A Theory of Bank Liquidity Requirements*

Charles Calomiris†  Florian Heider‡  Marie Hoerova§

Jan 2013

PRELIMINARY AND INCOMPLETE

Abstract

We develop a theory of bank liquidity requirements that considers the substitutability of cash requirements and capital requirements for prudential regulation. We present a model that encompasses three motives for requiring bank cash holdings as part of a prudential regulatory framework: (1) maintaining cash in advance saves on liquidation costs, (2) cash is observable and verifiable (while measuring capital requires a valuation of the loan portfolio, measuring cash does not), and (3) because the riskiness of cash is invariant to bankers’ decisions about whether to invest resources in risk management, greater cash holdings improve incentives to manage risk in the non-cash asset portfolio of risky assets held by the bank. In a stand-alone autarkic banking equilibrium, cash is held voluntarily by banks as a commitment device to manage risk properly; increasing cash holdings in response to adverse news stems depositors’ incentives to withdraw funds. In a model of multiple banks with diversifiable risk, the coalition of banks will commit to lend each other funds in response to bank-specific needs to accumulate cash; in that equilibrium, cash requirements will be imposed by the group to prevent free riding on efficient interbank liquidity assistance, and cash requirements will be less than the cash holdings voluntarily held in the autarkic equilibrium. In a model with government deposit insurance, cash holdings will be set higher than under either the autarkic equilibrium or the coalition equilibrium; in the presence of deposit insurance, higher cash requirements are necessary to prevent moral hazard and ensure proper risk management.

*We would like to thank our discussants Martin Hellwig and Christine Parlour, and seminar participants at the Cleveland Fed, NYU, the IMF Twelfth Jacques Polak Annual Research conference and the WFA conference. The views expressed do not necessarily reflect those of the European Central Bank or the Eurosystem.

†Columbia University, Graduate School of Business, email: cc374@columbia.edu
‡European Central Bank, Financial Research Division, email: florian.heider@ecb.int
§European Central Bank, Financial Research Division, email: marie.hoerova@ecb.int
1 Introduction

In response to the global financial crisis of 2007-2009, the Basel Committee has proposed a new set of liquidity requirements to complement its revised framework of capital requirements. The primary and obvious motivation for the new interest in managing banks’ liquidity is concern about liquidity risk, which we define as the risk that a solvent bank may find itself unable to manage its current flow of withdrawals from its own stock of liquidity and access to borrowed funds from others. After the heavy reliance of banks on central bank lending during the crisis, policy makers understandably would like to reduce the dependence of banks on the lender of last resort, and thus encourage banks to limit or self-insure (through cash asset holdings) some of their liquidity risk.

In this paper, however, we show that the role of liquidity requirements should be conceived in a more nuanced way, not just as an insurance policy to deal with liquidity risk that can arise in a financial crisis, but as a prudential regulatory tool alongside capital requirements which, like capital requirements, has important consequences for limiting default risk and for encouraging good risk management. It may be that the primary benefits derived from liquidity requirements (like capital requirements) relate to improvements in bank risk profiles and incentives towards risk during normal times, and the consequences of those behaviors for reducing the probability of a liquidity crisis by making banking systems more resilient from a default risk perspective. After all, systemic liquidity crises, historically and in the recent past, result from asymmetric information about the possibility that one or more banks may be insolvent, which underlies the collapse of short-term credit markets in which banks participate (unsecured, asset-backed commercial paper, and repos). The best way to deal with a liquidity crisis is to design a combination of capital and liquidity requirements that prevent liquidity crises by ensuring that banks do not become weak enough so that their default becomes imaginable.

This paper provides a simple multi-period framework for conceiving of the role of liquidity requirements alongside capital requirements, and considers the interactions between liquidity
and capital requirements and the optimal combination of the two in the context of a model in which liquidity and capital requirements both affect default risk and liquidity risk. We show that, under the assumptions of a parsimonious model that allows bankers to choose their balance sheet structure and their risk profile, there is an optimal combination of capital and liquidity. Bankers are not indifferent (as they would be, say, under the Merton 1977 model of risk management) between combinations of liquidity and capital that have the same static consequences for default risk. This equilibrium reflects the dynamic consequences of liquidity and capital decisions for the management of default risk and liquidity risk in the future.

There are two essential differences between capital on the liability side and cash on the asset side that drive our results: (1) The value of cash, unlike capital, is observable (verifiable) at all times; the value of capital, in contrast, is not observable, but depends on the value of risky assets. (2) Cash is a riskless asset and so, when banks hold cash, they commit to removing default risk from a portion of their portfolio. Because cash is both observable and riskless, the commitment to hold cash has important implications for bankers’ incentives toward risk in the future. That commitment affects the way outsiders – who lack information about bank assets and bankers’ behavior – view the risk management of the bank, which has important immediate implications for the bank’s access to funding.

In our model, not only does cash mitigate the liquidity risks attendant to exogenous shocks (which are modeled using a simple Diamond-Dybvig 1983 cost of liquidation), it also mitigates endogenous (banker-chosen) default risk. In the model, costly state verification makes debt (withdrawable deposits, as in Calomiris and Kahn 1991) the optimal form of outside finance. There is a conflict of interest between the banker/owner and the depositors with respect to risk management; the banker suffers a private cost from managing risk, and does not always gain enough as the owner to offset that cost (as in Tirole 2010). Greater cash holdings increase the marginal gain to the banker from managing risk, and thereby encourage greater risk management.
In our model, the default-risk-mitigation consequence of cash holdings can be more important than the liquidity-risk-mitigation consequence. We demonstrate the relative importance of cash for mitigating these two risks by solving our model under three alternative policy regimes: (1) autarky, in which a single bank chooses liquidity and capital ratios in combination to manage its risks in an environment without interbank cooperation or government regulation, (2) collective action, where a group of banks forms a coalition to coinsure against liquidity risk, and (3) government deposit insurance, in which the government establishes a government-regulated and financed deposit insurance system.

Under autarky, regulation of capital or liquidity plays no role. The bank voluntarily chooses the socially optimal combination of cash holdings and capital holdings, trading off the costs and benefits of each. Under collective action, the coalition of banks (which we refer to as the “clearinghouse,” since bank clearinghouses historically were the form of collective action that managed liquidity risk) is able to achieve a better outcome than banks can achieve under autarky because liquidity risk is diversifiable across banks. The bank coalition will impose a cash reserve requirement on its members to limit free riding on collective protection (as discussed, for example, in Gorton 1985). Government deposit insurance is an alternative means for eliminating liquidity risk. The government also sets prudential regulations, including a cash reserve requirement, optimally to limit the moral-hazard consequences of government protection of deposits and bailouts of other bank claimants, which can be especially pronounced when risks to the banking system are correlated (Farhi and Tirole 2009).

In our framework, because bank coalitions have strong incentives and the capability to monitor each other’s risks, they are able to reduce liquidity risk and reduce the amount of cash banks hold in equilibrium; thus, collective action not only reduces risk, it also reduces the cost of intermediation (through a law-of-large-numbers diversification effect analogous to Diamond 1984). In contrast, government deposit insurance – which does not entail the same incentives for interbank monitoring – can only achieve the same outcome of risk reduction
with a higher cash asset requirement than would occur under autarky.

That cross-regime comparison illustrates the relative important of cash as a device for managing default risk. In the presence of deposit insurance, there is no liquidity risk in the banking system, and yet the role of cash in the prudential regulation is even greater than in the autarkic regime where liquidity risk is still present.

Section 2 provides a brief review of the literature. Section 3 describes the physical assumptions of the model. Section 4 presents the benchmark case in which effort is observable and there is no moral hazard problem. We then consider the problem with moral hazard. Section 5 solves for equilibrium under moral hazard and autarky. Sections 6 and 7 solve for equilibrium under collective action and under government deposit insurance, respectively. Section 8 concludes.

2 Literature Review

A helpful place to begin any discussion of cash reserve requirements is with the Black-Scholes-Merton framework (e.g., Merton 1977). In that framework, there are no transaction or information costs, all information that can be known is known equally to all parties, and all securities are equally (perfectly) liquid – by which we mean that they can be sold for their true value, which is common knowledge, without incurring any physical liquidation cost. In that framework, the only special property of cash is its lack of risk. From the standpoint of prudential regulation – which seeks to avoid the default risk of banks – greater cash holdings and greater reliance on equity finance each reduce the default risk of the bank. Ceteris paribus, allocating more assets to cash reduces asset risk (defined as the standard deviation of asset returns), while increasing equity reduces default risk for any given level of asset risk. Cash and equity, therefore, are two alternative ways to skin the cat of reducing default risk.

Given that the ultimate prudential goal of controlling default risk can be achieved through many different combinations of the two, there is no unique optimal combination of cash-to-
asset ratio and equity-to-asset ratio. Figure 1 illustrates this fact. A one basis point level of actuarially fair default-risk level can be achieved with many different combinations of equity-to-asset ratios and asset risk, as can a 50 basis point level of default risk.

When one relaxes the assumptions of the Black-Scholes-Merton framework, however, the effects of cash and equity on default risk are not the same. For example, in Diamond and Dybvig (1983) physical costs of liquidation make liquidity risk (the possible need to finance early consumption) costly, which could motivate the holding of inventories of liquid assets. In Calomiris and Kahn (1991), depositors receive noisy and independent signals about the risky portfolio outcome of the bank. By holding reserves, banks insulate themselves against the liquidity risk of a small number of “misinformed” early withdrawals in states of the world where the outcome is actually good. Without those reserves, banks offering demandable debt contracts (which are optimal in the Calomiris-Kahn model) would unnecessarily subject themselves to physical liquidation costs when they fail to meet depositors’ requests for early withdrawal.

In many models of banking, equity is an undesirable means of controlling risk because of the high costs of raising equity. Typically, the optimal contract between the banker and his or her funding sources is a debt contract, and sometimes a demandable debt contract. Debt contracts economize on the cost of ex post verification (Townsend 1979, Diamond 1984, Gale and Hellwig 1985, Calomiris and Kahn 1991), and also can be beneficial by reducing the negative signaling of bank type (Myers and Majluf 1984). Equity is assumed to be supplied by the banker, and not by outsiders because of its prohibitively high cost.

Given the high costs of external finance via equity – for example, in a simple Myers-Majluf model with adverse-selection costs of equity issuance – cash holdings can provide a more cost-effective means of reducing default risk (Calomiris and Wilson 2004).

In Tirole (2010), asymmetric information about risk management effort creates a conflict of interest between managers and debtholders. Biais, Heider and Hoerova (2010) show that, under asymmetric information, optimal hedging contracts may benefit from the use of cash
margins, as a commitment device to manage risk properly. As we show in the model below, when high costs of raising equity (e.g., resulting from verification costs) favor high leverage, the agency problem of risk management is pronounced (Jensen and Meckling 1976, Myers 1977). In that environment, cash holdings can be an especially useful means of reducing risk through the effects of cash holdings on the incentives of managers to expend effort on risk management. Because cash holdings limit the extent to which debtholders lose from high-risk strategies, more cash helps to better align managers’ incentives with the interest of debtholders.

Our paper builds on the insights of these various models. Like Diamond and Dybvig (1983) we assume that exogenous demand for early liquidation can be physically costly, which gives rise to liquidity risk in our model. And like Tirole (2010) and Biais, Heider and Hoerova (2010), we assume that risk management entails private costs to the bank manager. We also follow the banking literature by considering the effects of verification costs. In such a framework, cash is doubly important, not just because of its lack of risk, but because of its observability. Outsiders (those who face high costs of observing outcomes) can be confident that managers will exert effort for proper risk management irrespective of the unobservable state, if and only if cash margins are sufficiently high to ensure that managers have the incentive to exert that effort. We show that, in equilibrium, withdrawable deposit contracting with outsiders, and a sufficient amount of cash reserves, provides an optimal contractual solution to the banking problem: That solution minimizes the overall costs associated with early liquidation, shirking in risk management, foregone opportunities for profits (from cash holdings), and costs of verification. Costs of verification are not essential to most of the model’s key results, and so we derive our model without costly verification initially, and later consider its additional implications.

It is also interesting to note the long tradition in banking – despite the absence of formal modeling – that has focused on cash requirements as a prudential device. Indeed, with few exceptions, historical prudential regulation prior to the 1980s has concentrated on require-
ments for cash rather than requirements for capital. For example, the New York Clearing House maintained a 25% cash reserve requirement against deposits for its members.\footnote{An additional motivation for New York City banks to maintain high required reserves was that city’s position at the peak of the “reserve pyramid” that connected banks throughout the United States. Financial panics gripped the US frequently during the pre-World War I era, which largely reflected the fragmented structure of US banking (Calomiris and Gorton 1991, Calomiris and Schweikart 1991, Calomiris 2000, 2011). As numerous scholars have noted, although the New York bank Clearing House sometimes acted collectively during financial crises, it generally did not employ its reserves to protect banks in other regions or non-clearing house members during financial crises (Goodhart 1995, Wicker 2000).} In the famous 1873 Coe Report, authored by George Coe, the President of the New York Clearing House, cash was seen as the essential tool for managing systemic risk (Wicker 2000). In contrast, neither regulators nor bank coalitions set minimum equity capital-to-asset ratios for banks, with few exceptions, until the 1980s.\footnote{Some of those exceptions were the state experiments with deposit insurance, for example, in the early 20th century. See Calomiris (1989, 1990).}

In private interbank coalitions, it is easy to see why cash would be especially valued over equity capital as a tool. Not only does cash mitigate the costs of raising equity ex ante and mitigate liquidity risks ex post, a greater reliance on observable cash (rather than hard-to-observe equity) to manage free-riding problems has additional advantages in the context of a private coalition of competing banks. For bankers to verify each other’s capital in a precise way they would have to be able to demonstrate the value of each bank’s capital to the group. Given the private information inherent in bank lending, doing so would require a significant expenditure of cost. Furthermore, a full analysis of the value of a bank’s loan portfolio would require the sharing of information about loan customers, and such sharing would destroy quasi rents banks earn from investing in private information (Rajan 1992); thus, a policy of collective monitoring of each bank’s capital would entail huge information costs to banks, and erode the incentive for banks to invest in client-specific information. Our model of collective action developed in Sections 5 and 6 does not consider these quasi-rent-erosion costs of enforcing capital requirements in a banking coalition, but obviously, if they were included, they would magnify the gains from the use of cash to manage default risk.

The observability advantage of cash has been particularly obvious in the recent history of banking.
of financial crises in our current highly regulated and protected banking environment, where
government both insures deposits and regulates banks’ risk-based capital requirements. In
this environment, depositors or other banks have little incentive to monitor banks or penalize
high risk. Supervisors, who are charged with protecting taxpayers from the costs of bailing
out banks, typically have not identified bank losses in a timely way. That fact reflects a com-
bination of problems, including the absence of private incentives for supervisors to invest in
monitoring, and the presence of political pressures on supervisors to avoid recognizing losses
(Calomiris and Herring 2011). For example, in December 2008, as an insolvent Citibank
was bailed out by the U.S. government, its ratio of book capital to risk-weighted assets still
exceeded 11%.

In the current environment, therefore, it is particularly important to consider the advan-
tages of requiring cash relative to requiring book equity capital, since cash may be a more
reliable risk mitigator. Bankers can “evergreen” bad loans, and supervisors can “forebear”
from recognizing loan losses, but if bankers and supervisors falsify the accounting of cash
reserves, they will be committing fraud, and will be subject to credible ex post penalties.
That fact gives cash a special role for protecting taxpayers, not only from liquidity risks of
orises, but from the moral-hazard problems of the abuse of government safety nets within a
taxpayer-protected banking system.

3 The Model

There are five dates, $t = 0, 1, 2, 3, \text{ and } 4$, and two types of agents, a banker and depositors.

The banker is risk-neutral. At time 0, the banker is endowed with knowledge about a
prospective loan making opportunity, and with own equity $E_0 > 0$.\(^3\)

The banker also accepts deposits $D$ at time 0. The supply of deposits is perfectly elastic
to the banker up to a maximum supply of $\bar{D}$.\(^4\) Depositors are risk-neutral and can store

\(^3\)Alternatively, we could have several perfectly correlated individual loan making possibilities.
\(^4\)For simplicity, we assume a maximum supply of deposits. This is a reduced form to capture an increasing
their funds, which yields a per-unit return of one for sure.

At time 0, the banker can invest the bank’s resources, \( D + E_0 \), in cash (the amount invested in cash is denoted \( C_0 \)) or in loans (the amount invested in loans is denoted \( L_0 \), which is also the number of loans since the loan size is normalized to 1). Hence, \( C_0 + L_0 = D + E_0 \).

The payoff from the loan will occur at time 4 if the loan is allowed to mature. That payoff per unit invested will either be \( Y > 1 \) or 0. The probability of the high payoff depends on the banker’s costly risk-management effort at time 3. If he exerts effort, the high payoff \( Y \) occurs for sure and if he does not exert effort, \( Y \) occurs only with probability \( p < 1 \). When the banker chooses not to exert effort, he derives a private benefit \( B \) per loan (equivalently, when not exerting effort, the banker saves on costly risk management). The effort choice of the banker is unobservable by depositors.

At time 1, there are two possible aggregate states \( s \), good or bad, \( s = \{g, b\} \). The probability of the good state occurring is denoted by \( q \) (the bad state occurs with probability \( 1 - q \)). The aggregate state is observed for free by the banker and by any depositors who choose to invest in the signal, at a cost to depositors of \( m \). The signal is perfectly revealing. In the good aggregate state, the private benefit of the banker is smaller than in the bad aggregate state, \( B_g < B_b \). That is, we assume that during times when aggregate downside risk is greater, risk management is more costly. This assumption could be viewed as capturing, in a reduced-form, a set-up with costly loan monitoring (whereby the frequency of monitoring should be higher in bad times to achieve the same probability of repayment) or a set-up with searching for prospective loan applicants and/or screening out bad borrowers (which is more difficult in bad economic times). We let \( B \) denote the expected private benefit of the banker, i.e. \( B \equiv qB_g + (1 - q)B_b \). We assume that

\[
Y > 1 > qY + (1 - q) (pY + B_b).
\]

(1)

cost of attracting more deposits due to, e.g., geographical distance. Alternatively, we could assume a limited amount of loan-making opportunities.
The first inequality implies that when the banker exerts risk-management effort, loan making has a positive NPV. The second inequality implies that unless the banker does risk-management effort in both aggregate states, loan making is socially wasteful: even after accounting for the (maximum) private benefit of the banker, it is more profitable to store funds.\footnote{The expression on the right-hand side of (1) is the return from exerting effort in the good state and not exerting effort in the bad state. Since $B_b > B_g$ and, by (1), $Y > pY + B_b$, this is a maximum expected return that can be generated if the banker does not exert effort in both aggregate states.}

At $t = 2$, after learning the aggregate state, the banker can increase his cash holdings by $\Delta C(s)$ by liquidating an amount $\Delta L(s)$ of loans. But converting loans into cash at $t = 2$ is costly, $\Delta C(s) = (1 - l) \Delta L(s)$, where $l$ is the cost of liquidation per unit of loans. The banker’s cash holdings at time 2 and state $s$ are $C_2(s) = C_0 + \Delta C(s)$, while his loan holdings are reduced to $L_2(s) = L_0 - \Delta L(s) = L_0 - \Delta C(s) \frac{1}{1 - l}$. Note that the liquidation of loans reduces the value of equity, $E_2 = C_2 + L_2 - D = E_0 - \Delta C(s) \frac{1}{1 - l}$.\footnote{This is analogous to the mechanism in Diamond and Rajan (2005, 2011) whereby the need to raise cash, either by increasing the rate on deposits or by selling assets to a cash-constrained market, reduces the value of the bank pushing it closer to insolvency.}

The purpose of having liquid assets prior to time 3 is to incentivize the banker to exert risk management effort. The banker’s incentive to exert effort, and thus increase the probability of high payoff, is lower in the bad aggregate state. By increasing liquid assets at time 2, the banker will be able to respond to the observed bad aggregate state by increasing liquid assets, and thereby make risk-management effort at time 3 credible.

At the end of $t = 2$, depositors choose whether or not to “run” the bank, i.e., whether or not to shut down the bank. We assume that one depositor’s announcement is enough to shut down the bank. This can be understood, e.g., in a set-up in which one depositor has a low cost $m$ of acquiring a signal about the aggregate state, while the other depositors fact prohibitive costs (so that only one of the depositors is informed). The informed depositor can be seen as causing the bank to close by withdrawing his or her funds, and triggering imitative withdrawals by other depositors (who, in equilibrium, would not otherwise run the bank). Alternatively, one can imagine the monitoring cost $m$ as a shared cost by a finite
number of depositors with equal monitoring cost who randomize in their decision whether to monitor (as in Calomiris, Kahn and Krasa 1992). For simplicity, we adopt the simple liquidation rule that when the bank is shut down, the banker surrenders all claims to the bank’s assets. Depositors liquidate the loan portfolio at a cost \( v \) per unit of loans. We assume that depositors face a higher cost of liquidation cost than the banker, i.e. \( v > l \). Hence, if the bank is shut down, depositors receive the liquidation value of loans plus the cash in the bank, \((1 - v) L_2(s) + C_2(s)\).

At time 3, the banker chooses whether or not to exert risk-management effort. Effort is not contractible since it is not observable by anyone but the banker.

At time 4, the loan outcome \( Y \) or 0 is realized. If the high payoff is realized, then depositors receive \( R \) and the banker receives the residual, \( Y L_2(s) + C_2(s) - R \). If the low payoff is realized, depositors receive \( R_f \leq C_2(s) \). Note that holding cash allows the banker to pay depositors a positive amount if the loan returns nothing.\(^7\)

In sum, the timeline is as follows:

Time 0: \( E_0 \) and \( D \) are invested in \( L_0 \) and \( C_0 \) and contract promises \((R, R_f)\).

Time 1: The state is revealed and the banker updates his private benefit \( B_s \) based on the aggregate state. Depositors choose whether or not to invest \( m \) to observe the state.

Time 2: The banker can increase reserves by liquidating some loans, where the new amount of cash is given by \( C_2(s) \). Depositors choose whether or not to shut down the bank.

Time 3: The banker decides whether to exert unobservable risk-management effort.

Time 4: The banker pays \( R \) if outcome is \( Y \) and \( R_f \) if outcome is 0.

\(^7\)Note that with our two-point distribution, there is no real difference between debt and equity for the outside claim held by depositors. We can explicitly derive the outside claim to be a debt contract using costly state verification. Cash would then serve an additional function in the model, namely economizing on verification costs. For the sake of simplicity, and without loss of generality, we avoid this complication in our initial formulation of the model.
4 First-best Equilibrium

We now consider the case in which the banker’s risk-management effort is publicly observable so that there is no moral hazard problem. While implausible, it offers a benchmark against which we will identify the inefficiencies generated by moral hazard.

In the first-best, efficiency requires that the banker does risk-management effort since loan making is only productive when the banker exerts effort (condition (1)). Moreover, cash is not used. This is because cash is dominated in terms of rate of return compared to loan making when accompanied by risk management effort, and there are no incentive benefits of cash since incentive problems are absent in the first-best. Hence, $E_0 + D = L_0$, $C_0 = C_2(s) = 0$ and loans always return $Y$ per unit invested. Note also that depositors do not need to spend $m$ to observe the aggregate state (and hence gauge how costly risk-management is for the banker) since they can observe the banker’s effort directly and write contracts that guarantee the banker will exert effort.

Depositors are willing to deposit with the banker as long as the contract offers them at least the same payoff as storage. That is, the participation constraint of depositors is given by:

$$R \geq D.$$  \hfill (2)

Since the banker has a unique ability to make loans, he will not leave depositors any rents, i.e., $R = D$. Every deposit brought in earns $Y - 1 > 0$ implying that it is optimal to operate the bank at the maximum scale. Hence, if the banker chooses to take in deposits, his payoff at $t = 4$ is

$$YE_0 + (Y - 1)\bar{D}.$$  \hfill (3)

The banker prefers to take in deposits and operate the bank rather than invest solely his own equity as long as two conditions are satisfied. First, he must earn more by operating
the bank with deposits compared to investing his equity only:

\[ YE_0 + (Y - 1)\bar{D} \geq YE_0, \]  

(4)

which is satisfied since \( Y > 1 \). Second, he must be sure that depositors will not shut down the bank at \( t = 2 \), in which case his payoff is zero. If depositors do not run at \( t = 2 \), they obtain \( \bar{D} \) for sure, while if they run, they obtain the liquidation value of the bank, or \((1 - v)(E_0 + \bar{D})\). Hence, depositors do not run in the first-best as long as:

\[ \bar{D} \geq \frac{1 - v}{v}E_0 \]  

(5)

This is a condition on exogenous parameters: the liquidation cost should be large and the banker’s equity small relative to the amount of deposits available. Interestingly, this imposes a lower threshold on the bank’s leverage.

Proposition 1 summarizes the first-best outcome.

**Proposition 1 (First-best)** When risk-management effort is observable, the banker exerts effort, cash is never used, and depositors do not invest in the signal about the aggregate state.

## 5 Equilibrium under Moral Hazard

In this section, we solve for the autarkic banking equilibrium when risk-management effort is unobservable. We show that the banker may choose to hold cash as a commitment device for proper risk management.

Unlike in the first-best, the banker must now be induced to exert risk management effort since his effort is no longer contractible by depositors. There are two incentive-compatibility constraints, one for each aggregate state, which ensure that the banker prefers to exert risk-management effort. The constraints are conditional on the aggregate state since the banker observes the realization of the aggregate state before he makes his effort decision.
The incentive-compatibility constraint in state $s$ is given by:

$$YL_2(s) + C_2(s) - R \geq p[YL_2(s) + C_2(s) - R] + (1 - p)[C_2(s) - R_{fs}] + B_s L_2(s)$$

The expression on the left-hand side is the banker’s payoff when he exerts risk-management effort. The expression on the right-hand side is his (out-of-equilibrium) expected payoff if he does not exert effort. With probability $1 - p$, the loans return zero and the payments to depositors cannot exceed the banker’s cash holdings, $R_{fs} \leq C_2(s)$. Simplifying the incentive constraint we get:

$$\mathcal{P}_s L_2(s) \geq R - R_{fs} \quad (6)$$

where

$$\mathcal{P}_s \equiv Y - \frac{B_s}{1 - p} \quad (7)$$

Following Tirole (2006), we refer to $\mathcal{P}$ as the “pledgeable income” of the banker, i.e., the share of the return per unit of loans that can be pledged to depositors without jeopardizing the incentives of the banker to properly manage the loan portfolio. Note that $\mathcal{P}_s > 0$ (with $\mathcal{P}_g > \mathcal{P}_b$) under our assumption that effort is productive in both aggregate states (see (1)).

Note that setting the (out-of-equilibrium) payoff to depositors $R_{fs}$ as high as possible relaxes the incentive constraint. Hence, it is optimal to set $R_{fs} = C_2(s)$. Moreover, the participation constraint of depositors is given by (2) and, as before, the banker does not want to leave any surplus to depositors, so the participation constraint binds, $R = D$. Using $L_2(s) = L_0 - \Delta C(s) \frac{1}{1 - l}$, $C_2(s) = C_0 + \Delta C(s)$, and $L_0 = D + E_0 - C_0$, we can re-write (6) as
\[ P_s (D + E_0) + (1 - P_s) C_0 + \left(1 - \frac{P_s}{1 - l}\right) \Delta C (s) \geq D. \]  

Note that \( 1 - P_s > 1 - \frac{P_s}{1 - l} \). If \( P_s \geq 1 \) in both aggregate states, then cash holdings cannot help with incentives as they tighten the incentive constraints. Hence, cash will not be held in this case, implying that the incentive constraints simplify to \( P_s (D + E_0) \geq D \). This is always satisfied for \( P_s \geq 1 \). We can thus state the following lemma.

**Lemma 1** When risk-management effort is not observable, but \( P_s \geq 1 \) in both aggregate states, the first-best is always reached.

Moreover, note that even when \( P_s < 1 \), as long as \( \frac{P_s}{1 - P_s} \geq \frac{D}{E_0} \), the first-best can be reached with \( C_0 = \Delta C (s) = 0 \). In what follows, we focus on the case in which the first-best is not attainable, i.e., \( P_s < 1 \) and \( \frac{P_s}{1 - P_s} < \frac{D}{E_0} \) in at least one aggregate state. Since \( P_g > P_b \), we require that the first-best be not attainable in the bad aggregate state.\(^8\)

### 5.1 Aggregate State Observable at No Cost

We start by characterizing the case in which depositors can observe the aggregate state at zero cost, \( m = 0 \). We consider the case in which \( m > 0 \), and the depositor’s choice of whether or not to expend \( m \) and observe the state in the next subsection.

When \( P_s < 1 \), we can re-write (8) as

\[ \frac{P_s}{1 - P_s} E_0 + \frac{1}{1 - P_s} \left(1 - \frac{P_s}{1 - l}\right) \Delta C (s) \geq D - C_0 \]  

An increase in \( t = 2 \) cash holdings, \( \Delta C (s) > 0 \), can only relax the incentive constraint when the term in brackets is positive, which is the case when

\(^8\)Note that incentive-compatibility could potentially be ensured by reducing the promised payout to depositors in the bad state. This would economize on the cost of liquidating loans at \( t = 2 \). We assume that depositors are dispersed and that such renegotiation cannot take place.
\[ P_s < 1 - l. \]  

Liquidating loans only helps with incentives when the liquidation return is higher than the pledgeable return. Moreover, (9) shows that while both cash at \( t = 0 \) and extra cash at \( t = 2 \) relax the incentive constraint, they do not have the same role in relaxing incentives. \( C_0 \) has a per unit benefit of one but is not contingent on the state. \( \Delta C(s) \) is contingent on the state (i.e., it is flexible), but its per unit benefit is less than one:

\[
\frac{1}{1 - P_s} \left(1 - \frac{P_s}{1 - l} \right) < 1.
\]

In equilibrium, the banker always exerts effort, and his payoff from taking in deposits and operating the bank is:

\[
q [YL_2(g) + C_2(g) - D] + (1 - q) [YL_2(b) + C_2(b) - D]
\]

which, using \( L_2(s) = L_0 - \Delta C(s) \frac{1}{1 - l} \), \( C_2(s) = C_0 + \Delta C(s) \), and \( L_0 = D + E_0 - C_0 \), re-writes as

\[
YE_0 + (Y - 1)(D - C_0) - \left( \frac{Y}{1 - l} - 1 \right) [q\Delta C(g) + (1 - q)\Delta C(b)]
\]

The banker maximizes (11) subject to the two incentive constraints (8), the feasibility constraints \( D \leq \tilde{D} \) and \( 0 \leq \Delta C(s) \leq (1 - l)(D + E_0 - C_0) \), the banker's participation constraint (requiring that he does better by taking in deposits rather than by investing only his own equity)

\[
YE_0 + (Y - 1)(D - C_0) - \left( \frac{Y}{1 - l} - 1 \right) [q\Delta C(g) + (1 - q)\Delta C(b)] \geq YE_0
\]

and the no run condition for depositors requiring that \( (1 - v) L_2(s) + C_2(s) \leq D \) or, after substituting for \( L_2(s) \) and \( C_2(s) \) and simplifying:
\[
\frac{1 - v}{v} E_0 + \Delta C(s) \frac{v - l}{v(1 - l)} \leq D - C_0
\] (13)

For ease of exposition, we will now consider the case in which there is no incentive problem in the good aggregate state, \( c \bar{P}_g > 1 \), and in which the liquidation costs for depositors are sufficiently high (see condition (14) below). Then, \( \Delta C(g) = 0 \). At this point, it is useful to introduce some notation. Define the net return on loans as \( r \equiv Y - 1 \), the outside financing capacity of equity in the bad state as \( f_b \equiv \frac{P_b}{P_b^1} \), the ratio of return-to-cost of liquidation for the banker as \( \lambda \equiv \frac{1}{1 - l} \), and the ratio of return-to-cost of liquidation for depositors as \( \nu \equiv \frac{1 - v}{v} \), with: \(^9\)

\[
\nu < \lambda - (1 - q) \left( \frac{1 + r}{r} + \lambda \right) \quad (14)
\]

With this notation, we can now spell out our assumption about the maximum available amount of deposits, \( \bar{D} \). We posit that

\[
\frac{\bar{D}}{E_0} < \lambda < \frac{\lambda + f_b}{1 + f_b}
\] (15)

This assumption avoids the full liquidation of the bank in the bad aggregate state and makes it impossible for bankers with severe moral hazard (low \( f_b \)) to attract any deposits.

The incentive constraint in the bad aggregate state must be binding since cash holdings are costly and hence the banker will choose to hold just enough cash to satisfy the incentive constraint. The binding incentive constraint in the bad state (9) can be written as:

\[
f_b E_0 + \frac{\lambda - f_b}{\lambda} \Delta C(b) = D - C_0
\] (16)

while the no run condition (13) can be written as:

\(^9\)This condition rules out a set of parameters for which deposit-taking is not feasible at intermediate levels of financing capacity \( f_b \).
\[ \nu E_0 + \frac{\lambda - \nu}{\lambda} \Delta C(b) \leq D - C_0 \]  \hspace{1cm} (17)

From (11), we get that the net surplus of the banker (over the return from investing only his own equity, \( YE_0 \)) is given by:

\[ r(D - C_0) - (1 - q) \frac{1 + r + r\lambda}{\lambda} \Delta C(b) \]  \hspace{1cm} (18)

Substituting the binding incentive constraint (16) into (18) yields:

\[ rf_bE_0 + \left[ \frac{\lambda - f_b}{\lambda} r - (1 - q) \frac{1 + r + r\lambda}{\lambda} \right] \Delta C(b). \]  \hspace{1cm} (19)

It follows that if the expression in the square brackets above is non-positive,

\[ \frac{\lambda - f_b}{\lambda} r - (1 - q) \frac{1 + r + r\lambda}{\lambda} \leq 0 \]

or, equivalently,

\[ f_b \geq \lambda - (1 - q) \left( \frac{1}{r} + \lambda \right), \]  \hspace{1cm} (20)

then the banker will never increase his cash holdings at \( t = 2 \) since his surplus is decreasing in \( \Delta C \).

Suppose condition (20) holds. From the binding incentive constraint (16), we get that:

\[ f_bE_0 = D - C_0 \]  \hspace{1cm} (21)

It follows that no run condition (17) is satisfied since \( f_bE_0 = D - C_0 \geq \nu E_0 \) holds by (20) and (14). By (19), the banker’s surplus in this case is given by:

\[ rf_bE_0 \]
Note that cash holdings do not affect the banker’s surplus since he only earns the net return \( r \) on loans. Hence, when condition (20) holds, the banker will raise deposits and invest \( L_0 = E_0 + f_b E_0 \) into loans. He can raise deposits beyond \( f_b E_0 \) (the financing capacity of his equity in the bad state) as long as he invests the extra deposits into cash at \( t = 0 \) so that the incentive constraint (21) continues to hold.

Suppose condition (20) does not hold. As the banker’s surplus (19) is increasing in \( \Delta C(b) \) in this case, the banker would want to set \( \Delta C(b) \) as high as possible. From the binding incentive constraint (16), we get:

\[
\Delta C(b) = \frac{\lambda}{\lambda - f_b} (D - C_0 - f_b E_0)
\]  

To make \( \Delta C(b) \) as high as possible, the banker will set \( D = \bar{D} \) and \( C_0 = 0 \). Hence, \( \Delta C(b) = \frac{\lambda}{\lambda - f_b} (\bar{D} - f_b E_0) \). By (15), the banker will not liquidate all the loans as he will exhaust the available deposits \( \bar{D} \) before full liquidation. That is, \( \Delta C(b) < (1 - l)(D + E_0) \) holds. Lastly, we check that the no run condition (17) is satisfied. Substituting for \( \Delta C(b), D, \) and \( C_0 \) in (17), we have that:

\[
\frac{\nu - f_b}{\lambda - f_b} [\bar{D} - \lambda E_0] \geq 0
\]

must hold. Since we assume that \( \bar{D} \leq \lambda E_0 \), the equilibrium with deposit-taking is only viable for \( f_b \geq \nu \), i.e., when the banker has a sufficiently high financing capacity of the inside equity compared to the liquidation return of depositors. For \( f_b < \nu \), the banker will not raise deposits and will run the bank using his own equity only.

The following proposition summarizes the case when depositors can observe the aggregate state at no cost.

**Proposition 2** When risk-management effort is unobservable and the first-best is unattainable, the banker is incentivized to exert effort by holding cash. If condition (20) holds, then adjusting cash at \( t = 2 \) is too costly and the banker builds incentive-compatible cash holdings...
at $t = 0$. If condition (20) does not hold and $f_b \geq \nu$, then the banker increases his cash holdings at $t = 2$ if the bad aggregate state is realized. For $f_b < \nu$, the banker will not raise deposits and will run the bank using his own equity only.

5.2 Aggregate State Observable at a Cost

In this subsection, we consider the case in which depositors can only observe the aggregate state at a cost $m > 0$. Hence, they have to decide whether or not it is in their interest to expend this cost.

If depositors spend $m$ and observe the aggregate state, then the solution to the banker’s problem is the same as the one characterized in the previous subsection, except that the promised payout to depositors has to be increased by $m$ (so that depositors still break-even on the contract, even though they have to pay $m$ to observe the aggregate state). Hence, the banker’s surplus is reduced by $m$.

If depositors decide not to spend $m$, then the solution to the banker’s problem is as follows. Since depositors cannot distinguish good aggregate state from bad, the banker has to hold the same amount of cash in both states. The banker’s cash holdings have to be high enough to ensure that he exerts risk-management effort in the bad aggregate state (then, he would also exert effort in the good state). If the banker did not hold enough cash for his incentive constraint in the bad state to be satisfied, the banking equilibrium would not be viable. It follows that the banker will find it optimal to satisfy his incentive constraint with $t = 0$ cash only. This is because there is no flexibility benefit of increasing cash at $t = 2$ (since depositors cannot tell the states apart) and increasing cash holdings at $t = 2$ is always more costly than building sufficient cash reserves at $t = 0$.

This logic leads to the following proposition.

**Proposition 3** If the cost of acquiring the signal about the aggregate state $m$ is low, depositors invest in the signal and the banker increases his cash holdings if the bad aggregate state
realizes, as long as the liquidation is not too costly. If the cost is high, depositors do not invest in the signal and the banker only uses cash at $t = 0$ as a commitment for proper risk management.

6 Mutual Liquidity Risk Insurance

Thus far in our model banks operate on a stand-alone basis. We now construct a model of mutual liquidity risk insurance among banks that approximates the role of bank clearing houses. Mutual liquidity risk insurance, in our model, improves efficiency by allowing banks to pool pure liquidity risk, defined as withdrawal risks that are unrelated to the aggregate state.

We will show that the optimal clearing house insurance contract takes the simple form of providing unlimited clearing house protection against pure liquidity risk to qualifying members. In order to qualify, members must maintain adequate required reserves, which vary with the aggregate state, and which are identical to the state-specific equilibrium levels of bank reserves derived in Section 5.

Pure liquidity risk is a new risk to the banking system, which was not considered previously: the possibility of depositor withdrawals for exogenous liquidity reasons at time 2. Exogenous demands for spending occur after each bank’s initial decision about how much to adjust reserves in period 2, after the revelation of the state, but prior to the bank’s effort decision, which occurs in period 3.

When a depositor in Bank A experiences a “spending demand shock” in period 2, he/she uses part of his/her deposit to purchase something from another depositor, who is a depositor with a different bank, Bank B. This is not a withdrawal of cash from Bank A, but rather a transfer of a check from a depositor in Bank A to a depositor in Bank B. When this “spending” takes place, in the absence of mutual liquidity risk insurance, Bank A must transfer cash to Bank B. Doing so, however, causes Bank A’s cash balance to fall. To restore
its equilibrium position prior to period 3, in the absence of mutual liquidity insurance, Bank A will adjust its period 2 liquidation of loans to preserve its period 2 state-dependent level of reserves (derived in Section 5). Thus, for either aggregate state, a spending shock implies that Bank A would have to liquidate additional loans and bear the liquidation cost \( l \) on the amount of deposits transferred at period 2. Under autarky, therefore, when banks face exogenous liquidity risk at time 2, they will choose to hold higher levels of reserves at \( t = 0 \) than the equilibrium quantity derived in Section 5, and depending on whether they experience withdrawals of sufficient size at time 2, they may have to liquidate additional loans at time 2.

Because all spending is assumed to be routed within the banking system, liquidity risk is fully diversifiable among banks. One bank’s loss of reserves is another bank’s gain. So long as there is no moral hazard in the liquidity risk insurance contract written among banks, banks will all benefit by fully insuring one another against period 0 withdrawal risk. For example, if Bank A and Bank B were members of a “clearing house” that fully insured liquidity risk, the clearing house would substitute its own deposit for the deposit withdrawal from Bank A, and ask Bank B to deposit its surplus reserves at the clearing house, which would be transferred to Bank A. By entering into such a contract all banks would be able to set their reserves at time 0 and at time 2 at the levels consistent with the equilibrium determined in Section 5. Liquidity risk insurance avoids any additional costs associated with higher reserve reserves at \( t = 0 \) or increased liquidation costs resulting from depositor spending at \( t = 2 \).

A simple contract that would provide full liquidity risk insurance without moral hazard would specify that all clearing house member banks would set their reserve levels in accordance with the state-dependent equilibrium derived in Section 5. All banks meeting those requirements would be eligible for receiving loans from the clearing house at \( t = 2 \), as needed, to transfer reserves from reserve-surplus to reserve-deficit banks, which would effectively substitute clearing house deposits for the deposits lost as the result of depositor spending.
Of course, the clearing house would not be willing to provide insurance against withdrawal risks to banks that failed to maintain the state-dependent required level of reserves. Banks unwilling to meet membership reserve requirements would be attempting to free-ride on the loan liquidation costs of other banks, by seeking reserves transfers as a substitute for their own need to liquidate loans in the bad state in period 2. By tying mutual liquidity insurance among members to maintaining appropriate reserve requirements, liquidity insurance can be provided without moral hazard.

7 Deposit Insurance

Now suppose that the government requires banks to participate in a deposit insurance system, which guarantees that all deposits will be repaid in full, either if demanded in period 2 or in period 4. Suppose also that government prudential regulators do not face sufficient incentives to monitor banks to determine aggregate states. This assumption is consistent with the basic criticism that unlike “market discipline” regulatory discipline does not make use of market information, because regulators and supervisors do not face incentives to invest in information and enforce proper behavior (Calomiris and Herring 2011).

Under these assumptions, how should prudential regulation be designed? Note that under deposit insurance, banks face no incentive to voluntarily hold significant levels of reserves, since depositors do not demand them.\(^{10}\)

If the regulatory system is incapable of enforcing state-dependent reserve requirements on banks, then regulators are unable to require additional reserves after the revelation of bad states. Under that circumstance, the optimal regulatory arrangement would mandate a uniformly high reserve requirement (the level required in the bad state), which banks would maintain from time 0. Although this high level of reserves would be wasteful - since it would

\(^{10}\)Banks will be willing to lend to one another to insure against liquidity risk in period 2 so long as interbank deposits are covered by deposit insurance. If not, then liquidity risk insurance among banks would be avoided; banks would hold a small amount of reserves to economize on liquidation costs associated with exogenous expenditure-related demands by depositors at time 2, but this would be much less than the amounts of reserves banks would hold for prudential reasons.
mandate unnecessarily high reserves in good states - it would be optimal in the presence of deposit insurance because a lower level would produce losses to taxpayers in bad states, as bankers would choose poor risk management in those states.

In contrast to the liquidity-insurance contract of Section 6 - which reduced costs of maintaining reserves - deposit insurance raises the social costs of reserves. It is also important to recognize that it does so despite the fact that deposit insurance eliminates all liquidity risk in the banking system. That fact illustrates the key role of reserves as a prudential regulatory tool, not just to mitigate liquidity risks.

Deposit insurance accomplishes nothing in our model and is not part of an optimal banking arrangement. Its existence, therefore, must be attributed to exogenous (“political”) reasons despite its social costs. We recognize that there are modeling environments in which deposit insurance can be part of an optimal banking arrangement, but we also recognize that there is substantial empirical evidence consistent with the idea that deposit insurance may, on net, be a socially costly mechanism that is chosen for political, not economic, reasons (Demirguc-Kunt, Kane and Laeven 2008, Calomiris 2011).

8 Conclusion

We consider the role that cash reserves play in stabilizing banks. Risk-based capital requirements define a margin of substitution between cash and equity, which reflects the role of cash in reducing asset risk. In our model, however, cash plays a unique role, and is not just useful as a risk reducer in the mechanical sense. Compared to book capital, cash has unique advantages, including the ease with which it can be observed, and the inability for bankers to unobservably increase its riskiness. The fact that cash is impervious to “risk shifting” makes its possession by banks uniquely useful in promoting good risk management by banks on their risky assets. Banks that hold sufficient cash are able to gain market confidence in their risk management, and thereby attract and retain deposits. In bad states of the world (i.e.,
recessions) banks that might otherwise be tempted to allow risky assets to become riskier will raise their cash holdings to preserve market confidence in their low risk.

While our initial model emphasizes this prudential role of cash reserves, rather than their role in mitigating exogenous liquidity risks, we also consider the effect of adding liquidity risk to our model. When liquidity risk is added to our autarkic model of banking, it raises the optimal amount of reserves held by banks. As before, in the autarkic equilibrium, banks voluntarily choose to hold the optimal level of reserves.

In a multi-bank model, which includes diversifiable exogenous liquidity risk, there is a role for the group of banks to contract to share liquidity risk. Banks are able to insure effectively against liquidity risk through a clearing house arrangement in which the clearing house recycles deposits that exogenously move from one bank to another. To prevent abuse of clearing house protection, reserve requirements must be instituted to force members to hold adequate reserve balances. Because liquidity risk is diversifiable, we show that the formation of the clearing house economizes on reserve holdings in comparison with autarkic (self-insurance) of liquidity risk.

We also consider the effects of government-financed deposit insurance, which is not welfare-improving in our framework. Deposit insurance creates a potential moral hazard because it removes incentives for private monitoring. In the presence of deposit insurance, optimal prudential regulation will set the cash reserve ratio higher than it would be in the absence of deposit insurance. That outcome, in spite of the absence of liquidity risk under deposit insurance, reflects the fact that in our framework cash reserves primarily serve a prudential role.
References


