

# How Should Bank Liquidity be Regulated?\*

Franklin Allen  
Imperial College London

and

Douglas Gale  
New York University

January 27, 2017

## **Abstract**

Before the crisis, bank regulation relied to a large extent on capital regulation. Liquidity regulation was not widely used. The liquidity problems during the crisis led to calls for liquidity regulation. As a result, the Basel III accord introduced global liquidity standards. An important issue in the construction of such liquidity regulations is the exact nature of the problem they are trying to solve. What is the market failure they are designed to correct? Why is the provision of liquidity that the market provides insufficient? This paper considers the literature analyzing liquidity regulation. There is no wide agreement on the rationale for liquidity regulation.

---

\* Franklin Allen is Professor of Finance and Economics, and Director of the Brevan Howard Centre at Imperial College London. Douglas Gale is Silver Professor and Professor of Economics at New York University. The paper was prepared for the Federal Reserve Bank of Chicago and the European Central Bank's co-sponsored 2016 International Banking Conference. We are grateful for comments from Douglas Diamond.

## **1. Introduction**

One reason why the 2007–2009 financial crisis was so severe and had a global impact was massive illiquidity in many markets, particularly interbank markets. This combined with an extreme exposure of many financial institutions to liquidity needs meant 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 injected huge amounts of liquidity into financial systems.

Before the crisis, bank regulation relied to a large extent on capital regulation. Liquidity regulation was not widely used. The liquidity problems during the crisis led to calls for liquidity regulation. As a result, the Basel III accord introduces global liquidity standards. These comprise a Liquidity Coverage Ratio (LCR) to withstand a stressed funding scenario and a Net Stable Funding Ratio (NSFR) to address liquidity mismatches.

The LCR is a measure of an institution's ability to withstand a severe liquidity freeze that lasts at least 30 days. Liabilities are categorized in terms of the degree of difficulty in rolling them over. Each category is assigned a percentage representing the portion of that liability that remains a source of funding during the next 30 days or is replaced by funds in the same category. Assets are also sorted into categories with each category being assigned a percentage haircut representing the loss that would be incurred if the asset were to be sold in the middle of a severe financial crisis. The LCR is defined as the ratio of High Quality Liquid Assets (HQLA) to total net cash outflows over the next 30 calendar days. The total net cash outflow equals total expected cash outflows minus the minimum of total expected cash inflows and 75 percent of total expected cash outflows. The idea is that the ratio should exceed 100 percent so the financial institution can survive at least 30 days. The LCR regulation was introduced during 2015-16.

The NSFR is designed to reveal risks that arise from significant maturity mismatches between assets and liabilities. It is the ratio of the available amount of stable funding to the required amount of stable funding over a one-year horizon. Stable funding includes customer deposits, long-term wholesale funding, and equity. The required amount of stable funding is calculated by weighting assets (longer-term assets receive higher weights but assets which mature within one year do not necessarily receive a zero risk-weight). Again, the idea is that the ratio exceeds 100%. The NSFR is due to be introduced starting in 2018.

One important issue with these and other liquidity regulations is the exact nature of the problem they are trying to solve. In other words, what is the market failure they are designed to correct? Why is the provision of liquidity that the market provides insufficient? Are the market failures such that the numerical values in the regulations are sensible?

This paper starts in Section 2 with a discussion of a benchmark model due to Allen and Gale (2004a) that builds on the framework developed by Bhattacharya and Gale (1987). The model has both financial institutions and markets and it is shown how the invisible hand leads to an efficient provision of liquidity without any intervention. The key question is then which of the assumptions of the model are not satisfied and lead to market failures. The literature on possible market failures is then discussed.

Section 3 considers the role of the central bank in solving the problems associated with market failures that prevent efficient provision of liquidity by the market. These include Lender of Last Resort (LOLR) facilities and other similar kinds of intervention. Bagehot (1873) laid out central banks' principles of intervention in times of crisis in his famous book *Lombard Street*. Rochet and Vives (2004) provide an interesting analysis justifying these principles while Repullo (2005) considers the moral hazard that such principles can lead to.

These models, like most models in banking, treat banking as a real activity with no role for fiat money. While "real" models have provided valuable insights into the nature of financial fragility, they do not capture important aspects of reality, such as the role of fiat money in the financial system. In practice, financial contracts are almost always written in terms of money. This fact has important consequences for the theory. Because the central bank can costlessly create fiat money in a crisis, there is no reason why the banking system should find itself unable to meet its commitments to depositors. As Willem Buiter (2007) has argued,

“Liquidity is a public good. It can be managed privately (by hoarding inherently liquid assets), but it would be socially inefficient for private banks and other financial institutions to hold liquid assets on their balance sheets in amounts sufficient to tide them over when markets become disorderly. They are meant to intermediate short maturity liabilities into long maturity assets and (normally) liquid liabilities into illiquid assets. Since central banks can create unquestioned money at the drop of a hat, in any amount and at zero cost, they should be the money providers of last resort both as lender of last resort and as market maker of last resort....”

Again a benchmark model is needed to understand the effects of incorporating money into the analysis. Allen, Carletti and Gale (2014) provide such a model. It is shown in the context of this model, that if the central bank accommodates the demand of commercial banks for liquidity then there will be an efficient allocation of resources. The problem of market failures is discussed in the context of this framework.

The final part of the paper, Section 4, considers the policy implications of the analysis of liquidity provision and the role of regulations such as those in Basel III.

## **2. Liquidity Provision in the Financial System**

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) was 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 coinsure 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. 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 and 2007) 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 may have more or less liquidity needs. The second type of uncertainty is aggregate liquidity risk which banks have to face. In some periods, 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 such a framework is the ability of banks to hedge themselves against these liquidity shocks. They show in Allen and Gale (2004a) 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 of course interbank markets, is necessary for an efficient functioning of the financial system. Only 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 (2004b) explore in further detail the ramifications of this framework. By using a simplified version of the general equilibrium model introduced in Allen and Gale (2004a), they investigate the role of liquidity in determining asset prices. The incompleteness of markets leads to an inefficient provision of liquidity by the financial system. This can generate cash-in-the-market pricing or fire sale pricing as it is often referred to, which implies that the prices of long-term safe assets can fall below their fundamental value, and leads to financial fragility, which means that even small shocks can have large price effects.

The intuition for this result is as follows. When markets are incomplete, liquidity provision is achieved by selling assets when liquidity is required. That implies, asset prices are determined by the available liquidity or in other words by the cash in the market in case liquidity is scarce. This is often referred to as a “fire sale.” Due to the incompleteness of markets, liquidity cannot be traded in a state-contingent manner and hence suppliers of liquidity can no longer be compensated for the cost of providing liquidity state by state. Instead they must be compensated on average across all states and that causes inefficiency. Providers of liquidity have the alternative of investing in a productive long-term asset. Therefore holding liquidity is associated with an opportunity cost as this has a lower return than the productive long-term asset. Hence, providers of liquidity must be able to make a profit in some states in order to get compensated for holding the liquidity. However, in equilibrium they can only make a profit in those states where liquidity is scarce. In states of abundant liquidity where banks do not need liquidity, liquidity suppliers have to bear the opportunity cost of holding liquidity. Accordingly it must be the case that in states where banks face high liquidity demand, the equilibrium price of the long-term asset is low enough to compensate the liquidity providers for all the other states where they do not make any profit. Otherwise, they will not hold the liquidity in the first place. Hence, prices are low in states where banks need more liquidity. Asset price volatility is a consequence of compensating liquidity suppliers. But from an efficiency point of view there is a transfer from the banks who need liquidity to the providers of liquidity. Negative insurance and suboptimal risk sharing will be the result. Allen and Gale (2004b) show that the only equilibria that are robust in such a setting involve stochastic consumption as well as volatile asset prices.

There are many deviations from the efficient framework in Allen and Gale (2004a) that constitute market failures. Interbank markets did not function well during the crisis and this has

led to some very interesting recent contributions analysing how they malfunctioned. These include Acharya, Gromb and Yorulmazer (2012), Heider, Hoerova and Holthausen (2009), Freixas and Jorge (2008), Diamond and Rajan (2011), and Acharya, Gale and Yorulmazer (2011).

Acharya, Gromb and Yorulmazer (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, Hoerova and Holthausen (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, Gale and Yorulmazer (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.

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 asset market liquidity. They show that, under certain conditions, margins are destabilizing 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, Steffen, Hackethal and Tyrell (2009) 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 focussing 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 crisis.

Although most of these papers are concerned with understanding the failure of markets and their effect on liquidity provision, they do not typically consider interventions such as liquidity regulations to counter the effects of these failures. Papers that are focused on liquidity regulation are considered in Section 4.

### 3. The Role of Central Banks in Providing Liquidity

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 examined the Bagehot principles in recent times. In particular, the authors focus on Bagehot's assertion that the lender of last resort (LOLR) should lend to 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 about the validity of this argument. For this reason, it is even more interesting that Rochet and Vives (2004) provides a theoretical foundation supporting Bagehot's doctrine regarding this dimension. An important problem in the banking literature in the spirit of Diamond and Dybvig (1983) and Bryant (1980) 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 make 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 this proposition is not true in general. He assumes a bank which is funded with insured deposits and equity capital is subject to capital requirements and can invest, similarly to the Diamond-Dybvig framework, in two assets: a safe liquid asset and an illiquid asset, whose risk 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 the 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.

When there are insufficient opportunities for banks to hedge aggregate and idiosyncratic liquidity shocks, Allen, Carletti and Gale (2009) show that the interbank market is characterised

by excessive price volatility. 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 and Gale (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.

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. There are a number of papers that consider 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 bank note 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 United States experiences from 1880–1910.

Antinolfi, Huybens and Keister (2001) build on the model of Champ, Smith and Williamson (1996) by replacing the private issue of bank notes 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.

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 precommit 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.

Allen, Carletti and Gale (2014) develop a model in which fiat money is issued by the central bank. Deposit contracts and loan contracts are denominated in terms of money and money is used in transactions. In other words, money is both a unit of account and a medium of exchange. In contrast to the previous literature, it is shown that the combination of nominal contracts and a central bank policy of accommodating commercial banks' demand for money leads to first best efficiency in a wide range of circumstances. The result holds when there are aggregate liquidity and asset return shocks and also when there are idiosyncratic (bank specific) liquidity shocks.

Time is represented by a sequence of three dates and, at each date there is a single good that can be used for consumption or investment. Assets are represented by constant returns to scale technologies that allow the consumers' initial endowment of the good to be transformed into consumption at the second and third dates. The short-term asset is represented by a storage technology: one unit of the good invested in this technology yields one unit of the good at the next date. The long-term asset is represented by a technology that requires an investment at the initial date and yields a random return at the final date. The expected return of the long-term asset is greater than the return of the short-term asset.

There is a large number of ex ante identical consumers, each of whom is endowed with one unit of the good at the initial date. At the beginning of the second date, each consumer receives a time-preference shock that makes him either an early consumer, who wants to consume only at the second date, or a late consumer, who wants to consume only at the third date. The proportion of early and late consumers is itself random, an important source of aggregate uncertainty.

The first best allocation is characterized as the solution to a planner's problem. The planner invests the consumers' endowments in a portfolio of short- and long-term assets and then distributes the returns to these assets to the early and late consumers. The portfolio is chosen before the realization of the aggregate state, that is, the fraction of early consumers and the return on the risky asset. The consumption allocation is determined after the realization of the aggregate state and is therefore state contingent. It is shown how this allocation can be implemented using a simple institutional structure and non-contingent nominal contracts.

In the decentralized economy, there are three types of institutions -- a central bank, commercial banks and firms. At the initial date, the central bank makes money available to the commercial banks on an intraday basis at a zero interest rate. The banks make loans to the firms and the firms in turn use the money to buy the consumers' endowments and invest them in the short- and long-term assets. At the intermediate and final dates, the central bank again makes intraday loans to the banks. The banks use this money to pay for depositors' withdrawals. The depositors in turn use the money to purchase goods from the firms. Then the firms use the same money to repay their loans to their banks and the banks use it to repay the central bank. The central bank's policy is passive: at each date it supplies the amount of money demanded by the commercial banks. Commercial banks and firms are assumed to be profit maximizing but in a competitive equilibrium they earn zero profit. Consumers are expected utility maximizers, but in equilibrium

their decision problem is simple: they deposit the money received in exchange for the sale of their endowments at the first date and withdraw and spend all their money at the second or third date, depending on whether they are early or late consumers.

The main result is to show that a competitive equilibrium implements the same state-contingent allocation as the planner's problem, even though deposit contracts represent a fixed claim (in terms of money) on the banks. In spite of the debt-like nature of the deposit contract, it is possible to implement a state-contingent allocation because deposit contracts are written in terms of money. Regardless of the liquidity and asset return shocks, banks are able to meet their commitments as long as the central bank supplies them with sufficient amounts of fiat money. The price level adjusts in response to aggregate shocks in order to clear markets. When the number of early consumers is high, the amount of money withdrawn from the banks is also high and this increases the price level. When the returns on the long asset are low, the supply of goods is also low and this increases the price level. The adjustments in the price level ensure that early and late consumers' receive the efficient, state-contingent levels of consumption.

A central bank policy of passively accommodating the demands of the commercial banks for money is sufficient to eliminate financial crises and achieve the first best. The role of the central bank is simply to provide the necessary money so that each bank can meet withdrawals by its depositors. Price level adjustments lead to the optimal level of real balances and the optimal allocation of consumption at each date.

The baseline model can be extended in a number of ways. Idiosyncratic (bank-specific) liquidity shocks can be introduced without upsetting the efficiency results. The interbank market allows banks to reshuffle money between banks that receive high and low liquidity shocks at the second date so that each bank can meet the required level of withdrawal by its depositors, without

being subject to distress. The process is reversed at the third date, so that banks with a large proportion of late consumers can meet the higher number of withdrawals then. First best efficiency can be achieved by monetary policy alone when the model is extended to allow for idiosyncratic (bank-specific) liquidity risk and multiple periods.

Accommodative monetary policy alone is not always sufficient to achieve efficiency, however. In particular, it does not allow the sharing of idiosyncratic (bank-specific) asset return risk. If the banks' asset-specific returns are observable, the government could introduce an insurance scheme. Alternatively, a private scheme could achieve the same end by securitizing the assets and allowing banks to hold a diversified portfolio of asset backed securities. Such schemes are vulnerable to moral hazard if there is asymmetric information about asset returns. Insuring low returns gives banks an incentive to engage in asset substitution and to misrepresent the realized returns of the assets. Clearly, pooling idiosyncratic return risks is more difficult than implementing an accommodative monetary policy.

Both in the case of real models of central bank interventions and monetary injections, benchmark models have been identified and the role of market failures considered. Although liquidity is the key issue in these analyses, liquidity regulation of banks is usually considered only in passing, if at all. The next section considers contributions to the policy discussion on liquidity regulation.

#### **4. Policy Discussion of Liquidity Regulation**

While there has been a great deal of academic literature on capital regulation there has been little on liquidity regulation, particularly before the crisis. The literature discussed in the previous

sections considered various aspects of liquidity provision including the role of the central bank but did not consider liquidity regulation except in passing.

Rochet (2004, 2008) are early contributions. These papers focus on a number of market failures that can justify liquidity regulation. These include potential problems in payment systems, moral hazard problems at the individual bank level due to opaqueness of assets, and moral hazard at the aggregate level due to expectations of a generalized bailout if there are macro shocks. Rochet argues that while simple liquidity ratios can potentially deal with the first two, more complex regulation based on a banks' exposure to macros shocks may be necessary for the third problem.

Perotti and Suarez (2011) develop a formal model of liquidity regulation based on Pigovian taxes. The basic market failure is due to an externality. Even though each individual bank takes into account its own exposure to refinancing risk, it does not internalize the system-wide effect of its decision. This externality results in too much short term funding and the problem is to mitigate this effect. Banks differ in their ability to extend credit and their incentives to take risk. Depending on which of these types of heterogeneity is dominant, the socially efficient allocation can be obtained with some combination of Pigovian taxes and quantity regulations. When banks differ in credit opportunities, Pigovian taxes are best. However, when they differ in their risk taking incentives, net funding ratios are best. If capital controls can be used as well as liquidity ratios, then taxes can again be optimal.

In a related contribution, Stein (2013) develops a framework where the market failure is that banks do not take into account all the social benefits of increased liquidity reserves in terms of enhanced financial stability and lower costs to taxpayers. The central bank acting as LOLR is one way to solve this problem. However, Stein argues that it is socially costly to use LOLR capacity because it is difficult to distinguish between illiquidity and insolvency. As a result, it

may be better to have liquidity regulation. In addition, it may, in cases where high quality collateral is in short supply, be optimal to price access to the LOLR as well.

In addition to these papers analysing the optimal form of regulation, Bech and Keister (2013) consider the effect of liquidity regulation on the implementation of monetary policy. Since monetary policy is typically implemented by central banks targeting the rate in the market for central bank reserves, liquidity regulation may change the relationship between market conditions and the interest rate. In Bech and Keister's model, this happens because banks worried about violating the liquidity regulation are more likely to seek term funding in the market. This results in a steeper yield curve at short maturities.

Heider, Hoerova and Calomiris (2015) develop an interesting theory of liquidity regulation based on the substitution possibilities between liquidity and capital. They provide evidence that this substitution occurs empirically. Their theory is based on the idea that it is much easier to verify the value of liquidity on a bank's balance sheet than that of other assets that contribute to its capital position. This allows liquidity regulation to supplement capital regulation in an effective way. Goodfriend (2016) argues that having banks hold reserves at the central bank is a much more effective way of pre-positioning liquid assets on a bank's balance sheet to promote financial resilience than the LCR requirement. The latter in its current form is a very difficult regulation to enforce. Central banks can achieve the same goals much more easily by varying the interest rate on bank reserves held with them.

Diamond and Kashyap (2016) develop a rationale for liquidity regulation based on depositors having incomplete information about the bank's ability to survive a run. The model is an extension of the standard Diamond and Dybvig (1983) framework. The asymmetric information between depositors and the bank leads to a market failure where the bank does not have the correct

incentives to hold the right amount of liquidity. It is shown that regulations corresponding to an LCR and NSFR can make runs less likely. However, optimal regulation in this model does not involve these rules.

Allen and Gale (2016) show that financial regulation in general and liquidity regulation in particular needs to be constructed taking into account the fact that there are incentives to innovate around it. Such innovations are difficult to forestall short of specifying exactly what kinds of financing instruments firms can use. However, this type of regulation will rule out beneficial innovations as well and may be quite costly in the long run.

Bouwman (2014) provides an overview of how banks create liquidity and a survey of the literature on how liquidity should be regulated. She recounts the history of liquidity regulation and points to the paucity of theoretical analysis. In her view, the need for regulation arises because of moral hazard associated with deposit insurance and the discount window. One of the points she stresses is the importance of the interaction between capital and liquidity regulation and the need for both to be done in concert.

The account of the literature on liquidity regulation in this section indicates that it is still at an early stage. In particular, the models the analysis is based on are quite simple and capture only a small part of the market failures that have been identified in the literature on liquidity provision.

At a more basic level, there is no clear analysis of whether liquidity should correspond to short term real assets or to monetary instruments. Real models of banking focus on the former while monetary ones focus on the latter. Distorting the allocation of real assets through regulation may be socially costly. In contrast, if the central bank creates liquidity that is held by commercial banks then that is potentially costless as argued by Buiter (2007).

Much more research is required in this area. With capital regulation there is a huge literature but little agreement on the optimal level of requirements. With liquidity regulation, we do not even know what to argue about.

## References

- Acharya V, D Gale and T Yorulmazer (2011), 'Rollover Risk and Market Freezes', *Journal of Finance* 66(4), pp 1177–1209.
- Acharya V, D Gromb and T Yorulmazer (2012), 'Imperfect Competition in the Interbank Market for Liquidity as a Rationale for Central Banking', *American Economic Journal: Macroeconomics* 4(2), pp 184–217.
- Allen F and D Gale (1998), 'Optimal Financial Crises', *Journal of Finance* 53(4), pp 1245–1284.
- Allen F and D Gale (2004a), 'Financial Fragility, Liquidity, and Asset Prices', *Journal of the European Economic Association* 2(6), pp 1015–1048.
- Allen F and D Gale (2004b), 'Financial Intermediaries and Markets', *Econometrica* 72(4), pp 1023–1061.
- Allen F and D Gale (2007), *Understanding Financial Crises*, Clarendon Lectures in Finance, Oxford University Press, Oxford.
- Allen F, E Carletti and D Gale (2009), 'Interbank Market Liquidity and Central Bank Intervention', *Journal of Monetary Economics* 56, pp 639–652.
- Allen, F, E Carletti and D Gale (2014), 'Money, Financial Stability and Efficiency', *Journal of Economic Theory* 149, 100-127.
- Allen F and D Gale (2016), 'Liquidity Regulation, Extended Repo and the Real Economy,' working paper, Imperial College London.
- Antinolfi G, E Huybens and T Keister (2001), 'Monetary Stability and Liquidity Crises: The Role of the Lender of Last Resort', *Journal of Economic Theory* 99(1–2), pp 187–219.
- Bagehot W (1873), 'Lombard Street: A Description of the Money Market, London: H. S. King'.
- Bech M and T Keister (2013), 'Liquidity Regulation and the Implementation of Monetary Policy', BIS Working Paper No. 432.
- Berger AN and CHS Bouwman (2009), 'Bank Liquidity Creation', *Review of Financial Studies* 22(9), pp 3779–3837.
- Berger AN and CHS Bouwman (2013), 'How Does Capital Affect Bank Performance During Financial Crises?', *Journal of Financial Economics* 109(1), 146–176.
- Bhattacharya S and D Gale (1987), 'Preference Shocks, Liquidity, and Central Bank Policy', in WA Barnett and KJ Singleton (eds), *New Approaches to Monetary Economics: Proceedings of the Second International Symposium in Economic Theory and*

- Econometrics, International Symposia in Economic Theory and Econometrics, Cambridge University Press, Cambridge, pp 69–88.
- Bolton P, T Santos and JA Scheinkman (2011), ‘Outside and Inside Liquidity’, *Quarterly Journal of Economics* 126(1), pp 259–321.
- Bouwman, C (2014), ‘Liquidity: How Banks Create It and How It Should Be Regulated’, forthcoming in *The Oxford Handbook of Banking* (2nd edition; AN Berger, P Molyneux, and JOS. Wilson (eds.)).
- Brunnermeier M and LH Pedersen (2009), ‘Market Liquidity and Funding Liquidity’, *Review of Financial Studies* 22(6), pp 2201–2238.
- Bryant J (1980), ‘A Model of Reserves, Bank Runs, and Deposit Insurance’, *Journal of Banking and Finance* 4(4), pp 335–344.
- Buiter, W (2007), ‘Maverecon’, *Financial Times*, December 13, 2007.
- Champ B, B Smith and S Williamson (1996), ‘Currency Elasticity and Banking Panics: Theory and Evidence’, *Canadian Journal of Economics* 29(4), pp 828–864.
- Diamond DW and PH Dybvig (1983), ‘Bank Runs, Deposit Insurance, and Liquidity’, *Journal of Political Economy* 91(3), pp 401–419.
- Diamond DW and AK Kashyap (2016). ‘Liquidity Requirements, Liquidity Choice and Financial Stability,’ *Handbook of Macroeconomics*, forthcoming.
- Diamond DW and RG Rajan (2001), ‘Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking,’ *Journal of Political Economy* 109(2), pp 287–327.
- Diamond DW and RG Rajan (2006), ‘Money in a Theory of Banking’, *American Economic Review* 96(1), pp 30–53.
- Diamond DW and RG Rajan (2011), ‘Fear of Fire Sales, Illiquidity Seeking, and Credit Freezes’, *Quarterly Journal of Economics* 126(2), pp 557–591.
- Freixas X and J Jorge (2008), ‘The Role of Interbank Markets in Monetary Policy: A Model with Rationing’, *Journal of Money, Credit and Banking* 40(6), pp 1151–1176.
- Freixas X, A Martin and D Skeie (2011), ‘Bank Liquidity, Interbank Markets and Monetary Policy’, *Review of Financial Studies* 24(8), pp 2656–2692.
- Goodfriend M (2016), ‘Liquidity Regulation, Bank Capital, and Monetary Policy,’ Testimony before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate, June 7, 2016.

- Goodfriend M and R King (1988), 'Financial Deregulation, Monetary Policy and Central Banking', *Federal Reserve Bank of Richmond Economic Review* 74(3), pp 3–22.
- Gorton GB (1988), 'Banking Panics and Business Cycles', *Oxford Economic Papers* 40(4), pp 751–781.
- Heider F, M Hoerova and C Holthausen (2009), 'Liquidity Hoarding and Interbank Market Spreads: The Role of Counterparty Risk', ECB Working Paper No 1126.
- Heider F, M Hoerova and C Calomiris (2015), 'A Theory of Bank Liquidity Requirements,' Working Paper SSRN id247701.
- Morris S and H Shin (2009), 'Illiquidity Component of Credit Risk', Working Paper, Princeton University.
- Perotti, R and J Suarez (2011), 'A Pigovian Approach to Liquidity Regulation', *International Journal of Central Banking* (December), pp. 3-41.
- Rauch C, S Steffen, A Hackethal and M Tyrell (2009), 'Savings Banks, Liquidity Creation and Monetary Policy', Working Paper, Goethe University Frankfurt.
- Repullo R (2005), 'Liquidity, Risk Taking, and the Lender of Last Resort', *International Journal of Central Banking* 1(2), pp 47–80.
- Rochet J-C (2004), 'Macroeconomic Shocks and Banking Supervision', *Journal of Financial Stability* 1(1), 93-110.
- Rochet J-C (2008), 'Liquidity Regulation and the Lender of Last Resort,' *Financial Stability Review* 11(February), pp. 45-52.
- Rochet J-C and X Vives (2004), 'Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?', *Journal of the European Economic Association* 2(6), pp 1116–1147.
- Skeie D (2008), 'Banking with Nominal Deposits and Inside Money,' *Journal of Financial Intermediation* 17(4), pp. 562–584.
- Smith B (2002), 'Monetary Policy, Banking Crises, and the Friedman Rule', *American Economic Review* 92(2), pp 128–134.
- Stein J (2013), 'Liquidity Regulation and Central Banking', Remarks made at 'Finding the Right Balance', 2013 Credit Markets Symposium, Federal Reserve Bank of Richmond, Charlotte, NC.