"Liquidity requirements, liquidity choice and financial stability"*

Douglas W. Diamond

Anil K Kashyap

University of Chicago Booth School of Business and National Bureau of Economic Research

revised October 2015

Abstract

We study a modification of the Diamond and Dybvig (1983) model in which the bank may hold a liquid asset, some depositors see sunspots that could lead them to run, and all depositors have incomplete information about the bank's ability to survive a run. The incomplete information means that the bank is not automatically incentivized to always hold enough liquid assets to survive runs. Regulation similar to the liquidity coverage ratio and the net stable funding ratio (that are soon be implemented) can change the bank's incentives so that runs are less likely. Optimal regulation would not mimic these rules.

* Prepared for the Handbook of Macroeconomics. We thank Franklin Allen, Gary Gorton, Guido Lorenzoni, Annette Vissing-Jorgenson, Nancy Stokey, Nao Sudo John Taylor, Harald Uhlig and seminar participants at the Asian Development Bank Institute, East Asian Economic Seminar, Imperial College, National Bureau of Economic Research Monetary Economics meeting, and the University of Chicago for helpful comments and Adam Jorring for expert research assistance. We thank the Initiative on Global Markets at Chicago Booth, the Fama Miller Center at Chicago Booth and the National Science Foundation for grants administered through the NBER for research support. All errors are solely our responsibility.

Introduction

In September 2009 the leaders of 20 major economies created the Financial Stability Board (FSB) whose purpose is to "coordinate at the international level the work of national financial authorities and international standard setting bodies (SSBs) in order to develop and promote the implementation of effective regulatory, supervisory and other financial sector policies." Since that time the financial system has undergone a regulatory overhaul. Much of the public attention has focused on changes to the rules regarding capital requirements for banks. Yet by 2019, via the Basel Committee on Bank Supervision, the major economies have also agreed also to implement new rules governing banks' debt structures and requirements to hold certain types of liquid assets.

To date there is a remarkable asymmetry in the economic analysis of the capital and liquidity regulations. The pioneering work of Modigliani and Miller (1958) provides a solid theoretical framework for analyzing capital regulation. Any student taking a first course in corporate finance will encounter this theory and there is a massive empirical literature that explores the theory's predictions. Capital regulations for banks at the international level go back to 1988 and there many empirical examinations of the impact of these regulations.

In contrast, there is no benchmark theory regarding liquidity provision by intermediaries. Indeed, financial economists even have competing concepts that they have in mind when discussing liquidity, so that there is no generally accepted empirical measure of liquidity economists study. Allen (2014), in his survey of the nascent literature on liquidity regulation, concludes by writing "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."

Nonetheless, the global regulatory community has agreed on certain liquidity requirements (Basel Committee on Bank Supervision (2013, 2014)). Two new concepts, the liquidity coverage ratio and the net stable funding ratio, have been proposed and banks by 2019 will be compelled to meet requirements for these ratios. Thus, it seems fair to say we are in a situation where practice is ahead of both theory and measurement.

In this paper we survey the existing work on liquidity regulation and develop a framework for discussing the regulation. The theory that we propose suggests, in certain parameterizations, regulations bearing some resemblance to the liquidity coverage ratio and net stable funding ratio can emerge as ones which will improve outcomes relative to an unregulated benchmark. However, the regulations that arise in our model would naturally differ across banks, depending on certain bank characteristics, so they do not mimic exactly the ones that are on track to be implemented.

The critical ingredients in our model are the following. First, we consider banks which are spatially separated and hence do not compete aggressively for deposits. Treating the bank as monopolist simplifies the analysis by allowing us to side-step some complications that arise from having to model the deposit market equilibrium. The model can also be interpreted as a description of the aggregate banking system, which for many financial stability and regulatory discussions is the object of primary concern and under this interpretation ignoring the deposit competition is perhaps more natural.

Second, we assume that intermediaries provide liquidity insurance for customers who have uncertain withdrawal needs (or consumption desires). We build on the Diamond and Dybvig (1983), henceforth DD, model of banking in which banks provide this insurance by relying on the law of large numbers to eliminate idiosyncratic customer liquidity needs. For those familiar with DD, we make two modifications. We allow the bank to invest in a liquid asset that has a positive rate of return and can be used to pay customers that need liquidity. This introduces a tradeoff between lending and holding liquidity as in Bhattacharya and Gale (1987) and several papers of Allen and Gale (1997 and others).

The other change from DD is the form of run risk that the banks face. Banks are assumed to have a good assessment of the aggregate needs of their customers for fundamental reasons. But, they also know that some customers will receive a signal about the bank which could lead to a run. The sunspots that we consider are a metaphor for people being concerned with the health of the bank, but not having a fully formed set of beliefs about the bank's solvency status. In making their decisions we assume that customers are unable to fully evaluate the ability of the bank to honor deposits. Given the complexity of modern banks it seems realistic to presume that most customers cannot precisely determine their bank's maturity mismatch and hence its vulnerability to a run. The imperfect information creates a challenge for the banks because their customers will not necessarily know if the bank is prudently holding liquidity or not, which reduces the incentive to hold liquidity.

In the event that a run does occur, we depart from DD and allow for the possibility that not all customers seek to withdraw their funds. We believe it is useful to analyze partial runs for two separate reasons. One is that in practice there do seem to be some sticky deposits that do not flee even in times of considerable banking stress. In addition, even before troubles occur it is usually clear which types of deposits are prone to running. So this allows us to talk about policies for different types of withdrawal risk.

Within this environment we can assess the vulnerability of the financial system to runs under different regulatory arrangements. In the baseline case, we assume that banks simply maximize their profits and see which types of equilibria arise. As usual in DD style models, the outcomes depend critically on how depositors form beliefs. It is possible, under certain parameter configurations, that the pure self-interest motives of the banks will sufficient to insure that the system will be run proof.

Given that depositors cannot be sure about how robust the banks are, the banks will typically face a tension in deciding how much to fortify themselves against the risk of a run. They can always choose to be sufficiently conservative to be able to withstand a worst case scenario. But in order to do that, they will engage in very little lending, and the forgone profits from deterring the run will be high. Hence, it is possible they will make more profits from taking more risk and living with the consequence that they may be wiped out.

We next allow regulatory interventions that place restrictions on bank portfolio choices. In the baseline set up, the banks have perfectly aligned incentives to prepare to service fundamental aggregate withdrawal needs. So the regulatory challenge is to determine whether a requirement that distorts their private incentives towards being more robust to a run will improve outcomes. We allow for regulation that is inspired by the two impending Basel rules.

One variant requires an initial liquidity position that must be established before depositors make their intentions clear. This can function like the "net stable funding ratio" that is proposed as part of the Basel reforms. A second option is a mandate to always hold additional liquid assets beyond those needed for the fundamental withdrawals. This regulation looks like a traditional reserve requirement for the bank, but can also be interpreted as a kind of "liquidity coverage" ratio that is part of the Basel reforms.

One point of contention regarding the liquidity coverage ratio that has emerged is whether required liquidity can be deployed in the case of crisis. Goodhart (2008) framed the issue nicely with a now famous analogy of "the weary traveller who arrives at the railway station late at night, and, to his delight, sees a taxi there who could take him to his distant destination. He hails the taxi, but the taxi driver replies that he cannot take him, since local bylaws require that there must always be one taxi standing ready at the station."

One way to interpret the Goodhart conundrum is to recognize that, broadly speaking, there are two ways to think about the purpose behind liquidity regulations. One motivation can be to make sure that banks can better withstand a surge in withdrawals should one occur. From this perspective mandating that the last cab cannot depart the station seems foolish. Another possible motivation is to design regulations aimed at reducing the likelihood of a withdrawal surge. Our model helps highlight the potential incentive properties of regulation and can potentially explain why mandating the presence of some unused liquidity could be beneficial.

In studying how private and social incentives for liquidity choices diverge, our main conclusion from analyzing the two Basel-style regulations is that they may improve outcomes relative to the ones that arise from pure self-interest, but each brings potential inefficiencies. Hence, we also describe the solution of the mechanism design problem for a social planner who has less information about withdrawal risk than the bank does and seeks to optimally regulate banks to avoid runs. That solution provides a natural benchmark against which to judge the Basel-style regulations.

The remainder of the paper is divided into five parts. Section two contains our selective overview of previous work. As mentioned already, there is enormous and rapidly growing literature on capital regulation. We note several surveys on pure effects of capital regulation. Our emphasis is instead on papers that focus specifically on liquidity regulation.

Section three introduces the benchmark model. We explain how it works under complete information. We also derive a generic proposition that holds with incomplete information that describes when the bank's preferred liquidity choice will be sufficient to deter a run. Generically, however, privately chosen levels of liquidity need not be sufficient to deter runs. So this opens the door for regulations that might do so.

In section four we analyze the two types of liquidity regulation that are akin to the ones contemplated under the Basel process. We first demonstrate that a particular type of regulation that requires the bank to hold liquid assets equal to a fixed percentage of deposits at all times can potentially deter runs. This works because the liquidity mandate, combined the bank's self-interest to prepare to service predictable deposit outflows, leads the bank to hold more overall liquidity than it would otherwise. Because depositors understand this, it removes the incentive to run in some cases. We also consider alternative assumptions about depositors' knowledge and the information available to regulators and assess the vulnerability of the bank to runs in these scenarios.

In section 5, we describe several extensions of the baseline model. The first sketches a mechanism design problem where the regulator does not have all of the bank's information and seeks to implement run-free banking. We fully characterize the solution to this problem in Diamond and Kashyap (2015), here we describe the main findings from this exercise. It turns out that a regulator with sufficient tools can induce the bank to hold the proper amount of liquidity despite the private information advantage possessed by the bank.

We next explore the effects of allowing for competition in the deposit market. Once banks can compete for deposits, a bank may opt to try to attract deposits from competitors to satisfy withdrawals rather than choosing to hold liquidity. What banks believe about the likely success of such a strategy will govern whether regulations need to be adjusted for this possibility.

Finally, we briefly discuss capital regulation. We first explain why as a tool for managing liquidity problems, capital requirements can be relatively inefficient compared to the other regulations that we have reviewed. Obviously in a richer model where both credit risk and liquidity risk are present, capital and liquidity regulations can serve different purposes. We describe a couple of these differences.

Section six presents our conclusions. Besides summarizing our findings, we also pose a few open questions are natural next steps to consider in addressing the issues analyzed in this paper.

2. Literature Review

In considering capital regulation, the literature can be organized by sorting papers along two dimensions. The first regards what is assumed regarding the Modigliani-Miller (1958) (henceforth MM) capital structure propositions. As in all models of corporate finance, absent failures of one of the MM propositions any choices regarding capital structure will be inconsequential. There have been four primary MM violations that have drawn attention in the literature.

One concerns that existence of deposit insurance. If certain parts of a bank's capital structure is protected from losses by the government, that can create risk-shifting incentives for equity holders. In many models, bank managers working on behalf of the equity owners face an incentive to gamble after adverse shocks that goes unchecked because depositors are immune from losses that they would suffer if the gamble fails.

A second distortion is concerns over guarantees to protect equity holders of banks from losses. Usually this is couched as a problem of having some banks that are assumed to be "too big" or "too-interconnected" to fail. But, in the recent global financial crisis, there were also cases in some countries where equity owners of smaller, non-systemic banks were insulated from losses due to political connections.

A third violation regards the MM assumption of complete financial markets. With incomplete markets, an institution that creates new securities could be valuable. In the banking context, deposits are a leading example of special security that banks might create.

Finally, there are many models where either asymmetric information or moral hazard problems are considered. Some of the prominent examples include the possibility that borrowers know more about their investment opportunities than lenders, or that borrowers can shift the riskiness of their investments after receiving funding.

So unlike much of the literature research on non-financial corporations, the trade-off theory of capital structure, whereby firms prefer debt for its tax advantages and balance those benefits against costs of financial distress, has not figured prominently in the banking research on capital regulation. Rather, regulation is usually justified on the grounds of addressing one of these other four problems. The type of regulation that can be welfare improving will differ depending on which of these other frictions is assumed to be present.

The second important dimension one which the literature can be organized concerns the economic services that banks are assumed to provide. [1] Broadly, there are three types of services that have been modeled. The first presumes that certain financial institutions can expand the amount of credit that borrowers can obtain (say, relative to direct lending by individual savers).

[1] The next few paragraphs are taken from Kashyap, Tsomocos and Vardoulakis (2014)

The micro-founded theories typically assume that borrowers can potentially default on loans and so any lender has to be diligent in monitoring borrowers (Diamond (1984)). By concentrating the lending with specialised agents, these monitoring costs can be conserved and the amount of credit extended can be expanded.

A second widely posited role for intermediaries is helping people and businesses share risks (Benston and Smith (1976), Allen and Gale (1997)). There are many ways to formalise how this takes place, but perhaps the simplest is to recognise that because banks offer both deposits and equity to savers, they can create two different types of claims that would be backed by bank assets. These two choices allow savers to hedge some risks associated with lending and this hedging improves the consumption opportunities for savers. More broadly, these theories suppose that banks help pool and tranche risks.[2]

A third class of models, which complements the second, supposes that the financial system creates liquid claims that facilitate transactions. There are various motivations behind how this can be modelled. In DD style models, an intermediary can cross-insure consumers' needs for liquidity by exploiting the law of large numbers among customers. But doing so exposes banks to the possibility of a run, which can be disastrous for the bank and its borrowers and depositors. Calomiris and Kahn (1991) and Diamond and Rajan (2001) explain that the very destructive nature of a run is perhaps helpful in disciplining the bank to work hard to honour its claims. So the fragility of runs is potentially important in allowing both high amounts of lending and large amounts of liquidity creation.

Depending on which of these three services is presumed to be operative, and which of the MM failures are present, one can reach very different conclusions about the efficacy of capital regulation in improving welfare. For instance, in models where liquidity creation is not one of the services provided by banks, the costs of mandating higher amounts of equity financing are often modest. Likewise, the benefits of protecting taxpayers from having to bail out banks or depositors by forcing more equity issuance are potentially substantial.

Martynova (2015) offers a recent, comprehensive survey of this literature we will not duplicate this summary. However, it is worth emphasizing that many of these papers are not very informative regarding liquidity regulation, or the potential interactions of liquidity and capital regulation because in the environment being analyzed there is no value to liquidity creation (and hence no cost to limiting it).

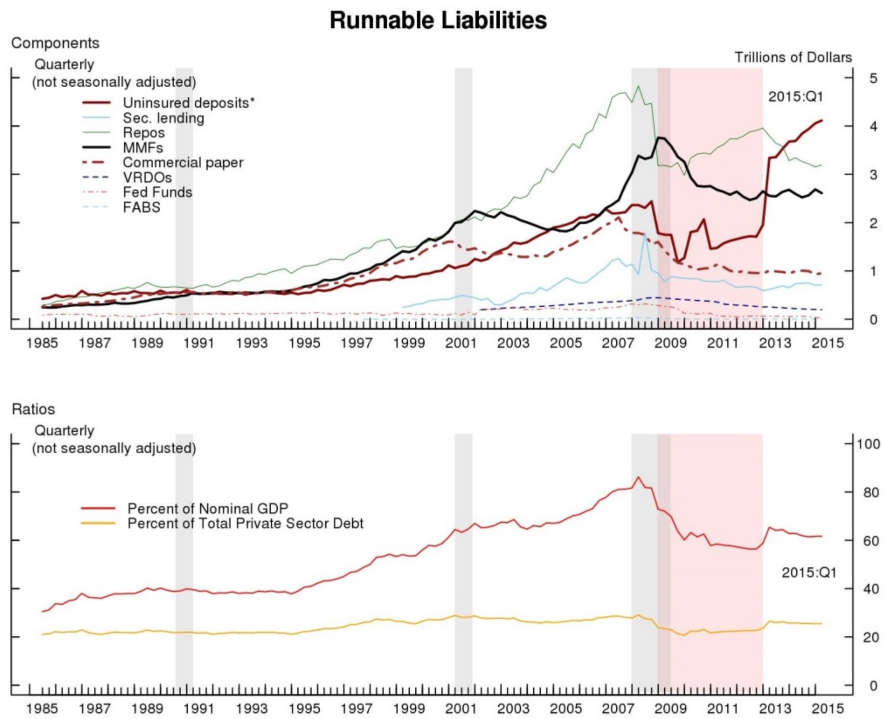
As mentioned in the introduction, there are far fewer papers that seek to investigate the purpose and effect of liquidity regulation. Allen (2014) offers a survey of this nascent literature and we share the sentiment of the concluding paragraph of his survey. He writes, "much more research

[2] For instance, if there are transactions costs associated with buying securities, a bank that makes no loans but holds traded securities could still be valuable.

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."

It is possible to again use a similar kind of two-way to classification regarding capital regulation to describe much of the thinking on liquidity. Trivially, if the economic services offered by a bank do not include the provision of liquidity, then regulation that focuses on liquidity will not be particularly interesting to consider. It is possible that in such environments regulating liquidity could make sense to achieve other aims, such as supplementing or substituting for capital requirements. However, if maturity transformation is not one of the outputs of the financial system, assessments of the efficacy of liquidity regulation in such models will be incomplete. Put bluntly, if there are no costs to limiting liquidity provision per se, then obviously the cost of regulations that have this effect cannot be fully assessed.

It is worth noting that will most of the literature on liquidity and liquidity regulation label the institutions that that undertake this activity as "banks". However, as became evident in the global financial crisis this activity is hardly limited to banks. Figure 1, reproduced from Bao, David and Hong (2015) shows the total amount of runnable funding inside the U.S. financial system over the past 30 years.



* Uninsured deposits equal the difference between total deposits and insured deposits. The quarterly insured deposits series between 1985 and 1990 is obtained by interpolating the available annual data. For 2008:Q4-2012:Q4 (red shades), insured deposits increased due to the Transaction Account Guarantee (TAG) Program. For 2008:Q4-2009:Q2, some insured deposits were not accounted for because the FDIC did not collect data on insured amounts for those TAG accounts with balances between $100,000 and $250,000.
Note: The gray shades, which overlap the red shades, indicate NBER recession dates.
Source: Staff calculations using data from RMA, DTCC, SIFMA, Call Reports, Financial Accounts, M3 monetary aggregates, and Bloomberg Finance LP.

Figure 1: Bao, David and Han (2015) Estimates of Runnable Funding in the U.S.

We draw three conclusions from their estimates that are worth bearing in mind throughout the rest of the discussion. First, there has been a sizable increase in the amount maturity transformation over the last 20 years. From 1995 until 2015, the scale of such activity rose by 50% as measured relative to Gross Domestic Product. Second, as far back as 1985 as much of this activity has occurred outside the banking system as inside it. Third, the decline immediately after the GFC was sizable. The drop in repurchase agreements and money market funds were especially pronounced, but even as a percent of GDP, the level in 2015 is very similar to the level in 2005 (just before the frenzied period ahead of the GFC). Hence, maturity transformation is still happening on a substantial scale even after the GFC and all of the various regulatory reforms that have been introduced.

Given this evidence, we focus only on papers where one of the services of the financial system is to provide liquidity. Among these it is helpful to separate them into papers that model liquidity provision in the same way or similarly to DD, and those that introduce other mechanisms.

Among the DD style models, we focus on three that are closely related to our analysis. Ennis and Keister (2006) have a DD style model and determine how much liquidity banks need to hold to deter runs. They compute the amount of excess liquidity the bank must hold to buffer it against a run by all depositors, and also determine the optimal amounts to promise depositors. In their model with full information, when depositors desire safe banks, there will be private incentives to hold enough excess liquidity to deter a sunspot-based run. They do not study regulation because there is no need for any under their assumptions, but we will see that some of the same forces that are present in their model arise in ours.

Vives (2014) analyzes a question similar to that in Ennis and Keister (2006): what are the efficient combinations of equity capital and liquidity holdings to make a bank safe when it subject to runs based on private information about its solvency? He studies a global game where a bank can be insolvent or illiquid. The need for regulation is not considered explicitly, but he does examine what capital and liquidity levels would make the bank safer. He finds that capital and liquidity are differentially successful in attending to insolvency and illiquidity. In particular, if depositors are very conservative (and which makes them more inclined to run in the model), increased liquidity holdings which reduce profits by investing more in liquid assets can enhance stability.

Farhi, Golosov, Tsyvinski (2009) investigate a DD model where consumers need banks to invest and where the consumers can trade bank deposits. Absent a minimum liquidity regulation, it is profitable to free ride on the liquidity held by other banks, because banks offer rates which subsidize those who need to withdraw their deposit early (which is the spirit of Jacklin(1987)). A floor on liquidity holdings removes the incentive for this free riding.

Among the non-DD models, one that is related is Calomiris, Heider, Hoerova (2014). They have a six period model where banks can potentially engage in risk-shifting so that when banks suffer loan losses they may not be able to honor their deposit contracts. Cash is observable and mandating that banks must have minimum levels of cash reserves can limit the risk-shifting.

More generally, our approach is closely related to the mechanism design approach to regulation of monopolists in Baron and Myerson (1982). They also were interested in investigating how regulation could be structured to induce the party being regulated to efficiently use information that is private.

3. Baseline Model

The timing and preferences are as in DD. There are three dates, T= 0, 1, and 2. The interest rates that bank must offer are taken as given, motivated by a monopoly bank which must meet the outside option of depositors to attract deposits. Equivalently, the single bank can be thought of as representing the overall banking system.

For a unit investment at date 0, the bank offers a demand deposit which pays either r_1 at date 1 or r_2 at date 2. This effectively offers a gross rate of return r_2/r_1 between dates 1 and 2 which is equal to the exogenous outside option (such as government bonds) for depositors between these dates. Essentially, the bank offers one period deposits which equal the interest rate on the outside option. We will assume that depositors are sufficiently risk averse that they would like the banking system to supply one period deposits that are riskless. Hence, when we consider interventions they will be designed to deliver as this as the only possible equilibrium.

The residual claim after deposits are paid is limited liability equity retained by the banker. All equity payments are made at date 2.[3]

The bank can invest in two assets with constant returns to scale. One is a liquid asset (which we will interchangeably refer to as the safe asset) that returns $R_1>0$ per unit invested in the previous period. The other is an illiquid asset for which a unit investment at date 0 returns at date 2 an amount that exceeds the return from rolling over liquid assets ($R_2> R_{1*} R_1$). The illiquid asset (which we will interchangeably refer to a loan) can be liquidated for θR_2 date 1, where $\theta R_2< R_1$ and $\theta\geq 0$. These restrictions imply that when the bank knows it must make a payment at date 1, it is always more efficient to do that by investing in the safe asset rather than planning to liquidate the loan.

We also assume that banking is profitable even if the bank invests exclusively in the liquid asset, so that $r_1<R_1$ and $r_2<R_1^2$. This is a sufficient condition to guarantee that requiring excess liquidity will not make the bank insolvent (though it still will reduce the efficiency of investment). In addition, we assume that bank profits from investing in illiquid assets when depositors hold their deposits for two periods is greater than from investing in liquid assets when

[3] We could introduce another incentive problem for the banker to motive a minimum value of equity at all dates and states, but for now the bank will operate efficiently as long as equity remains positive in equilibrium.

depositor's hold their deposits for only one period (or $\frac{r_2}{R_2} < \frac{r_1}{R_1}$). The latter assumption is used only in some results on optimal regulation.

There are many possible reasons to presume that the illiquid asset can be liquidated for only θR_2. For instance, in DD liquidation can be thought of as a non-tradable production technology. Alternatively it could reflect the bank's lending skills, implying that it would be worth less to a buyer than to the bank because (compared to the bank) the buyer would be able to collect less from a borrower, as in Diamond-Rajan (2001). Nothing in our analysis hinges on why this discount exists, though we do insist that it is operative for everyone in the economy including a potential lender of last resort. Also, our assumption that θ is a constant implies that we are not modeling a situation where the sale price depends only on the amount of remaining liquidity held by potential buyers (as in Bhattacharya-Gale (1987), Allen-Gale (1997) and Diamond (1997)).

If there were uncertainty regarding θ, and its value could be learned before assets had to be liquidated, then there could be an additional option value associated with owning liquid assets if they could be used to buy time to learn about θ. This channel is foreclosed in our set-up, so the only role for liquidity requirements is to provide incentives for the bank to alter its asset investment decisions.

For fundamental reasons, a fraction t_s of depositors want to withdraw at date 1 and $1-t_s$ want to withdraw at date 2 in state s. The realizations of t_s are bounded below by $\underline{t} \geq 0$ and above by $\overline{t} \leq 1$ The banker will know the realization of t_s when the asset composition choice is made. This assumption is meant to capture the fact that banks have superior information about their customers. Indeed, some early theories of banking supposed that the advantage of tying lending and deposit making was that by watching a customer's checking account activities a bank could gauge that customer's creditworthiness (Black(1975)). So the assumption that banks are well-informed about their customers' withdrawal tendencies has a long tradition.

If the ex-post state is s and there is not a run, a fraction $f_1=t_s$ will withdraw r_1 each, requiring r_1t_s in date 1 resources, this will leave a fraction $1-t_s$ depositors at date 2 and the required date 2 resources to pay them equals $r_2(1-t_s)$. If we let α be the fraction of the bank's portfolio that is invested in the liquid asset and $(1-\alpha)$ be the portion invested in the illiquid one, then the bank's profits, and hence its value of equity in general will be

$$\text{Value of equity} = (1\text{-}\alpha)R_2 + (\alpha R_1 - f_1 r_1)R_1 - (1-f_1)r_2 \text{ if } \alpha R_1 \geq f_1 r_1 \quad (1)$$

Because we are assuming that the bank knows t_s, its own self-interest will lead it to make sure to always have enough invested in the liquid asset to cover these withdrawals. So absent a run, the profits are very intuitive and easy to understand. The first term in (1) represents the returns from the illiquid investment, the second reflects the spread on the safe asset relative to deposits (recognizing that any leftover funds are rolled over), and the third term reflects the funding costs of two period deposits.

The more interesting case to consider is what happens when a run is possible. We suppose that a fixed number Δ of the patient depositors are highly likely to see a sunspot. All depositors (and the bank) know Δ and upon seeing the sunspot they must decide whether they believe that the others who see it will decide withdraw their funds early. As mentioned earlier, the sunspot is intended to stand in for general fears about the solvency of the bank, so the inference problem relates to their conjecture about whether others investors might panic. In that case, they have to decide whether to join the run. So in general $f_1 > t_s$ is possible.

If the bank will be insolvent with a fraction of withdrawals of any amount less than $t_s + \Delta$, then we assume each depositor who sees then sunspot will all withdraw and $f_1 = t_s + \Delta$. This will give zero to all who do not withdraw, and the goal of bank or its regulator is to prevent this outcome from ever being a Nash equilibrium.

In addition, we will assume that if the bank is exactly solvent at $f_1 = t_s + \Delta$, no depositor who does not need to withdraw (and only sees the sunspot) will withdraw. This condition establishes how much liqudity is needed to deter a run. A natural justification for why someone who does not need to withdraw until date 2 would not want to withdraw from a solvent bank at date 1, is that by doing so the person can only obtain a lower return than they get from the bank. Put differently, we suppose that self-insurance such as sticking money under a mattress is not an attractive option.

If the bank observed the fraction of people withdrawing their funds (for fundamental reasons or because of panic), prior to choosing its liquid asset holdings, then it would always invest enough to cover the planned withdrawals. This directly follows from the assumption about the inefficiency of liquidating illiquid assets to pay early depositors relative to using the liquid asset. Holding just enough liquid assets pay off exactly $f_1 r_1$ is profit maximizing because doing so maximizes its long-term investments (which are its most profitable option). The profits for the bank in this case will again be given by (1).

Proposition 1: The bank optimally chooses $\alpha^{AIC} = \frac{t_s r_1}{R_1}$, under the belief that no one runs,

and if $t_s + \Delta < \dfrac{t_s r_1 + (1 - \frac{t_s r_1}{R_1})\theta R_2 - r_2\theta}{r_1 - r_2\theta}$, investors will never run.

Proof: The bank's equity is positive when $\left((1-\alpha) - \frac{(f_1 r_1 - \alpha R_1)}{\theta R_2}\right) R_2 - (1-f_1) r_2 \geq 0$. So, when the automatically incentive compatible level of initial liquidity α^{AIC} is chosen ($\alpha^{AIC} = \frac{t_s r_1}{R_1}$), the value of equity is decreasing in f_1 and equal zero when $f_1^* = \frac{\frac{t_s r_1}{R_1} R_1 + (1-\frac{t_s r_1}{R_1})\theta R_2 - r_2\theta}{r_1 - r_2\theta}$. Therefore, if $t_s + \Delta$ is less than f_1^*, then the depositors always know the bank will be solvent and there is no Nash equlibrium with a run.

QED

The proposition simply states the condition when the bank is sufficiently profitable and not too illiquid, even if it only holds enough of the liquid asset to service fundamental withdrawals, it would still be solvent in the event of a run. If the depositors know Δ and t_s, then they can check whether this condition holds and if it does, then it never is individually rational to react to the sunspot.

If $t_s + \Delta > \frac{t_s r_1 + (1-\frac{t_s r_1}{R_1})\theta R_2 - r_2\theta}{r_1 - r_2\theta}$ then the bank must increase α_s to α_s^{stable} to definitely deter the run, where α_s^{stable} is such that $t_s + \Delta = \frac{\alpha_s^{stable} R_1 + (1-\alpha_s^{stable})\theta R_2 - r_2\theta}{r_1 - r_2\theta}$. This yields $\alpha_s^{stable} = \frac{(t_s+\Delta) r_1 + \theta((1-t_s-\Delta) r_2 - R_2)}{R_1 - \theta R_2}$.

So merely preparing to service fundamental withdrawals will not always be enough to deter a run.

One useful benchmark is to notice that when runs can be complete, $t_s + \Delta = 1$, then the amount needed to necessarily deter a run simplifies to $\overline{\alpha}^{stable} = \frac{r_1 - \theta R_2}{R_1 - \theta R_2}$. The result in this special case seems particularly intuitive since it says that if the bank is at risk for a total run, it can only prevent that by promising to pay out to its early customers no more than it is earning on the liquid investment. This should not be surprising because in a full run none of the rents from funding the illiquid investment for two periods will be realized so that the bank cannot use any of those rents to subsidize the early withdrawals (as occurs in the more general case).

This means we can say a bit more if t_s is known and all parties understand the bank's incentives. If runs are impossible, the bank will choose $\alpha_s = \frac{t_s r_1}{R_1}$ and no unused liquidity is held from dates 1 to 2. When runs are possible and must be avoided, then this level of liquidity holdings might not be high enough to leave the bank solvent in a run. To always deter a possible run bank will have to hold $\alpha_s=\alpha_s^{stable}$. This will require that some unused liquidity be held from date 1 to 2 (after normal withdraws are met at date 1). If the bank is free to use all of this unused liquidity if a run should occur, then depositors can see that the liquidity is present and will never choose to run.

Alternatively, suppose the depositors do not observe α_S, but the bank and regulator (or planner) could. Then the following arrangement is possible.

Proposition 2: With full information a bank (or a regulator) seeking to deter runs will choose $\alpha_s^*=\max\{\alpha_s^{AIC},\alpha_s^{stable}\}$

Proof: The bank is automatically stable when $\frac{r_1 t_s}{R_1} \geq \alpha_s^{stable}$ so the regulator would always want to maximize lending and allow the bank to follow its self-interest and select that level ($\frac{r_1 t_s}{R_1}$) of liquidity. Otherwise, the minimum amount of liquidity that is needed is α_s^{stable}.

More generally, for arbitrary anticipated withdrawals of t_S, α_s^{AIC} and α_s^{stable} will differ and will look something like what is shown in Figure 2. For very low levels of anticipated withdrawals, where the condition in Proposition 1 holds, the bank is sufficiently solvent that chooses to hold more ex-ante liquidity than is need to be stable, so that runs are impossible. At some point, however, this ceases to be true and the amount needed to just be solvent in a run is higher than the bank would hold out of pure self-interest. So in this case run deterrence would require a higher level of initial liquidity. This observation will be helpful in understanding some of the regulatory tradeoffs that we subsequently explore.

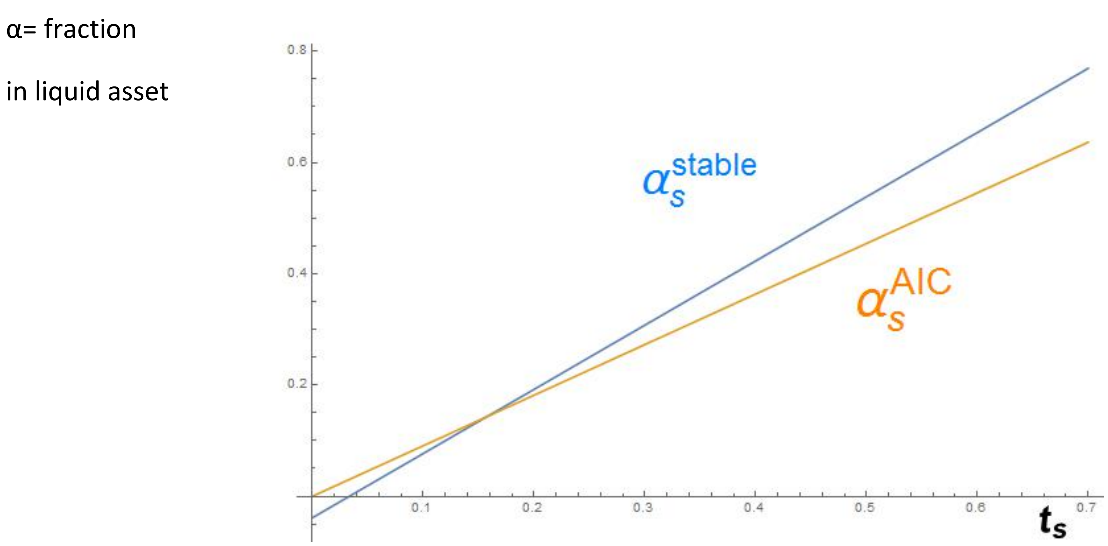


Figure 2: Comparison of Automatically Incentive Compatible and Necessarily Stable Liquidity Choices

While the full information benchmark is helpful, we think it may be too extreme to be realistic. It is plausible that the bank knows much more about its customers than anyone else. Moreover, depositors are likely to have trouble interpreting α. They may only be able to access stale information and also may not be able to tell how many withdrawals had occurred as of the time that α was measured.

So we need to understand what happens when either α is unobserved, or when t_s is not observed. Remarkably, if t_s observed but α is not, so long as $\alpha_s^{AIC} \geq \alpha_s^{stable}$ runs are deterred by bank self-interest. If t_s is unknown, then the runs will only be automatically deterred if this condition is true for all s.

Provisionally, we will assume that α is unobservable. One way to interpret this is that liquidity is disclosed only on single date (such as year end), and that it can be altered quickly on other dates. We will eventually relax that assumption, though it will turn out that most of the key intuitions and results still carry through even under weaker assumptions. So as our baseline assumptions are as follows:

i) Depositors cannot observe α, but do know t_s and Δ and they know that some of the patient depositors are highly likely to see the sunspot. They also know that the bank maximizes its profits conditional on what it knows.

ii) The bank knows t_s when it picks α. The bank knows Δ and it knows that the depositors know what it knows.

Our analysis seeks to implement choices by the bank where there is never a bank run. Under these assumptions, when the required condition on $t_s + \Delta$ in Proposition 1 does not hold, there is essentially never an equilibrium where the bank is always solvent during a run.[4] The problem arises because if bank anticipates that depositors will never run, the bank will not hold any extra liquidity to deter a run. We show in the appendix that even with mixed strategies by depositors, this remains true.

4. Basel Style Regulatory Options

In cases where self-interest by the bank will not automatically eliminate runs, we next ask whether some simple forms of regulation might do so. For these purposes, we continue to assume that the depositors are sufficiently risk averse so that this is social optimum. We recognize that one consequence of trying to eliminate runs is that we are not maximizing lending,

[4] If for some reason all depositors always chose to ignore the sunspot, that could be a pure strategy equilibrium where runs are never anticipated, but in addition to being uninteresting in our context, it would not be unique because the equilibrium we describe in the appendix would also always exist.

which might be another possible social objective. Given that the model does not have fully endogenous general equilibrium interest rates we hesitate to use it to explore situations where depositors are less risk averse which would imply tradeoffs between arrangements that might deliver extra lending at the expense of additional run risk. This would be a natural topic for further study; nonetheless, we believe the insights from this model on how different regulations operate would still carry over to a richer environment.

We consider two potential approaches that a regulator (who at this stage could also be described as an auditor) could pursue. These are inspired by the kinds of regulations that are proposed as part of Basel III. We suppose that she can credibly certify that the bank has some level of the safe asset is present (as a percentage of deposits). One option is to report on this ratio at the time when the liquid assets are acquired at time zero. This would amount to regulating α and this is similar in spirit the net stable funding ratio (NSFR). The NSFR requires "banks to maintain a stable funding profile in relation to the composition of their assets and off-balance sheet activities" (Basel Committee on Bank Supervision (2014)). Loosely speaking, the NSFR can be thought of as forcing banks to match long term assets with long term funding

Alternatively, she could insist that the bank will always have a certain amount of liquid assets relative to deposits at all times, including after any withdrawals. This kind of regulation is more like the liquidity coverage ratio (LCR). The LCR requires "that banks have an adequate stock of unencumbered high-quality liquid assets (HQLA) that can be converted easily and immediately in private markets into cash to meet their liquidity needs for a 30 calendar day liquidity stress scenario (Basel Committee on Bank Supervision (2013)).

As a first step, consider imposing a LCR regulation that says the bank must always hold a fraction ρ of deposits in liquid assets. The important consequence of this is that regulation would even apply after first period withdrawals (f_1), when the bank would have to have a minimum level of safe assets equal to $\rho(1-f_1)$.

If the bank is subject to this requirement, and it conjectures that f_1 depositors will withdraw, then its optimal initial level of safe assets (α) will satisfy $\alpha R_1=f_1 r_1+\rho(1-f_1)$. This choice follows trivially because it is never efficient to make loans with intention of liquidating them and this is the minimum amount of liquid assets that will satisfy the regulation. Accordingly, the bank knows that the depositors will know this (and also understand that the bank is trying to maximize its profits). The residual value of the bank's equity will be:

$$E_2(f_1;.,\rho)=\begin{cases}(\alpha R_1-f_1r_1)R_1+(1-\alpha)R_2-(1-f_1)r_2 & \text{if } f_1<\dfrac{\alpha R_1-\rho}{r_1-\rho},\\[2ex] \left((1-\alpha)-\dfrac{f_1r_1-\alpha R_1+\rho(1-f_1)}{\theta R_2}\right)R_2+(\rho R_1-r_2)(1-f_1) & \text{if } f_1\geq\dfrac{\alpha R_1-\rho}{r_1-\rho} \text{ and if}\\ & f_1\leq\dfrac{\alpha R_1+(1-\alpha)\theta R_2-\rho(1-\theta R_1)-r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta},\\[2ex] 0 & \text{if } f_1>\dfrac{\alpha R_1+(1-\alpha)\theta R_2-\rho(1-\theta R_1)-r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta}.\end{cases}$$

Each branch of the expression is intuitive. The top branch shows the profits that accrue when withdrawals are small enough that the bank can pay all depositors without liquidating any loans and still satisfy the LCR; this will be the case whenever $f_1r_1<\alpha R_1+\rho(1-f_1)$, which when rearranged is the threshold condition that is listed. In this case, the bank has two sources of revenue, one coming from rolling over the residual safe assets after paying early depositors and the other coming from the return on the loans. The date 2 depositors must be paid and the banker keeps everything that is left.

The second branch represents a case where the bank must liquidate some loans to service the early withdrawals. In this case, the bank liquidates just enough loans so that after the deposits are paid, it exactly satisfies the LCR. The same two sources of revenues and deposit cost are present, but the formula adjusts for the liquidations. Recall that each loan that is liquidated yields θR_2 at date 1. Hence rather than having the revenue from the full set of loans $(1-\alpha)$ that were initially granted, the bank only receives returns on the portion that remains after some loans that were liquidated in order to pay the depositors and comply with the LCR. Because the LCR is binding from date 1 until date 2, the bank has exactly $\rho(1-f_1)$ of the safe asset that is rolled over and that money can also be used to pay the remaining patient depositors.

The third possibility is that the level of withdrawals is sufficiently large that the bank becomes insolvent. Insolvency occurs when $f_1>\dfrac{\alpha R_1+(1-\alpha)\theta R_2-\rho(1-\theta R_1)-r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta}$ because at that point the depositors can see that the liquidations do not generate enough to fully cover the promised repayments.

The bank knows that depositors consider all these possibilities in trying to infer what the bank will do. If the coverage ratio can be set such that the bank chooses to hold sufficient liquidity to remain solvent during a run, then runs will be deterred.

Proposition 3: If $\rho \in [0,1]$ satisfies $t_s + \Delta \le \dfrac{\dfrac{t_s r_1+\rho(1-t_s)}{R_1} R_1 + (1-\dfrac{t_s r_1+\rho(1-t_s)}{R_1})\theta R_2 - \rho(1-\theta R_1) - r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta}$ then a regulator who knows t_s can choose ρ so as to deter runs.

Proof: Suppose the bank believes that $f_1 = t_s$, then it will pick α to satisfy: $\alpha R_1 = t_s r_1 + \rho(1-t_s)$. The bank will be solvent until $f_1 = \overline{f}_1 = \dfrac{\alpha R_1 + (1-\alpha)\theta R_2 - \rho(1-\theta R_1) - r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta}$. So regulator can pick ρ such that it delivers $t_s + \Delta < \overline{f}_1$ and $\alpha_s = \dfrac{t_s r_1+\rho(1-t_s)}{R_1}$. If $\theta>0$ and $t_s + \Delta > \overline{f}$ at $\rho = 0$, then the lowest required ρ satisfies

$$t_s + \Delta = \frac{\dfrac{t_s r_1+\rho(1-t_s)}{R_1} R_1 + (1-\dfrac{t_s r_1+\rho(1-t_s)}{R_1})\theta R_2 - \rho(1-\theta R_1) - r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta} \qquad (2).$$

If such a $\rho \in [0,1]$ exists, then it is given by $\rho = \dfrac{\theta R_1((1-t_s-\Delta)r_2 - R_2) + r_1(\Delta R_1 + t_s\theta R_2)}{\Delta R_1 + (1-t_s-\Delta)\theta R_1^2 - (1-t_s)\theta R_2}$.

From our assumptions that $r_1 \le R_1$ and $r_2 \le R_1^2$, there will always be a be a value of ρ between 0 and 1 which satisfies $t_s + \Delta \le \dfrac{\dfrac{t_s r_1+\rho(1-t_s)}{R_1} R_1 + (1-\dfrac{t_s r_1+\rho(1-t_s)}{R_1})\theta R_2 - \rho(1-\theta R_1) - r_2\theta}{r_1-\rho(1-\theta R_1)-r_2\theta}$. If $\theta > 0$ and the bank is not solvent given a run with ρ=0 , then a $\rho \in [0,1]$ exists where (2) holds with equality and then it is given by $\rho = \dfrac{\theta R_1((1-t_s-\Delta)r_2 - R_2) + r_1(\Delta R_1 + t_s\theta R_2)}{\Delta R_1 + (1-t_s-\Delta)\theta R_1^2 - (1-t_s)\theta R_2}$.

If the regulator chooses an appropriate level of $\rho \le 1$ knowing t_s, then depositors can be sure that the bank is stable and will never want to join a run, even though they cannot observe or interpret the level of liquidity at any instant.

The intuition for why the regulation (which is a combination of a rule which can be enforced and credibly auditing) might be sufficient to foreclose a run, even when the bank's liquidity choice is unobservable, is straightforward. The LCR forces the bank to invest in more safe assets than it would voluntarily prefer to hold and the depositors know that the regulator is doing this to try to prevent runs. The bank's own self-interest continues to insure that it plans to always hold enough safe assets to cover its anticipated fundamental withdrawals and we are assuming that it can do that perfectly. Consequently, knowing that the extra liquidity cannot be avoided reduces the incentive to run.

Importantly, once the run has been prevented the liquidity still will have to remain on the bank's balance sheet. So, under these assumptions it is beneficial to force the last taxi cab to always remain at the train station.

To better understand the model works, consider the following example (which is not calibrated in any particular way). Suppose the maximum value of t is t= ½, and θ= ½, R_1= 1.1, R_2= 1.5, r_1=r_2= 1, then it is possible to solve for the ρ needed to deter the run as a function of Δ. Figure 3 shows this correspondence.

For these parameters, there are two interesting regions. First, up until the point when Δ reaches about 0.32, the optimal value of ρ is zero. Runs that are smaller than that cutoff are such that the condition in Proposition 1 holds and the bank selfishly will always hold enough safe assets so as to deter a run.

At certain point, however, the condition in Proposition 1 no longer applies and profits are no longer sufficient to prevent the run. For potential runs that are this size (or larger), ρ must be positive and it increases as the size of the potential run does, up until the point where a full run is a possibility.

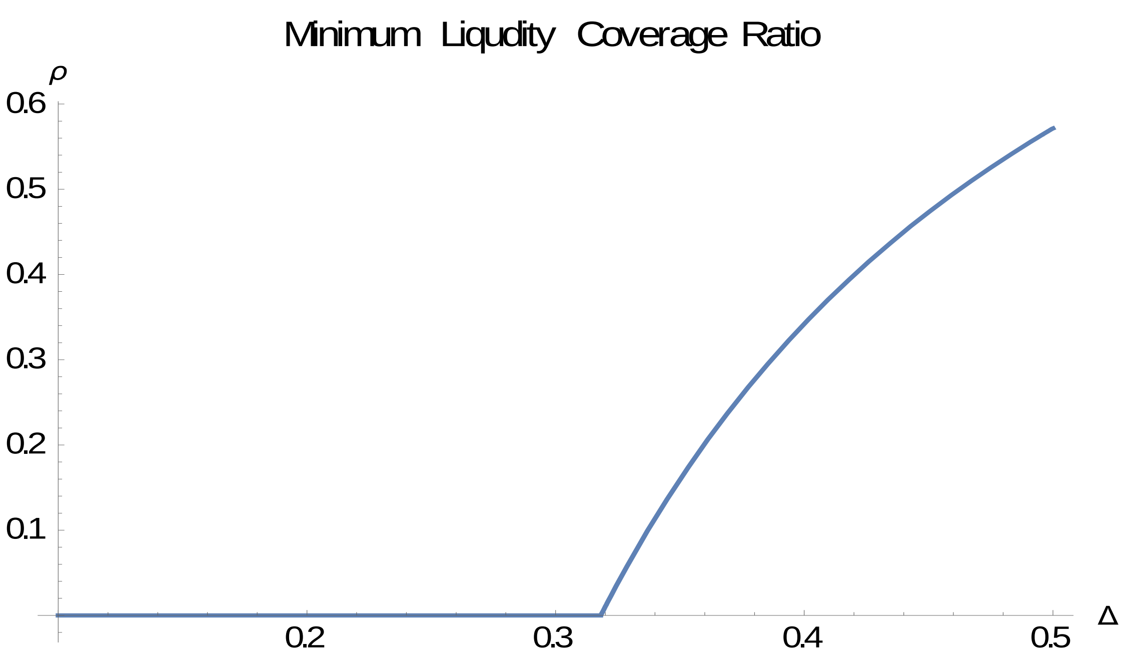


Figure 3: Liquidity Coverage Ratio as a Function of the Potential Run Risk

Proposition 4 characterizes the optimal LCR when t_s is private information to the bank.

Proposition 4: If the regulator must specify an LCR with a constant ρ knowing only the distribution of outcomes, then a value which leads the bank to be stable for all t_s must be specified. The worst case for solvency given a run is the bank with anticipated withdrawals of $\bar{t}$ (the highest possible value of t_s). A LCR ratio which makes the bank with $\bar{t}$ anticipated withdrawals just solvent in a complete run will make all types banks safe.

Proof:

A bank of type t_s, subject to an LCR of ρ will choose $\alpha_s = t_s r_1 + \rho(1-t_s)$ and given a run, the value of its equity when withdrawals exceed t_s and $f_1 = t_s + \Delta$ is:

$$E(\rho, t=t_s, f_1=t_s+\Delta) = ((1-(\frac{t_s r_1 + \rho(1-t_s)}{R_1})-\frac{t_s r_1 - \frac{t_s r_1 + \rho(1-t_s)}{R_1} R_1 + \rho(1-t_s-\Delta)}{\theta R_2})R_2 + (\rho R_1 - r_2)(1-t_s-\Delta)$$

Choose lowest ρ for a type $\hat{t}_s$, such that the value of equity given a run for that type will be exactly zero (it will just be solvent). To determine the solvency of types $t_s < \hat{t}_s$ subject to this fixed this ρ, note each will choose $\alpha_s = \frac{t_s r_1 + \rho^*(1-t_s)}{R_1}$. Differentiating $E(\rho, t=t_s, f_1=t_s+\Delta)$ with respect to t_s yields: $\frac{\partial E(\rho, t=t_s, f_1=t_s+\Delta)}{\partial t_s} = r_2 + \frac{(\rho - r_1)R_2}{R_1} - \rho R_1$. From the assumption that it is more profitable to finance illiquid assets with deposits absent a withdrawal than to finance liquid asset with one period deposits, $\frac{r_1}{R_1} > \frac{r_2}{R_2}$ we know $r_2 < \frac{R_2 r_1}{R_1}$, which implies that

$r_2 + \frac{(\rho - r_1)R_2}{R_1} - \rho R_1 < \frac{R_2 r_1}{R_1} + \frac{(\rho - r_1)R_2}{R_1} - \rho R_1 = \rho(\frac{R_2}{R_1} - R_1) < 0$. The final inequality follows from the profitability of the illiquid asset (i.e., $R_2 > R_1^2$). This implies that for all $t_s \leq \hat{t}$, banks are stable and no one would join an anticipated run. A LCR ratio ρ which makes the bank with anticipated withdrawals of $t_s = \bar{t}$ just solvent in a complete run will therefore make all types banks stable.

QED

To consider a net stable funding ratio we have to drop the assumption that initial liquidity is completely unobservable – otherwise it could not be enforced. As an alternative, suppose instead that depositors can perfectly observe α_s, but do not know how many people need to withdraw for fundamental reasons (t_s) and only know the distribution of its support (where we denote the maximum value by $\bar{t}$). Initially, we suppose that the regulator has the same information as the depositors. Suppose that the bank can continue to see t_s and that all parties know Δ.

While these assumptions allow for regulations akin to the NSFR, the regulation still must be very crude. The only certain way to assure the depositors that adequate ex-ante liquidity is being held is to insist that the bank invests in enough safe assets to cover the worst case withdrawals, $\bar{t}+\Delta$. Otherwise there will be an equilibrium where there is a run under the belief that other depositors conjecture that $t_s=\bar{t}$. Only covering this worst case will definitely remove the incentive to run, but whenever fewer fundamental withdrawals are required, the bank is left with many safe assets that must be rolled over.

This allows us to compare a NSFR which is sufficient to make stable a bank with $t_s=\bar{t}$ to a LCR which will make that same type of bank stable. Either will make stable banks of all values of t_s (and no lower values will achieve this). To illustrate the possible disadvantages of a constant NSFR, we show what happens when the worst case is $\bar{t}+\Delta=1$, and where the best possible LCR is implemented

Proposition 5: An LCR regulation can potentially support more lending than a NSFR regulation when depositors and regulators cannot condition on t_s.

Proof: The simplest way to see that this might occur is to suppose that in the worst case the run is complete, $\bar{t}+\Delta=1$. In this case, we know that $\alpha=\alpha^*=\frac{r_1-\theta R_2}{R_1-\theta R_2}$ is the optimal NSFR. But in this case, the regulator can choose ρ=ρ*, where ρ* and implement the same outcome such that $\alpha^*=\frac{\bar{t}r_1+\rho^*(1-\bar{t})}{R_1}$. Because a run on a bank with $\bar{t}+\Delta=1$ will be complete, all its liquidity can be released in a run (the LCR becomes $\rho^*(1-\bar{t}-\Delta)=0$). From Proposition 4, this LCR will make stable the other types of banks with lower $t_s<\bar{t}$, and they will be able to invest a smaller amount in liquid assets. Because they are stable, there will not be runs and they will never need to liquidate illiquid assets.

Each bank will choose $\alpha_s=\frac{t_s r_1+\rho^*(1-t_s)}{R_1}$ while a bank subject to the NSFR would still have to hold α*.

QED.

The complete run case is some sense the most favorable environment for the LCR style regulation because in the event of a full run, the requirement to maintain extra liquidity after the first date is irrelevant. In this case, the last taxicab is allowed to depart (because $\rho^*(1-t_s-\Delta)=0$). If the worst possible case involves only a partial run, there would then be a tradeoff because the incentive effects of the LCR require that some liquid assets remain on the balance sheet and the NSFR ratio does not.

More generally, when a bank is subject only to a NSFR, it gets to release all of its liquidity in the event of a run. If the regulator knows all the information as in Proposition 2, then the best NSFR is $\alpha_s^* = \max\{\alpha^{AIC}, \alpha_s^{stable}\}$.

These polar cases provide some general guidance about the relative efficacy of the two types of regulations. The LCR will work well when monitoring the bank's liquidity is difficult because the regulation forces the bank to carry more safe assets than it would prefer to. Depositors understand this and in some cases this will be enough to quell any concerns about the bank having insufficient funds to withstand a run.

The main cost of the LCR is that deterring the run requires the bank to always have some funds invested in safe assets, even if a run has occurred. Ex-post this liquidity is inefficient and everyone would be better off if more loans had been made instead. But, the incentive effects vanish if the depositors are not convinced that the liquidity will always be present. The only situation when this is not true in the case of a full run.

Conversely, the NSFR is an attractive run deterrent when the regulator is well informed about the fundamental deposit outflows, so that initial liquidity requirement can be varied. In this case, the bank can be forced to hold just enough to survive a run, but never have to hold more than is needed. Importantly, during a run a bank subject to a NSFR can always use all of its liquid assets to serve depositors. So this kind of regulation does not require the bank to liquidate any more loans than is necessary, and hence in the best case it avoids the inefficiency associated with the LCR.

Once the regulator does not have good knowledge about the fundamental needs of the depositors, using the NSFR becomes less efficient. In this case, depositors cannot generally be confident that the bank will have a portfolio that guarantee solvency in all cases. The best the regulator can, therefore, accomplish is to protect against a worst case set of withdrawals. This can dis-incentivize the run, but doing so will mean that all but the worst case the bank over-invests in safe assets. The LCR potentially is less distorting in this case.

This intuition suggests that the relative advantages of the two approaches to regulation will hinge on two considerations. One is the variability of potential fundamental withdrawal requirements. When t_s fluctuates considerably, then regulation that relies on a fixed value of α will only deter runs if the liquidity requirement is set high enough to cover the worst case outcome. When the worst case does not materialize, this will result in the banking holding surplus liquidity. Because the LCR regulation exploits the bank's knowledge about impending withdrawals and relies on its incentives to plan for these withdrawals, variability of t_s is not a problem for this kind of regulation.

The other consideration is the size of the runs that are possible. The Achilles' heel of the LCR is that is that even after a run has taken place, the bank must continue to hold liquid assets. The

NSFR avoids this (ex-post) inefficiency because all the liquid assets that the bank has can be used in the event of a run. So if runs are not complete, the inefficiency associated with the LCR will be at a disadvantage.

It strikes us that the information requirements that would favor the NSFR are relatively onerous. One of the most difficult challenges in a real time crisis is gauging the extent of a run. In that case, even if it possible to verify and certify that some liquid assets are present at any given point in time, it make be difficult to forecast whether they will be adequate to meet potential subsequent withdrawals. Hence releasing all liquidity on hand can be risky.

5. Extensions

Having characterized the properties of Basel style regulations in this model, we now discuss the implications of extending the model in three directions. First, there is no reason to restrict regulations to only look like the NSFR and the LCR, so it makes sense to expand the regulatory tools. Diamond and Kashyap (2015) provide a complete analysis of how to optimally regulate liquidity in this kind of a model and we begin with a review of those results.

Next, we dispense with the monopoly banking assumption. When there are multiple banks, then it becomes necessary to specify how banks can compete for deposits and potentially share liquidity. The potential to attract deposits to service withdrawals can change a bank's liquidity choice, which in turn may alter the effectiveness or desirability of regulation.

Lastly, we discuss several of the issues that arise if the bank faces capital regulation. Allowing savers to have a choice between investing in deposits and equity greatly complicates the model. Part of the complication comes because our model abstracts from asset risk, and many of the benefits of capital regulation arise from creating a buffer against loan losses so that any discussion of capital without asset risk is necessarily incomplete. Nonetheless, there are a couple of interesting comparisons that are possible between capital and liquidity regulation that can be made even without developing a full blown model.

Stepping away from the Basel approach, how should liquidity optimally be regulated in this kind of environment? To find the most efficient set of choices which can be implemented, we describe the results from undertaking a mechanism design analysis.[5] This will achieve the best outcome by providing incentives for the bank to reveal to a regulator the information needed to implement run free banking most efficiently. Proposition 2 already describes the full information choices.

So the more interesting case to consider is what happens where t_s is known only to the bank.

[5] The analysis here is a special case of the more general treatment in Diamond and Kashyap (2015).

A bank with private information about t_s could have an incentive to misreport t_s. The condition for efficient investment without a run remains $\alpha_s = \frac{r_1 t_s}{R_1}$. A bank will choose this under its self-interest and this level is automatically incentive compatible. From Proposition 1 we also know that if this level of liquidity is also sufficient to deter a run, then runs will be avoided regardless of whether α is observable and without any regulation.

To make it incentive compatible to report honestly t_s, the bank must be provided an incentive for reporting high levels of withdrawals which offset any increased profits that could arise from underreporting and making more loans and hence having less unused liquidity which is held after normal withdrawals occur. Diamond and Kashyap prove that under certain conditions there is a way to implement a choice of α_s^* (from Proposition 2) and which is similar to, but not exactly the same as a liquidity coverage ratio requirement.

This is possible whenever the regulator has sufficient tools to penalize the banker when actual withdrawals deviate from those that the banker reports are anticipated. Such tools could include limits on compensation or fines that would be tied to the use of accessing extra liquidity (say from a lender of last resort). If such tools exist then the regulator can require the bank to hold α_s^* and can punish any cases where the unused liquidity after the withdrawals departs from what would be needed when the bank is run-free (i.e. when actual withdrawals are $f_1 = t_s$).

In essence, the reported value of t_s allows the regulator to determine whether the realized withdrawal f_1 is or is not due to a run and release liquidity only in a run. They critical decision by the regulator is to carefully choose how much excess liquidity must be held in all circumstances to avoid penalties so as to create the right incentives for the banker to truthfully report anticipated fundamental withdrawals. Diamond and Kashyap (2015) characterize these choices under various assumptions about the nature of run risk.

The robust conclusion from this analysis is that the optimal regulation requires less unused liquidity than the simple Basel style regulations because the excess liquidity can be released if a run should occur. If there are additional constraints on what the regulator can do, which limit the ability to release this liquidity, such as a stigma to the bank if it used the liquidity (or borrowed against it), then a regulation like the LCR could be nearly optimal. If all liquidity cannot be released in a run, then the regulations have the property that as anticipated withdrawals rise, the amount of required surplus liquidity falls.

It is also possible to describe a few issues that arise when there is competition for deposits. We consider the monopoly assumption as a simple way to think about the liquidity needs for the financial system as a whole. From that perspective, the only way in aggregate for the system to service withdrawals is to hold liquidity or liquidate loans. Once we look inside the system and begin modeling competition, other margins emerge. An individual bank can try to attract deposits from another one to handle withdrawals and it could also be possible to borrow excess liquidity from a competitor.

There are many possibilities that arise in this kind of environment. We believe it is helpful to focus on three considerations (and their associated implications) that are relatively generic. First, to make any progress we need to model something about the nature of competition for deposits. Recall that in our baseline, the rate of return on deposits held from date 1 until date 2 is $\frac{r_2}{r_1}$. Unless there is perfect competition in the deposit market, a bank that tries to attract deposits from a competitor will have to a pay premium over this rate. If that premium is sufficiently large, then the option of stealing deposits, though possible, becomes irrelevant because it is dominated by the existing options of holding liquidity or liquidating loans. In this case, the results and intuition from the monopoly version of the model carry over.

Second, it would be very natural to suppose that trying to attract outside deposits during a run would be particularly expensive. So it is plausible to think that competition would be more relevant for judging whether banks might to try pay off normal withdrawals by stealing deposits from competitors. In this case, a force similar to the kind of free-riding analyzed by Farhi, Golosov and Tsyvinki (2009) can emerge. If each bank makes its liquidity choice under the assumption that some withdrawals will met by attracting deposits from competitors, and the ex-ante cost of holding internal liquidity is higher than the cost of attracting external deposits, the system as a whole will be short of liquidity and it will be necessary to mandate that they hold extra to overcome this problem. In this case, one might need regulation even if information is complete. This would imply a particular level of liquidity to require, and the general approach we propose in the basic model can be used to provide incentives to hold this level.

Third, adding an interbank market to the analysis raises a number of thorny issues. For instance, if depositors can run on a particular bank and redeposit the funds in another bank, the run incentives will depend on whether the bank that receives the inflow can relend those funds immediately or not. Fully analyzing this and the related set of possibilities would be interesting but is beyond the scope of this paper.

Finally, it is worth noting several observations between interactions capital and liquidity regulation. In our baseline model, there is no credit risk associated with loans, so the usual arguments for capital requirements do not hold. Generically, however, the incentive to run is still related to depositors' assessments about the solvency of the bank so the presence of equity could still matter.

The role that capital would play in deterring a run is subtle. On the one hand, if the bank issued capital (non-demandable liabilities) and invested the proceeds in loans, this can leave the bank more solvent when a fixed number of deposits are withdrawn, moving the bank to a situation where it is solvent during a (potential) run of fixed size. This is due to the liquidation value of the additional loans made. On the other hand, added equity would be irrelevant if a (potential) run of given size given is still going to make the bank insolvent. In our framework, this is easiest to see if the liquidation value of the loans (θ) is zero. In that case, the future value of the assets that would otherwise be the basis of the equity value would be of no value in a run. So the liquidity requirements needed to deliver stability would be unchanged and capital requirements would be completely ineffective.

Once assets become risky, the analysis becomes much more complicated. In this kind of environment, depositors will make withdrawals based both on their fundamental liquidity needs and based on beliefs about the future value a bank's assets. In addition, if the bank can fail simply because loans turn out to default, the bank's choice between loans and liquid assets can also be distorted if there is limited liability; banks in this kind of an environment can in some situations have an incentive to shift risk on depositors.

A full analysis of this kind of model is beyond the scope of our survey, especially because there are so many additional assumptions that are needed to maintain tractability. However, Kashyap, Tsomocos and Vardoulakis (2015) have solved one particular version of this kind of model and their analysis does deliver one apparently general result about the interactions between capital and liquidity requirements in deterring runs that is worth mentioning.

They show that there is a fundamental asymmetry in the way that liquidity and capital regulations work in preventing runs. Capital requirements essentially work on the liability-side of a bank's balance sheet without directly constraining the bank's asset choices. Hence, when a bank is forced to have higher equity, it can on the margin reduce its reliance on deposit financing. The need for deposits means that the bank can marginally reduce liquid asset holdings too. This frees up the bank to make marginally more loans. While this marginal adjustment is not enough to raise the overall risk of a run, it does suggest that the bank's assets will become less liquid.[6]

Conversely, liquidity regulation, either in the form of an LCR or NSFR, work very differently. The LCR, as we have seen, directly forces the bank to substitute from illiquid assets towards liquid assets. So the run deterrence automatically is accompanied by having less liquidity risk. The NSFR forces the bank to finance illiquid assets with long-term liabilities. Therefore, if the bank wants to take on additional illiquid assets, it cannot fund them with runnable deposits. Instead, short-term deposits will shrink along with liquid assets.

[6] This is not arising because of a Modigliani Miller type fallacy whereby depositors fail to recognize that the bank's deposits are safer. Instead, this happens because the liquid assets are held only to deter runs and when capital requirements make them less likely the bank cut back on liquid assets.

Kashyap, Tsomocos and Vardoulakis (2015) describe many other complicated ways in which capital regulations and liquidity regulations can complement or substitute for each other. The asymmetry in how they marginally influence asset illiquidity is robust.

6. Conclusion

Our analysis provides some novel insights that can inform subsequent discussions of how to design liquidity regulation. Our starting point is the recognition that for a forward looking intermediary, access to future deposit or other funding influences the desired ex-ante, profit-maximizing choice of how much liquidity to hold. Absent any regulation, the bank will voluntarily opt to hold more liquidity when higher exogenous deposit reductions are anticipated. Hence it is helpful to understand whether, and when, this incentive alone will lead to banking stability when it is not directly a goal of the bank.

In the kind of model we have explored stability is not guaranteed because depositors may have doubts about whether the bank is sufficiently safe to withstand a panic. The lack of confidence that creates this problem can arise for various reasons. Banks are opaque and even for sophisticated counterparties assessing their balance sheet can be challenging. Information about the balance sheet is rarely available contemporaneously, so some forecasting (about the bank's condition and the decisions of other depositors) is inevitable. This will cause problems when the bank's incentives are not aligned with enhancing stability.

The imperfect information also creates a problem for the bank. Cutting back on lending and holding additional liquidity is not fully rewarded by the uninformed investors, so the bank's private incentive to become super-safe is limited. Regulation that mandates additional liquidity can potentially circumvent this problem.

Analogs to both of the two regulations contemplated as part of the Basel process, the net stable funding ratio and liquidity coverage ratio, are among the various types of regulations that we explore. These can arise as approximations of a general type of regulation that is optimally designed to resolve the information friction. All of the ones we consider are designed to eliminate runs.

The generic form of the optimal regulation specifies that the bank must hold a level of liquid assets that is tied to anticipated withdrawals, but which often will exceed the level that it would choose on its own. If the regulator is well informed about these withdrawals (and the risk of a run), then there are many equivalent ways to guarantee that adequate liquidity choices by the bank are made. In particular, stability can be achieved either by having the bank hold the correct amount of liquid assets up-front as with a NSFR, or by imposing restrictions that require liquidity be available even after withdrawals are underway (as with a LCR). Using combinations of these kinds of policies will work too.

To achieve the efficient outcome which could prevail with full information available to all, the regulator must be able to induce the bank to disclose everything it knows about the deposit risk that it faces (or have access to that information from some other way). With the ability to impose taxes on bank compensation, the regulator could elicit this from the bank. A lender of last resort policy that penalizes liquidity regulation violations by limiting compensation, but allows the bank to borrow can also implement this arrangement.

We also considered a case where the regulator has fewer tools and can only make the regulation depend on observed withdrawals. In that case, the best regulation shares some properties of a LCR type requirement. Some banks end up being forced to hold excess liquidity and by carefully choosing the required level, the regulator can assuage depositors concerns so that runs do not occur.

One generic property of all of the optimally designed regulations is that they often involve requiring the bank to hold liquidity that go unused. So even in the best possible case, the last taxi cab is often required to remain at the station. Fundamentally, this occurs because the unused liquidity is needed to deter the run.

There are two separate forces that lead to this result. First, a prudent provision that that forecloses a run necessarily requires that the bank has enough liquidity to be able to service depositors if they did run. This might be possible through liquidating loans. But liquidations are highly inefficient so this typically this will not be sufficient and the bank needs to have some liquid assets which could be deployed if needed. By mandating the “dry powder”, the regulator preserves solvency in a run and thus removes the depositors’ incentive to run.

The second consideration is that a regulator cannot count on being able to distinguish a run from a situation where fundamental withdrawal needs are simply high. The goal in preventing runs is to do so without mandating more dry powder than is needed. Unfortunately, even when exceptionally high levels of withdrawals are anticipated, some dry powder is needed.

Appendix: Mixed Strategies by Depositors

Any uncertainty about aggregate withdrawals given t_s would need to come from randomness in withdrawals by depositors who would otherwise prefer to withdraw at date 1 and who observe the sunspot. Individual randomness would require a mixed strategy on their part, but one where a belief that all of Δ of them run and make the bank insolvent is not self-fulfilling. If the bank would not fail in a run of fewer than Δ depositors, then because the return from running (and earning less than $\frac{r_2}{r_1}$ from money put under the mattress) is less than remaining in the bank, only t_s depositors will withdraw, and the bank must expect $f_1 = t_s$ as in the main text. Alternatively, if the bank were to fail in a positive probability run, some of the very risk averse depositors receive zero (which gives them a much lower ex-ante payoff than safe deposits). Hence, this would not be an equilibrium which anyone would want to implement.

References

Admati, A R, P M DeMarzo, M F Hellwig and P Pfleiderer (2010), "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is not Expensive," Rock Center for Corporate Governance Working Paper No. 86, Stanford University.

Allen, F, 2014, "How Should Bank Liquidity be Regulated?", mimeo.

Allen, F and D Gale (1997), "Financial Markets, Intermediaries, and Intertemporal Smoothing", *Journal of Political Economy* 105(3), 523–546.

Basel Committee on Bank Supervision (2013), "Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools," Bank for International Settlements.

Basel Committee on Bank Supervision (2014), "Basel III: the net stable funding ratio", Bank for International Settlements.

Baron, David P and Roger B. Myerson (1982), "Regulating a Monopolist with Unknown Costs," *Econometrica* 50, 911-930.

Benston, G J and C W Smith (1976), "A Transactions Cost Approach to the Theory of Financial Intermediation", *Journal of Finance* 31(2), 215–231.

Bhattacharya, S. and D. Gale (1987), Preference shocks, liquidity and central bank policy, in: W.A. Barnett and K.J. Singleton, eds., New approaches to monetary economics (Cambridge University Press, Cambridge).

Black, F (1975), "Bank funds management in an efficient market," Journal of Financial Economics, Volume 2, Issue 4: 323–339.

Calomiris, C W, Heider F, and M Hoerova, 2014, "A Theory of Bank Liquidity Requirements", mimeo, July 2014.

Calomiris, C W and C M Kahn (1991), "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *The American Economic Review* 81(3), 497–513.

Diamond, D W and P H Dybvig (1983), "Bank Runs, Deposit Insurance and Liquidity", *Journal of Political Economy* 91(3), 401–419.

Diamond, D W (1984), "Financial Intermediation and Delegated Monitoring", *Review of Economic Studies* 51(3), 393–414.

Diamond, DW, (1997), "Liquidity, Banks and Markets," *Journal of Political Economy: 105, 928-956.*

Diamond, DW and A K Kashyap, 2015. "Optimal Regulation of Bank Liquidity", work in progress.

Diamond, D W and R G Raghuram, 2001, "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking", *Journal of Political Economy* 109(2), 287–327.

Ennis, H and T Keister, 2006, "Bank Runs and Investment Decisions Revisited" *Journal of Monetary Economics* 53(2), pp. 217-232.

Farhi, E, Golosov, M and A Tsyvinski, 2009, A Theory of Liquidity and Regulation of Financial Intermediation, *Review of Economic Studies* 76, no. 3: 973-992

Goodhart, C A E, 2008, "Liquidity Risk Management", Banque de France Financial Stability Review, pp. 39-44.

Jacklin C J, 1987, "Demand deposits, trading restrictions, and risk sharing," in Edward C. Prescott and Neil Wallace, editors, *Contractual arrangements for intertemporal trade*, University of Minnesota Press, pp. 26-47.

Kashyap, A K, Tsomocos D P, and A P Vardoulakis, 2014, "Principles for Macroprudential Regulation", *Banque de France Financial Stability Review*, 18, pp. 173-181.

Kashyap, A K, Tsomocos D P, and A P Vardoulakis, 2015, "How does macroprudential regulation change bank credit supply?" revision of National Bureau of Economic Research working paper 20165. .

Modigliani, F and M H Miller, 1958, "The Cost of Capital, Corporate Finance and the Theory of Investment", *American Economic Review*, 48(3), 261-97.

Vives X, 2014, "Strategic Complementarity, Fragility, and Regulation", *Review of Financial Studies*, 2014, 27, 12, 3547-3592.