emerge. Financial frictions, experienced by both borrowers (entrepreneurs) and lenders (consumers), are the fundamental source of inefficiency. Both have limited ability to commit to future repayments which implies that entrepreneurs face (external) financial constraints and consumers cannot fully insure entrepreneurs against aggregate liquidity shocks *ex ante*, as in Holmström and Tirole (1998). Lorenzoni motivates his analysis about the (in)efficiency of credit booms with the observation that, in particular in the last two decades, many developed and emerging countries have experienced episodes of credit expansion which were typically followed by a financial crisis with a collapse in asset prices. He analyses the conditions under which a credit boom arises, why the credit boom may be inefficient from an *ex ante* perspective, and whether any intervention is warranted. He shows that excessive borrowing activity by entrepreneurs can
arise in equilibrium. This, in turn, leads to an excessive contraction in investment activity and asset prices if the crisis takes place, even though all entrepreneurs are rational and correctly perceive the risks and rewards associated with different financial decisions. Key to the result is a pecuniary externality which arises from the combination of the financial constraints mentioned above with a competitive market for real assets.

The model uses a three-period framework. In the first period, entrepreneurs have to borrow funds for productive investment. In the second period, investment returns are subject to an aggregate shock. When a bad shock hits, entrepreneurs face operational losses, which means that they have to sell part of their assets to finance these losses. Assets are sold in a competitive market where they are bought by a traditional sector, i.e. other less productive entrepreneurs. In essence, this setup is closely related to Gorton and Huang (2004). In the third period, consumption takes place and capital fully depreciates.

Although entrepreneurs have access to state-contingent contracts, in the sense that they can *ex ante* decide how much to borrow in the first period and how much to repay in different states of the world in the following periods, an inefficiency will arise. Basically, entrepreneurs face the following trade-off: if they invest more in the first period, they can earn higher returns if the good shock is realised. However, they incur larger losses if the bad shock hits. Even with these state-contingent contracts and despite being fully rational, the atomistic entrepreneurs do not take into account the general equilibrium effect of asset sales on asset prices. They do not internalise this pecuniary externality. Therefore, from a social efficiency viewpoint, they invest too much. Accordingly, a social planner should reduce aggregate investment *ex ante*, because this reduces the amount of asset sales in the bad state. This has a positive effect on asset prices and leads to a reallocation of funds to more productive uses, i.e. from the traditional sector to the entrepreneurial sector, which due to the presence of financial frictions leads to a welfare gain. In essence, the paper formalises the notion that credit booms might be inefficient because they lead to higher systemic risk in the economy. Here, systemic risk increases because more individual borrowing by entrepreneurs creates a pecuniary externality which results in higher fluctuations of asset prices. Again, we have a general equilibrium feedback mechanism between financial distress and asset prices that drives the results.

### 3.2 Views after the crisis

Aggressive monetary policy in most countries meant that, with some exceptions, firms (particularly large corporations) did not face funding problems for a protracted period. This was in contrast to banks and other financial institutions that faced major problems for protracted periods. There were breakdowns in some corporate funding markets that overlapped with those for financial institutions, such as commercial paper markets.

However, the analysis of Holmström and Tirole (1998) and Gorton and Huang (2004) were relevant to the crisis in terms of the shortage of risk-free assets. Gorton (2008) argues that there was a large demand for risk-free assets that led to the creation of securitised mortgages. Although these were rated as very safe by the rating agencies, they in fact turned out to be risky and this contributed greatly to the severity of the crisis.
4. Real Estate Markets

4.1 Views before the crisis

There is extensive evidence that the most important cause of banking crises is real estate booms and busts. Herring and Wachter (1999) document a wide range of boom and bust real estate cycles and their effects on banks. These episodes include Boston in the 1970s and 1980s, Sweden in the 1980s and 1990s, the Japanese bubble of the same period, and Thailand in the 1990s.

Reinhart and Rogoff (2009, Chapter 13) summarise a broad range of episodes where real estate played an important role in causing banking crises. As shown in Table 1, these include the ‘Big Five’ in advanced economies (Spain in the 1970s, Norway in the 1980s, Sweden, Finland and Japan in the 1990s) and the ‘Big Six’ in the Asian Crisis in 1997 (Hong Kong, Indonesia, Malaysia, the Philippines, South Korea and Thailand). Other examples from emerging countries are Colombia in 1998 and Argentina in 2001. Two interesting historical episodes are Norway in 1898 and the United States in the Great Depression. In Chapter 16 they provide evidence of the important role real estate played in many countries during the Great Depression.

Table 1: Real Housing Price Cycles and Banking Crises
From Reinhart and Rogoff (2009)

<table>
<thead>
<tr>
<th>Economy</th>
<th>Crisis date</th>
<th>Peak</th>
<th>Trough</th>
<th>Duration of downturn</th>
<th>Magnitude of decline Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced economies: The Big Five</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>1991</td>
<td>1989:Q2</td>
<td>1995:Q4</td>
<td>6 years</td>
<td>−50.4</td>
</tr>
<tr>
<td>Japan</td>
<td>1992</td>
<td>1991:Q1</td>
<td>Ongoing</td>
<td>Ongoing</td>
<td>−40.2</td>
</tr>
<tr>
<td>Norway</td>
<td>1987</td>
<td>1987:Q2</td>
<td>1993:Q1</td>
<td>5 years</td>
<td>−41.5</td>
</tr>
<tr>
<td>Spain</td>
<td>1977</td>
<td>1978</td>
<td>1982</td>
<td>4 years</td>
<td>−33.3</td>
</tr>
<tr>
<td>Asian crisis: The Big Six</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>1997</td>
<td>1997:Q2</td>
<td>2003:Q2</td>
<td>6 years</td>
<td>−58.9</td>
</tr>
<tr>
<td>Indonesia</td>
<td>1997</td>
<td>1994:Q1</td>
<td>1991:Q1</td>
<td>5 years</td>
<td>−49.9</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1997</td>
<td>1996</td>
<td>1999</td>
<td>3 years</td>
<td>−19.0</td>
</tr>
<tr>
<td>Philippines</td>
<td>1997</td>
<td>1997:Q1</td>
<td>2004:Q3</td>
<td>7 years</td>
<td>−53.0</td>
</tr>
<tr>
<td>South Korea</td>
<td>1997</td>
<td>2001:Q2</td>
<td>1999</td>
<td>4 years</td>
<td>−20.4</td>
</tr>
<tr>
<td>Other emerging economies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>2001</td>
<td>1999</td>
<td>2003</td>
<td>4 years</td>
<td>−25.5</td>
</tr>
<tr>
<td>Colombia</td>
<td>1998</td>
<td>1997:Q1</td>
<td>2003:Q2</td>
<td>6 years</td>
<td>−51.2</td>
</tr>
<tr>
<td>Historical episodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>1898</td>
<td>1899</td>
<td>1905</td>
<td>6 years</td>
<td>−25.5</td>
</tr>
<tr>
<td>US</td>
<td>1929</td>
<td>1925</td>
<td>1932</td>
<td>7 years</td>
<td>−12.6</td>
</tr>
</tbody>
</table>

Sources: Bank for International Settlements; individual economy sources as described in Reinhart and Rogoff (2009, Data Appendixes)
Finally, Crowe et al (2011, Text Table 1, p 5) give data on the relationship between real estate boom and bust cycles, banking crises, credit crunches and macroeconomic performance, using a sample of 40 countries. For example, more than two-thirds of the 46 systemic banking crises for which house price data are available were preceded by boom and bust episodes. In addition, 35 out of 51 boom and bust episodes were followed by a banking crisis.

Real estate clearly played an important role in the current crisis. Figure 1 plots nominal housing prices in Ireland, Spain and the United States. It can be seen that the boom and bust in Ireland was particularly large and was the cause of Ireland’s severe banking crisis. Because the state guaranteed the banks’ debt, the boom-bust cycle has also caused a sovereign debt crisis. This led to the bail-out by other euro area and European Union governments. Spain also had a large run up in real estate prices. So far they have not fallen as much as in Ireland. However, as Taylor (2008) points out, Spain had the biggest housing boom as measured by the change in housing investment as a share of GDP. This is why Spain’s unemployment rate has been so high during the bust phase of the cycle.

![Figure 1: Nominal Housing Prices](image-url)

The plots in Figure 1 suggest that there might be positive serial correlation in housing returns. Case and Shiller (1989), Englund, Quigley and Redfearn (1998) and Glaeser and Gyourko (2007) have investigated this issue using a range of different datasets. They do indeed find evidence of positive serial correlation. For example, Glaeser and Gyourko (2007) find that a $1 increase in real estate prices in one year will on average be followed by a $0.71 increase the following year. Thus, once a real estate boom has started, it is likely that it will persist for some time. Similarly for a bust; once real estate prices have started to fall, this is likely to continue. This feature of real estate prices...
is very different from stock prices, where there is extensive evidence that stock returns are (at least to a first approximation) a random walk. The serial correlation in real estate returns is an important phenomenon that is not well understood.

A striking feature of Figure 1 is that the US boom and bust cycle was much less extreme than that in Ireland and Spain. However, this is misleading because the figures are for the country as a whole. There was very wide variation in experiences in different parts of the country. Figure 2 shows the experiences of the ten cities that make up the S&P/Case-Shiller 10-city index. Two things stand out. The first is that from the mid 1990s until the early 2000s, prices in all ten cities move together. But for the next few years the cities had widely different experiences. Miami and Los Angeles had massive booms and busts, while Denver had a relatively small change in prices. The other cities were somewhere in between. However, in all these cases, interest rates and many other features of the credit market were common. It is not well understood why the experiences were so similar and then so different.

![Figure 2: Nominal Housing Prices in Different US Cities](image)

One of the major debates about the boom and bust episodes in the current crisis is the extent to which the real estate bubbles in these countries were the result of loose monetary policy and global imbalances that led to excessive credit availability. Central banks, in particular the Federal Reserve in the United States, set very low interest rates during the early 2000s to avoid a recession after the bursting of the technology bubble in 2000 and the 9/11 terrorist attacks in
2001. As argued by Taylor (2008), these levels of interest rates were much lower than in previous US recessions relative to the economic indicators at the time captured by the ‘Taylor rule’.

Although the ECB did not set absolute rates as low as those set by the Federal Reserve, the different economies had very different conditions. As Figure 1 shows, Spain and Ireland had very large increases in property prices. For these economies the ECB’s policy was very loose. Figure 3 shows the wide disparity of movements in property prices in Europe. While Ireland and Spain had big run-ups and collapses, other countries like the United Kingdom and Sweden had large run-ups but did not experience big collapses. By contrast, France and Germany, which are quite similar in terms of their industrial structure, and are both members of the euro area, had very different experiences in terms of property prices. This underlines our lack of knowledge concerning the determinants of real estate prices.

Taylor’s position has been quite controversial. For example, Bernanke (2010) has argued that the Taylor rule is sensitive to the choice of inflation measure and to whether actual or forecast inflation and output gaps are used. Once changes in these measures are introduced, it is no longer clear whether interest rates were unusually low given the state of the economy, or whether house prices were unusually high given interest rates and the state of the economy. Bernanke concludes that Taylor’s claim is not persuasive enough. He suggests that what seems to have played a crucial role in setting the stage for the crisis is financial innovation in the form of mortgage contracts and securitisation. Rather than interest rates being set too low, the implications of financial innovation
for monetary policy transmission were not understood by monetary policymakers. This failure, together with weak financial regulation and supervision, set the stage for the crisis.

As Allen and Gale (2000, 2003, 2007) have argued, asset price bubbles are also caused by growth in credit. During the recent crisis, credit expanded rapidly in the countries with low interest rates that were partly the results of global imbalances. In particular, several Asian countries started accumulating large amounts of reserves in the late 1990s, which lowered interest rates and helped fuel the bubble in asset prices. It was the bursting of this bubble that started the crisis; in the summer of 2007, the fall in property prices triggered a fall in price and downgrading of securitised mortgages.

4.2 Views after the crisis

Before the crisis, mortgages in the United States were securitised by both public firms such as Fannie Mae and Freddie Mac, and private firms such as investment banks. Once the crisis started, the private market disappeared and since then it has not reappeared. The only sector that remains in the United States is the public one. In the long run, one of the major issues is how the private sector can be restored.

5. The Role of Central Banks in Funding Markets

5.1 Views before the crisis

5.1.1 The central bank as lender of last resort

At least since the work of Bagehot and the 19th and 20th century interventions by the Bank of England, it has been recognised that central banks have a crucial role to play in the prevention and management of financial crises. In his influential book, *Lombard Street*, Bagehot (1873) laid out his famous principles for how a central bank should lend to banks during a crisis.

- Lend freely at a high rate of interest relative to the pre-crisis period but only to solvent but illiquid borrowers with good collateral (i.e. any assets normally accepted by the central bank).

- The assets should be valued at between panic and pre-panic prices.

- Institutions without collateral should be allowed to fail.

Despite being written over 140 years ago, these principles are still widely quoted and used as the foundation for many central bank policies. However, their validity in terms of modern financial economics has only been considered in a few papers.

Rochet and Vives (2004) is one of the few papers that has recently examined the Bagehot principles. In particular, the authors focus on Bagehot’s assertion that the LOLR should lend to any solvent but illiquid banks. In the past, several authors, such as Goodfriend and King (1988), have dismissed this view as obsolete since in modern interbank markets it cannot be the case that a solvent bank is illiquid. Of course, in light of the recent crisis, one can have serious doubts as to whether such an argument is true. For that reason, it is even more interesting that Rochet and Vives provides a theoretical foundation supporting Bagehot’s doctrine regarding this dimension. An important problem in the banking literature in the spirit of Bryant (1980) and Diamond and
Dybvig (1983) is that the fragility of banks depends crucially on possible coordination failures between depositors that can trigger bank runs. Given the assumption of first-come, first-served, and costly liquidation of long-term assets, there are multiple equilibria, which makes it hard to base any policy recommendations on such a framework. Using the global games approach, Rochet and Vives develop a theory which does not rely on multiple equilibria. Instead, their model produces a unique Bayesian equilibrium that is characterised by a positive probability that a solvent bank cannot get enough liquidity assistance in the market. Hence, in this respect the Bagehot doctrine still has a solid theoretical foundation.

One of the criticisms of the kind of LOLR policy advocated by Bagehot is that it creates a moral hazard problem in the sense of increasing the incentives for banks to take more risk. Repullo (2005) investigates this claim about LOLR lending. By modelling the strategic interaction between a bank and a LOLR, he shows that in general this proposition is not true. He assumes a bank which is funded with insured deposits and equity capital, is subject to capital requirements, and can invest, like in a Diamond-Dybvig framework, in two assets: a safe liquid asset and an illiquid asset, the risk of which will be privately chosen by the bank. Since deposits are randomly withdrawn, the bank is subject to liquidity shocks. Because the bank optimally will not invest all its endowment in liquidity, in case of a large negative withdrawal shock it has to rely on emergency lending from a LOLR to avoid being forced into liquidation. In this setting, Repullo shows that in equilibrium the bank chooses a risk level that is decreasing in the capital requirement and increasing in the penalty rate charged by the LOLR. However, in the case where the LOLR does not charge the penalty rate, there is an irrelevance result regarding the risk choice. Irrespective of the existence of a LOLR, the bank chooses the same level of risk, but the liquidity buffer chosen is lower when a LOLR exists.

5.1.2 Money creation and financial stability

Fiat money does not play a role in most models of banking crises. Typically, banks contract with depositors in real terms, and if government-injected liquidity is essential in preventing a crisis or alleviating an aggregate liquidity shortage, it will be done using appropriate financial and fiscal instruments that have effects in real terms. However, it is apparent from many crisis experiences in the past that monetary policy also seems to be important in crisis situations. A number of papers before the crisis considered the relationship between money and financial stability.

Much of this early literature seeks to explain historical crises that occurred at a time when fiat currency played an important role in the financial system. An early contribution is Champ, Smith and Williamson (1996). They address the issue of why Canada had no banking crises in the late 19th and early 20th centuries while the United States had many. Their explanation is that Canada allowed the amount of money in circulation to expand to meet demand during harvest time while this could not happen in the US financial system. The effect of this difference was that in Canada liquidity shocks could be easily absorbed but in the United States they led to banking panics. Since currency played an important role during this period, the authors use an overlapping generations model with two-period lived consumers to justify the use of currency. The consumers live in two different locations. Instead of random preference shocks as in Diamond and Dybvig (1983), consumers are subject to relocation shocks. Each period a random proportion of young consumers in each location is forced to move to the other location. These shocks are symmetric so that the population in each place remains constant. Banks make risk-free loans, hold
reserves of currency, issue bank notes, and write deposit contracts that are contingent on the proportion of the consumers that relocate. When young consumers relocate they can transport currency or the notes issued by the banks with them but nothing else. The authors show that if the banks are allowed to vary their issuance of notes to accommodate different levels of relocation shocks then there exists a stationary Pareto-optimal equilibrium. In this equilibrium, currency and banknotes are perfect substitutes and the nominal interest rate is zero. However, if the banknote issuance is fixed such that the random relocation demand cannot be accommodated, there will be a banking crisis if the shock is large enough to exhaust the banks’ currency reserves. The authors interpret these two possibilities as being consistent with the Canadian and US experiences from 1880–1910.

Antinolfi, Huybens and Keister (2001) build on the model of Champ et al (1996) by replacing the private issue of banknotes with a LOLR that is willing to lend freely at a zero nominal interest rate. A stationary Pareto-optimal equilibrium again exists but in addition there is a continuum of non-optimal inflationary equilibria. Antinolfi, Huybens and Keister are able to show that these can be eliminated if the LOLR places an appropriately chosen upper bound on the amount that each individual bank can borrow or is willing to lend freely at a zero real interest rate.

Smith (2002) considers a similar model with two-period lived overlapping generations, where spatial separation and random relocation introduces a role for money and banks. He shows that the lower the inflation rate and nominal interest rate, the lower is the probability of a banking crisis. Reducing the inflation rate to zero in line with the Friedman rule eliminates banking crises. However, this is inefficient as it leads banks to hold excessive cash reserves at the expense of investment in higher yielding assets.

Cooper and Corbae (2002) consider a model with increasing returns to scale in the intermediation process between savers and entrepreneurs. This leads to multiple equilibria that are interpreted as different levels of confidence. A calibrated version of the model with low confidence levels is able to match many features of the Great Depression.

As discussed above, Diamond and Rajan (2001) develop a model where banks have special skills to ensure that loans are repaid. By issuing real demand deposits, banks can pre-commit to recoup their loans. This allows long-term projects to be funded and depositors to consume when they have liquidity needs. However, this arrangement leads to the possibility of a liquidity shortage in which banks curtail credit when there is a real shock. Diamond and Rajan (2006) introduce money and nominal deposit contracts into this model to investigate whether monetary policy can help alleviate this problem. They assume there are two sources of value for money. The first arises from the fact that money can be used to pay taxes (the fiscal value). The second is that money facilitates transactions (the transactions demand). They show that the use of money can improve risk sharing since price adjustments introduce a form of state contingency to contracts. However, this is not the only effect. Variations in the transaction value of money can lead to bank failures. Monetary intervention can help to ease this problem. If the central bank buys bonds with money, this changes liquidity conditions in the market and allows banks to fund more long-term projects than would be possible in the absence of intervention.

Allen and Gale (1998) develop a model of banking crises caused by asset return uncertainty with three dates, early and late consumers as in Diamond and Dybvig (1983), and initially, real
contracts. Building on the empirical work of Gorton (1988), it is assumed that at the intermediate date investors receive a signal concerning the return on the banks’ long-term assets. If the signal indicates returns are sufficiently low, the late consumers will withdraw their deposits along with the early consumers and there will be a banking crisis. Allen and Gale go on to show that if contracts are written in nominal terms and a central bank can supply money to commercial banks, the incentive-efficient allocation can be implemented: the central bank gives money to the banks and they then pay this out to depositors. The early depositors use their money to buy goods from early-withdrawing late consumers who then hold money until the final date. Variations in the price level allow risk sharing.

Skeie (2008) develops a standard banking model with nominal contracts and inside money where depositors are subject to preference shocks in the usual way. There is no aggregate liquidity risk or return uncertainty. In contrast to Diamond and Dybvig (1983), Skeie shows that there is a unique equilibrium and it is efficient. If deposits are withdrawn by late consumers at the intermediate date, the price of the consumption good adjusts and this discourages such withdrawals. In order for there to be runs on banks there must be some other friction, such as problems in the interbank market.

5.2 Views after the crisis

The dramatic interventions undertaken by many central banks since the onset of the crisis has led to many contributions trying to understand their role in funding markets. A number of studies focus on their intervention in interbank markets.

Allen, Carletti and Gale (2009) show that the interbank market is characterised by excessive price volatility when there are insufficient opportunities for banks to hedge aggregate and idiosyncratic liquidity shocks. They analyse how the central bank should intervene to restore efficiency. By using open market operations to fix the short-term interest rate, the central bank can prevent price volatility and implement the constrained efficient solution. Thus, the central bank effectively completes the market, a result in line with the argument of Goodfriend and King (1988) that open market operations are sufficient to address pure liquidity risk in the interbank markets. Interestingly, one implication of the model is that situations where banks stop trading with each other can be a feature of the constrained efficient solution implemented by central bank policy if aggregate uncertainty is high. Banks may hoard liquidity because they may need it to meet high aggregate demand. When aggregate demand is low, however, they have enough liquidity to meet idiosyncratic shocks and accordingly do not need the interbank market. As a result the volume in the market falls to zero, but there is no need for central banks to intervene since the freeze is consistent with constrained efficiency.

Freixas, Martin and Skeie (2011) develop a model with aggregate liquidity risk, which like Allen, Carletti et al (2009) also has idiosyncratic liquidity shocks to banks. They suggest that inducing low interbank market rates in states of financial disruptions is an optimal policy response of the central bank. As they argue, a primary role for banks in the presence of incomplete markets is to provide better risk-sharing possibilities and more liquidity than markets. Yet during financial disruption, the banks themselves face considerable uncertainty regarding their own idiosyncratic liquidity needs. Hence, they may have large borrowing needs in the interbank market. They show
that an interbank market can achieve the optimal allocation, which implies efficient risk sharing to consumers and effective insurance for banks against idiosyncratic liquidity shocks. In the optimum, however, the interest rate in this market must be state-contingent and low in states of financial disruption. This suggests a role for the central bank, which in their model can implement the efficient allocation by setting the interest rates in the interbank market.

The main part of Freixas et al (2011) presents a real analysis where they show that there can be multiple equilibria, but that the central bank can determine the interest rate that implements the equilibrium with the efficient allocation. In an appendix, they show that money can be introduced along the lines of Skeie (2008) and the same results hold.

Allen, Carletti and Gale (forthcoming) also focus on inside money. At the initial date the central bank makes an intraday loan to the banks. Money is loaned to the firms that produce the goods, to fund the purchase and the consumers’ endowment. The payment from the firms is made to the consumers’ deposit accounts and the bank uses this money to repay its loan to the central bank. At the intermediate and final dates, the reverse process occurs: banks use the intraday loan from the central bank to repay depositors who use the funds to buy goods from firms. The firms use the money to repay their loans to the bank and the banks in turn repay the central bank. All money balances are held within the banking system in deposit accounts. This feature is important, since it means that, in contrast to the papers using the Champ et al (1996) model, the central bank is free to set the nominal interest rate to control the level of inflation. In practice, fiat currency is used relatively little in modern financial systems. It could be added to the model if it was recognised that for some transactions it is more convenient than debit or credit cards.

In contrast to the previous literature, Allen et al (forthcoming) consider a wide range of risks. The banks’ long-term assets, which can be thought of as loans, are subject to aggregate uncertainty. There is aggregate risk in liquidity shocks to consumers. In addition, there are also idiosyncratic liquidity shocks to individual banks as in Bhattacharya and Gale (1987). In contrast to the Champ et al model, where contracts are state-contingent, deposit contracts are assumed to be fixed in nominal terms. In spite of the multiplicity of risks the authors are able to show that provided the central bank runs an accommodative policy in terms of granting intraday loans, the first-best efficient allocation will result from the decentralised banking system described. Moreover, this allocation is unique. In the efficient equilibrium that we characterise there are no banking crises. Banks never fail, no matter how low asset returns fall, or how large are the aggregate liquidity shocks. The central bank always provides enough money to the banks to allow them to repay their depositors. The central bank can also control the expected level of inflation between the initial date and the first date and the actual level of inflation between the intermediate and final dates.

Central banks adopted many unconventional policies during the crisis. These include direct intervention in some markets like the commercial paper market. Other interventions involved the purchase of non-government securities. Quantitative easing policies involving the purchase of large quantities of government securities have also been widely adopted. An important issue is the extent to which these policies have distorted asset prices. This is another very contentious area. A key remaining issue is the ease with which central banks will exit from these policies.
6. Concluding Remarks

There remains one other important set of funding markets that have operated imperfectly during the crisis. These are government debt markets. There is a large literature on sovereign default. However, it was widely assumed that this was an emerging country problem. The default of Greece in March 2012 showed that this assumption was incorrect. The possibility of sovereign default has had a significant effect on some countries’ funding opportunities, particularly those on the periphery of the euro area. This is a very large topic and raises rather different issues than those considered here, so we do not attempt to discuss funding of governments here.
References


Bagehot W (1873), Lombard Street: A Description of the Money Market, H. S. King, London.


Bernanke BS (2010), ‘Monetary Policy and the Housing Bubble,’ Speech to the Annual Meeting of the American Economic Association, Atlanta, 3 January.


Reinhart CM and KS Rogoff (2009), This Time is Different: Eight Centuries of Financial Folly, Princeton University Press, Princeton.


