

Models of Discount Window Lending: A Review

Huberto M. Ennis

*“I have made a systematic analysis of discounting in my little book, *A Program for Monetary Stability*. Any answers that I might give to your questions now would be more offhand and less satisfactory than that statement.”*

—Milton Friedman

In July 1968, the Federal Reserve released a report titled “Reappraisal of the Federal Reserve Discount Mechanism.” The report contained the conclusions from a series of studies conducted by the Federal Reserve System during a period of three years. One of these studies was a compilation of the responses by a large group of academics to a set of eight questions about the subject. Milton Friedman’s (complete) answer to the questions appears in the quote above. I interpret Friedman as saying that if one wants to assess the extent of knowledge on how a discount window institution should be structured and operated, the best way to proceed is to study the existing literature on the subject. The objective of this essay is to summarize some of the ideas that come from doing just that.

The academic literature on the lender of last resort (LLR) policy is extensive. There are (at least) two kinds of papers in that body of work: (1) papers that explain and formalize the barriers to perfect market functioning that the discount window is trying to address and then discuss how to best run a LLR policy given that situation; and (2) papers that start from the premise that there is a LLR and study in more detail particular aspects of its organization – such as, for example,

■ I would like to thank Beth Klee, Jeff Lacker, Allen Sirolly, Nico Trachter, Alex Wolman, John Weinberg, and Steve Williamson for comments. The views expressed in this article are my own and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

whether the LLR should conduct supervisory activities, and how, or whether those should be outsourced to a separate agency. The focus here will be mainly on the first set of papers.

The government can conduct LLR activities in different ways. One classic approach is to delegate such authority to the central bank. But this is not the only way: the government can, in principle, put in place lending programs administered by the fiscal authority. Furthermore, not all of the central bank's LLR interventions have to be channeled through the regular discount window. Other specially designed lending facilities could be put in place instead. At the level that the subject is discussed in the particular literature I will be summarizing here, the distinction between all of these different forms of lending is often not very sharp. For this reason, and for the sake of concreteness, I will often refer to discount window lending as the generic LLR policy aimed at intervening in financial markets.

In general, to evaluate the optimality of a given discount window policy, one needs to determine the problem that such policy is trying to solve. Discount window lending may play a role in (at least) two different situations: (1) when only one or a few firms need to borrow short-term funding for idiosyncratic reasons; (2) when many firms in the system need to borrow in a situation that could be considered an economy-wide (systemic) event. In the first case, in principle, other firms have funds available that could satisfy the demands of the few borrowing firms at interest rates close to the prevailing (risk-adjusted) rates. The second case is a situation where only a significant change in interest rates would allow the system to equilibrate itself without any intervention, and a crisis may ensue. When only some firms are looking to borrow, the discount window may play a (meaningful) role only if there are some impediments to the functioning of markets limiting the ability of those firms to obtain funding from other firms. When the economy is experiencing a systemic event, the discount window may be one channel through which the central bank can adjust the aggregate quantity of liquidity in the market to avoid undesirable spikes in interest rates – with open market purchases of assets by the central bank, in exchange for bank reserves, being a natural alternative to that.

As is evident from this discussion, taking a general equilibrium approach is essential to evaluate the potential role of a discount window. When a set of firms (small or large) have borrowing needs, market forces will, in principle, produce the necessary price-and-quantity adjustments to accommodate those needs. The question then becomes: Are those prices and quantities desirable? Or, in other words: Is the allocation of credit in the economy efficient? To answer this question, one needs first to understand how the system adjusts in

general equilibrium. Additionally, one needs to determine the ideal (efficient) allocation to use as a benchmark in evaluating the equilibrium allocation. So, one needs a full description of the economic system (i.e., a general equilibrium model) and a notion of efficiency applicable to the set of feasible allocations of resources in that system.

Holmstrom and Tirole (1996) subscribe to this approach when they say: “Modeling aggregate liquidity shortages and analyzing liquidity premia require a general equilibrium model in which no spurious demand for liquidity is introduced through ad hoc restrictions on asset trades.” In this quote, the ideas of “liquidity” and “liquidity shortages” are crucial elements. However, these are terms that are often used in different contexts, with different meanings. To be able to evaluate alternative policies we need to understand the specific phenomena that lie behind those terms.

One possibility is to broadly interpret the concept of “demand for liquidity” as the need of one firm, or a set of them, to borrow short-term funding. Relatedly, in some cases liquidity refers to the idea of “cash in the market” and the fact that under certain conditions there may be only a limited amount of nominal means of payments available to execute an appropriate amount of trade. These concepts intend to capture complex situations also best understood in the context of a well-specified explicit model. The intention in this review is to discuss in some detail several existing attempts in the literature to *formally* analyze the role of a LLR using such models.

The central bank discount window is generally considered an integral tool in monetary policy implementation frameworks (see, for example, Ennis and Keister [2008]). The idea there is that the interest rate charged at the discount window represents the costs of being short of reserves at the end of the trading day and hence determines the willingness of banks to pay for reserves during the day. Understanding the behavior of the daily demand for bank reserves is crucial to the implementation of monetary policy when the central bank follows the common practice of intervening in the market for reserves in order to target a level for some relevant interest rates. Furthermore, the discount window rate will also provide a virtual ceiling for the interest rate on interbank loans of reserves in such a situation. This more specific role of the discount window is not the main focus of this article (see Ennis and Weinberg [2016] for a brief discussion of this topic).

Separating the LLR function of the discount window from its role in monetary policy implementation has some theoretical backing. In a now classic paper, Goodfriend and King (1988) argued that unless there are significant barriers to the functioning of financial markets, central-bank open market operations – that is, buying and selling assets

in exchange for reserves in the open market – are sufficient to conduct monetary policy effectively.¹ In principle, this separation of functions can be analytically helpful. However, in this article we will review several possible obstacles to the perfect functioning of markets that will make the Goodfriend and King argument much less clear-cut (Flannery 1996). As a result, monetary policy issues will arise in the discussion even though these are not the main focus of the article.

The rest of the paper is divided in two sections. In Section 1, I review several general equilibrium models that have been used to address the question of how to conduct appropriate discount window policy. I try to follow (approximately) a chronological order in the presentation, and I make an effort in the discussion to identify features shared by some of the models. In Section 2, I summarize the main common themes that come out from reviewing the literature, and I provide some concluding remarks.

1. AN OVERVIEW OF THE MODELS

One of the first and most influential contributions to our understanding of the theoretical determinants of aggregate liquidity conditions is the model proposed by Diamond and Dybvig (1983). While the original model is not designed specifically to discuss the role of the discount window, many features of the Diamond-Dybvig framework have been used later in the literature to address the questions that interest us here. For this reason, I start this section with a brief discussion of this seminal contribution.

Diamond and Dybvig (1983) consider an abstract general equilibrium economy populated by a large number of agents facing idiosyncratic preference shocks that drive them to want to consume more or less, earlier or later. These shocks are private information. For simplicity, call the agents who want to consume early impatient. The rest of the agents are patient.

There is also available a productive technology that delivers positive returns but requires time to mature. The optimal arrangement is to provide insurance to consumers against their idiosyncratic preference shock, but since providing “liquidity” insurance is costly – as it reduces investment in the productive technology – insurance is only partial:

¹ One way to interpret Goodfriend and King’s (1988) discussion is as recasting some of the most compelling arguments in Friedman’s (1960) book, using a more modern perspective. Friedman, like Goodfriend and King, favored open market operations as the main monetary policy tool and went further in saying that, in the U.S., “rediscounting should be eliminated.”

impatient agents are able to consume more than in autarky but less than patient agents.

Diamond and Dybvig show that in the absence of aggregate uncertainty about liquidity needs, there is an optimal arrangement where agents pool deposits in a bank-like institution and receive payments according to their preference shock in an incentive compatible way – that is, given those payments, agents do not want to pretend they have experienced a different shock than the one they actually received. In fact, that arrangement produces the first-best allocation. They then move on to study the case of aggregate uncertainty and show that, under the assumption that payouts are executed sequentially in a first come, first serve fashion (sequential service), the first-best allocation is no longer implementable.

Diamond and Dybvig also discuss the possibility of self-fulfilling runs in their model. An extensive literature has developed that refines the insights about financial fragility that come out from the model. A detailed discussion of that literature is beyond the scope of this article (see, for example, Ennis and Keister [2010] for a survey). As it turns out, the original Diamond-Dybvig framework does not produce clear-cut prescriptions about the value of having in place a discount window facility. The model, however, has been extended and modified in various ways to address such issues. Some of those contributions are discussed below.

Liquidity Risk, Moral Hazard, and the Interbank Market

Bhattacharya and Gale (1987) use the Diamond and Dybvig (1983) environment as a starting point for their analysis. They consider the case with no aggregate uncertainty – that is, there is a continuum of depositors and the law of large numbers applies, so that the proportion of impatient depositors in the economy is equal to the known probability of each individual depositor being impatient. Also, in their setup, there are two technologies: a liquid and an illiquid, more productive technology. The liquid technology can be liquidated early or late but there is no extra return from waiting. The illiquid technology, instead, produces higher returns when waiting but cannot be liquidated early.²

² Another way to think about this is that, for the illiquid technology, the liquidation costs are so high that there is effectively no benefit from trying to access the invested resources early. As it turns out, this feature of the illiquid technology makes the traditional Diamond-Dybvig self-fulfilling bank runs not possible in the environment.

Bhattacharya and Gale restrict (exogenously) the way intermediation can be organized in the economy. They assume that there is a continuum of intermediaries and that those intermediaries are divided into two groups (two types): in one group, intermediaries have a low proportion of impatient depositors and, in the other group, they have a high proportion.

This (industry) structure with multiple intermediaries is not explicitly justified in the model. In fact, centralizing the process of intermediation appears to be a simple solution to some of the problems that arise. Furthermore, proposing policies to address those problems while abstracting from the un-modeled reasons that justify the presence of multiple private intermediaries can be regarded as problematic: it is likely that those un-modeled reasons have important implications for the design of the optimal intervention policy. Ignoring such possibilities constitutes an obvious challenge to the robustness of the results. Keeping in mind this qualification, let us proceed to describe the main insights from this influential paper.

The timing of events in the model is important. Initially, at the time of the investment decision, intermediaries do not know which type they will be. Its type gets revealed to the intermediary at the time when impatient depositors wish to consume (no sequential service), but this information remains private (is not observable by the other intermediaries in the economy). Furthermore, the decision of how much liquid and illiquid investment to undertake is also private information of the intermediary. There is, hence, a combination of two private information problems: a hidden state problem associated with the possibility that an intermediary could misrepresent its proportion of impatient depositors after types are realized, and a hidden action (moral hazard) problem associated with the ability of intermediaries to choose the level of liquidity in their portfolios.³

Taking as given the assumed industrial organization of the intermediation industry – with multiple intermediaries – and the information structure imposed on the model, Bhattacharya and Gale solve a planning problem subject to incentive compatibility constraints. The planner receives reports from each intermediary about their type and designs payments to depositors so that each intermediary has

³ Allen, Carletti, and Gale (2009) study a related model where banks face idiosyncratic liquidity shocks and can trade their long-term assets for liquidity in an interbank market. The authors also consider the possibility of aggregate liquidity shocks. Shocks and portfolios are observable, but deposit contracts are assumed incomplete (payoff cannot be contingent on the liquidity shocks). The model supports the Goodfriend and King (1988) insights in the sense that central-bank open market operations are the appropriate policy in such an environment.

incentives to: (1) truthfully reveal its type and (2) choose the recommended (optimal) amount of liquidity to hold in its portfolio. The specification of those payments to depositors (given type and liquidity) implicitly defines a transfer scheme across intermediaries (which are later interpreted in the decentralization as the result of a set of transactions in the interbank market).

Once intermediaries choose the level of investment and, hence, liquidity – all choose the same, since they are identical *ex ante* – the uncertainty about the proportion of impatient depositors constitutes a (“liquidity”) risk to each intermediary and its depositors. Since there is no aggregate uncertainty, this risk could be pooled across intermediaries and, at least partially, could be insured. The incentive constraints, however, put a limit to the amount of insurance that can be effectively provided. An intermediary, when insured, has incentives to invest more in the high return illiquid technology and rely on the insurance provider (the planner) to make payments to impatient depositors. Also, if too much insurance is provided, then the intermediary has an incentive to misrepresent its type and claim a proportion of impatient depositors higher than the true one. The planner deals with this trade-off between insurance and incentives and strikes the optimal balance.⁴

Since the total proportion of impatient depositors in the economy is known and fixed *ex ante*, in terms of economy-wide resources, providing insurance to intermediaries is not costly. It only involves redistributing resources from one set of intermediaries to the other. On the other hand, just as in the Diamond-Dybvig setup, it is generally costly to provide insurance to *individuals* because making payments to impatient depositors requires investing less in the more productive (but illiquid) technology. For this reason, consumption of the impatient depositors of a given intermediary is always lower than the consumption of its patient depositors.

With only two types, the limits on the ability of the planner to fully insure intermediaries come from the nonobservability of liquidity decisions. That is, the moral hazard problem is the crucial friction in the model. In fact, if investment in liquidity were observable, then full insurance would be possible even when intermediaries’ types remain private information.

The basic trade-offs determining the constrained-optimal allocation are the following. Liquidity is useful to pay to impatient depositors. An intermediary can underinvest in liquidity and then try to reduce the

⁴ Ratnovski (2009) considers a model where banks may have incentives to underinvest in liquidity *ex ante* to exploit the central bank’s tendency to serve as a LLR when trying to contain the damage from a systemic banking crisis.

associated cost of choosing low liquidity by falsely claiming a high proportion of impatient depositors. To control such behavior on the part of the intermediaries, it is optimal to limit the “net transfer” to those intermediaries that report a high proportion of impatient depositors. As a result, depositors in those banks (“illiquid” banks) consume less than depositors in banks with a low proportion of impatient depositors (“liquid” banks). In other words, only partial liquidity insurance is optimal, and both patient and impatient depositors in the “illiquid” banks (that is, not just the impatient depositors), by consuming less, share the cost associated with providing appropriate incentives to banks.⁵

In principle, one way to decentralize desirable allocations in this environment would be to have intermediaries borrowing and lending in an interbank market to “insure” their liquidity risk. However, an interbank market for loans operating under laissez-faire conditions would not implement the optimum – the solution to the planner’s problem. Instead, careful inspection of the constrained-optimal allocation suggests that “illiquid” banks should get a loan of a given (limited) size at a “subsidized” rate. The authors argue that a discount window could play a role in providing these loans. Unfortunately, there is no detailed discussion of the way the system would work, short of completely substituting for the private interbank market and having the discount window provide all loans.

Based on this logic, the results of Bhattacharya and Gale (1987) could be used as a justification for subsidized discount window lending. Under the optimal allocation, intermediaries would associate an interest rate R_b to the intertemporal trade-off between paying patient and impatient depositors. Implementing the constrained-optimal allocation would require providing loans to intermediaries with a high proportion of impatient depositors at a lower interest rate R_o .⁶ In this sense, the rate R_o would involve a subsidy relative to R_b .

The origin of this apparent subsidy is the following. At the time that impatient depositors are being paid, the economy is just reallocating funds from those intermediaries that need funds to those that do not. From an economy-wide perspective, these transfers could be made one-for-one (as they do not create an extra resource cost for the economy). That suggests that R_o could be equal to unity. The reason

⁵ Also working with a framework inspired in the Diamond and Dybvig model, Keister (2016) studies bailouts and financial fragility. In Keister’s model, a similar moral hazard effect is present: when intermediaries expect a bailout, they become less liquid. Keister argues that the optimal way to handle this problem is to introduce a tax on short-term liabilities.

⁶ Actually, depending on parameters, R_o could be greater than R_b in the solution to the planning problem. However, if the amount of liquidity redistribution necessary to implement the optimum is not too large, generally R_o will be lower than R_b .

that R_o is greater than unity is that, when R_o is relatively low, intermediaries do not have sufficient incentives to invest in liquidity *ex ante*. However, the value of R_o that accommodates the incentive problem may be lower than the intertemporal rate of transformation R_b (which directly depends on the return of the productive technology).

One way to summarize the main takeaway from this discussion is that if the necessary reallocation of liquidity across financial institutions does not involve an intertemporal reallocation of resources, then the interest rate on the implied loans does not have to reflect any intertemporal tradeoffs. As a result, the optimal interest rate on those loans can be different from the main intertemporal price prevalent in the economy and, for that reason, may appear to involve a subsidy.

Before closing the discussion, it is interesting to call attention to the remarks that Holmstrom and Tirole (1998, p. 35) make about Bhattacharya and Gale's article: "The paper characterizes the socially optimal mechanism for sharing risk across banks, noting that this mechanism cannot be implemented through an interbank lending market. Whether a market in bank shares or some other institution could implement the social optimum is left open. The authors suggest that a central bank might be the right institution for carrying out interbank risk sharing." The quote draws attention to the relative tenuousness of the proposed role for a discount window. It seems unclear, based on the analysis in the paper, whether other feasible trading activity could render discount window lending superfluous in such an environment.

Spatial Separation, Nominal Debt, and Illiquidity

Freeman (1996) studies a general equilibrium model where spatial separation among agents seeking to exploit gains from trade creates the need for the use of a specific type of debt instrument as a means of payment on some transactions. The debt instrument useful for transactions is one that promises to pay fiat money in the future. The particular way in which the meetings between agents occur makes promises of payment in fiat money the only meaningful ones. For this reason, in the equilibrium of the model, if fiat money has no value, those debt-financed transactions become not viable.

A basic description of the patterns of trade is the following. Initially, some agents (called debtors) meet with some other agents (called creditors). Debtors would like to acquire some goods from creditors but cannot engage in barter because they have goods that creditors do not like. At that point in the timeline of events, debtors also do not have money. They will, however, be able to sell their goods to another set

of agents later in time in exchange for money. In anticipation of this future transaction, debtors buy from creditors paying with an IOU that is a promise to pay cash in the future. Debt repayment happens at a central location where all agents who are present at the location can transact with each other.⁷

The problem is that arrivals and departures to and from the central location are not coordinated. In particular, some of the debtors arrive to the location only after some of the creditors have left, and who arrives and leaves when is unknown at the time that the IOUs are being issued. As a result, some debtors arrive to the central location after the creditors whom they need to make a payment to have left. This lack of coordination creates a motive for buying and selling debt claims – in other words, private debt “circulates” in the equilibrium of Freeman’s model.

Creditors leaving early and in possession of an IOU from a debtor who has not yet arrived will want to trade the IOU with a creditor who is not yet leaving. The issue then becomes whether there is enough *cash in the market* to buy at par-value all the debt claims held by those creditors leaving early and still holding unredeemed IOUs. If the proportion of creditors leaving the central location early is high relative to the proportion of debtors arriving early, then cash in the market will be scarce and debt will sell at a discount. Freeman calls this equilibrium a liquidity-constrained equilibrium.⁸

Since creditors expect that with some probability they will have to sell their holdings of IOUs at a discount, they are less willing to initially exchange goods for IOUs, and this distorts the real allocation of goods in the economy. In other words, illiquidity matters for welfare.

Note that the reason for the illiquidity is that some agents who have money that could be used to purchase debt are not present in the market at the time when the sales need to take place. In that sense,

⁷ Money has value in the economy because there are two-period-lived overlapping generations of agents and old creditors want to consume the goods possessed by young debtors. The only way that old creditors can purchase goods from young debtors is by using money. Promises to pay in the future do not work because old creditors will exit the economy in the following period. Debt arrangements are only feasible among the young agents: young creditors make loans to young debtors, and those loans are repaid the following period when the corresponding counterparties are old. Old debtors repay in cash to old creditors, who then use it to buy goods from the new young debtors – who keep the cash to use it next period in repaying their debts, which they contracted in the same period but before being able to transact with old creditors.

⁸ There are other formal treatments in the literature of this type of cash-in-the-market effect. For example, Acharya and Yorulmazer (2008) discuss how cash-in-the-market pricing may interact with bank failures to produce high asset-price discounts and an inefficient allocation of resources. These authors argue that providing liquidity to the appropriate potential buyers of assets is a better policy than trying to contain asset sales with a bailout to failing banks.

there is market segmentation in the model. Alternatively, one could think that this is a version of the “slow moving capital” idea discussed, for example, by Duffie (2010). Essentially, all of these are possible justifications for cash-in-the-market pricing.

The central bank could try to address this “cash shortage” by issuing extra cash at the time that asset sales need to happen and channeling it somehow to the market. To that end, Freeman proposes a discount window policy that would solve the illiquidity problem with the additional benefit of not creating inflation – as money is injected in and out within the period. Basically, creditors can rediscount debt with the central bank and get cash, which they then can use to buy more debt from the market and again rediscount that debt at the window. It is easy to see that this will completely undo the illiquidity problem. All the rediscounting of debt is done by creditors leaving the central location late. Once the debtors arriving late finally arrive, they repay the debt to the creditors still at the central location and with that cash the creditors at the central location pay back the discount window loans. Hence, the quantity of money relevant for the determination of inflation does not change with the rediscounting, and the central bank can deal with the liquidity problem without creating inflation.

Obviously, a discount window is just one possible arrangement to deal with the illiquidity problem in this situation. In principle, the central bank could just buy the debt in the market, wait for debtors to arrive at the central location, when they would pay back the debt directly to the central bank.

In the last section of the paper (before the conclusion), Freeman extends the model to introduce default risk. If creditors know the (average) default risk of debt but the central bank does not, then the central bank can rely on creditors to deal with the default risk. Creditors are assumed to be able to fully diversify across debtors, so they only care about the average default risk. The central bank makes loans to creditors who are responsible for repaying the loans in full to the central bank. Then, creditors would be willing to buy debt in the market as long as the price of debt reflects the default risk of that debt (and not any liquidity premium). This creates a distinction between discount window lending (to trustworthy parties) and outright purchase of debt in the market.

In Freeman’s model money plays two roles. It is used to trade goods between old creditors and young debtors at the end of the period, and it is used for the clearing and settlement of debt within the period. The real value of money is determined by the need to pay for goods, but the resulting amount of real balances may not be enough to permit the clearing of debt at par within a given period. Similar sources of

“illiquidity” appear in other setups in the literature, such as in the model by Holmstrom and Tirole (1998), which we discuss in the next subsection. In Freeman’s model, the way to deal with this intra-period illiquidity is to have an intra-period elastic supply of currency provided by the central bank through the discount window. The central bank’s lending facility adds enough flexibility to the quantity of money at each point in time, so as to allow money to appropriately play all its roles in the economy.

Freeman (1999) extends the model to include *aggregate* default shocks. The idea is that in some states of the world only a proportion of those debtors who have issued an IOU are able to travel to the central location, where repayment of debt happens. Initially, when debtors are trading with creditors, they do not know whether all or only some debtors will be able to travel to the central location later in the period, nor do they know which debtors will travel and which ones will not. The rest of the model is basically the same as in Freeman (1996).

In the model with aggregate default risk, how the central bank structures its liquidity provision can matter for welfare. In particular, open market operations result in price level fluctuations that can produce better risk sharing than the fluctuations that arise when the central bank intervenes through the discount window. The reason why price level fluctuations happen is that the central bank, due to the default shock, is not able to recover at the end of the period all the liquidity injected intraperiod to ameliorate the impact of cash-in-the-market pricing on transactions.

Note that by buying assets outright through open market operations, the central bank assumes the default risk directly and transfers it more evenly to all agents holding money through the resulting fluctuations in the price level. Instead, when the central bank provides loans to creditors in order for them to buy IOUs of debtors, in many situations, those creditors retain most of the default risk. The more uneven distribution of risk under discount window interventions can be detrimental to welfare.

Freeman’s analysis abstracts from what is potentially an important problem associated with liquidity provision by a central bank: moral hazard. Two follow-up papers investigate the issue in Freeman-like frameworks: Martin (2006) and Mills (2006). Martin (2006) considers a model where agents can choose between a safe and a risky investment. Provision of central bank credit can distort the investment decision of agents, and Martin shows that a collateral requirement on central-bank loans can help to minimize the ensuing moral hazard problem.

In Martin's model, whenever agents have the resources to repay the loans, they do so. In other words, there is no strategic default. Mills (2006), on the other hand, allows agents to control their loan-repayment decisions. That is, agents who take a loan from the central bank will repay only if they have the appropriate incentives to do so. At the same time, Mills allows the central bank to engage in costly confiscation of property if an agent does not repay a loan. Because enforcement is costly, moral hazard can compromise the ability of the central bank to fully resolve the liquidity problems that may arise. Mills also considers an opportunity cost of providing collateral, absent in Martin's model, and shows that costly collateral also can limit the ability of the central bank to costlessly address liquidity issues.

Spatial Relocation, Lending, and the Limits to Diversification

Williamson (1998) (following Champ, Smith, and Williamson [1996]) writes a general equilibrium model that combines some features of the Diamond and Dybvig model with some features of the Freeman model and studies discount window lending in such a setup. Williamson also studies the role of deposit insurance in his proposed environment and how it interacts with the discount window. To begin the discussion, I will describe a simplified version of Williamson's model and consider only the effects of having a discount window in place. Later in the discussion, the role of deposit insurance in the model will be briefly considered.

There are two groups of agents in Williamson's economy. One group of agents has some resources and the other group has productive investment projects. The members of the former group will be called lenders, and the members of the latter group will be called borrowers. Investment projects have positive expected returns and take time to mature. The returns of each project have an idiosyncratic and an aggregate component, both random.

Demand for liquidity is motivated as follows. After investment takes place (and before it matures), a fraction of the lenders discover that they need to relocate. Investment cannot be moved across locations. Only liquid, low-return assets can be transported. Given this situation, it is optimal for agents to form a banking coalition (similar to those in Diamond and Dybvig [1983]), which allows lenders to pool their (in this case) relocation risk with other members of the coalition.

If it is possible to set up a bank that can be present in all locations (with branches, say), then the relocation risk can be fully insured.⁹ That is, all agents, those relocating and those not relocating, consume the same amount. In other words, those agents with an immediate need of liquidity do not suffer a cost in the optimal arrangement. This is actually different from what happens in the Diamond-Dybvig model, where the impatient agents consume less than the patient agents in the optimal allocation.

The reason for this difference is that in Williamson's model, agents need the liquidity to take on their relocation trip but not to consume it immediately. As a result, when a bank can be in all locations, it does not need to liquidate investment to supply consumption to agents who relocate – each location loses and gains some agents due to the relocation process and symmetry implies that the total consumption needs in each location, after agents have relocated, remain balanced. In the optimal arrangement, agents do not take liquidity with them as they relocate. They instead have a claim on the bank associated with their deposit, and they can withdraw resources, according to that claim, in the locations that become their final destinations.¹⁰

Things are different, however, when banks can be present only in one location. Williamson motivates this limitation by resorting to the long-standing restrictions on bank branching throughout U.S. history.¹¹ In that case, when an agent relocates, she needs to take with her low-return liquid assets. For this reason, insurance becomes costly and the best implementable allocation does not provide full insurance to relocating agents.

The timing of events in the model is such that some agents need to relocate before the return on investment gets realized. For this reason, payments to relocating agents cannot be made contingent on aggregate productivity, and only agents not relocating can bear that risk. This implies that when productivity is high, agents not relocating

⁹ Champ, Smith, and Williamson (1996) pursue a different interpretation by assuming that banks can issue notes that are transportable and that banks from different locations can exchange notes among themselves (or in a central market) on a regular basis. The results are essentially the same under either interpretation: branching or note issuance.

¹⁰ While productive investment produces random returns, liquid low-return investment is riskless. For this reason, the optimal allocation still assigns some resources to the riskless (liquid) technology. This pattern is the result of pursuing the optimal portfolio allocation under uncertain returns and not due to optimal liquidity provision.

¹¹ Under the alternative interpretation of note issuance, the assumption would be that banks are constrained in their ability to freely issue notes. Champ, Smith, and Williamson (1996) motivate such restrictions on the system prevailing during the national banking era in the U.S. In the model, restrictions on branch banking or note issuance are not explicitly motivated.

consume more than agents relocating, but when productivity is low, agents not relocating consume less than agents relocating. This pattern of consumption is also a generalization over the one that takes place in the Diamond-Dybvig environment, where the agents experiencing the liquidity shock always consume less than the agents who do not have a liquidity shock (since providing insurance is costly).

Williamson then introduces a discount window in the model. The central bank has a discount window office in each location. The way the discount window works is that the central bank issues its own claims in exchange for productive bank assets and those claims can be transported by relocating agents to their new location (where they can be redeemed at the local discount window office). Under this arrangement, if there are enough assets suitable for rediscounting, then all agents again receive the same consumption levels, regardless of whether they are or are not relocating.¹²

When the level of productive assets pledgable at the discount window is not large enough, there may be a role for a deposit insurance system in the economy. Deposit insurance enhances the ability of the economy to insure agents against the aggregate productivity shock. The details of the interaction in the model between discount window lending and deposit insurance are complicated and not essential for the discussion here.

There is no fiat money in Williamson's model. Instead, the central bank issues IOUs that can be redeemed the following period at the local office of the central bank, with real assets as backing. In a closely related paper, Smith (2002) embeds a version of the Williamson model in a dynamic general equilibrium overlapping-generations economy, which allows him to introduce fiat money and discuss the implications of having the discount window provide loans in fiat money.

In Smith's model, there is only one investment technology (with nonstochastic returns), which cannot be moved across locations unless it is liquidated at a cost. However, in the equilibrium with valued fiat money, banks can still diversify their portfolio between liquidity and productive assets – just as in the Williamson model. Here, though, holding liquidity amounts to holding money and is only useful to deal with the relocation activities of agents (since, in contrast with Williamson's setup, there are no aggregate technology shocks). Money

¹² In the model, there are two types of productive assets, and the optimal contract requires that one type of asset be monitored after the loan is granted. Those assets are, then, deemed not suitable for rediscounting as this reduces the incentives of banks to monitor them.

is a recognizable asset that can be transported across locations and used in transactions.

Agents in the model form banks that cannot communicate across locations. As in Freeman's (1996) model, money plays two roles in the economy: it allows the banks to make payments to relocating depositors and it allows the old generation to buy goods from young agents and banks in the typical overlapping-generations pattern. Because transactions between the old, the young, and banks happen before the relocation shocks are realized, the value of money is not contingent on the size of the relocation shock (i.e., the proportion of agents relocating). In the absence of a discount window, then, full insurance is not always optimal: when the relocation shock is large enough, agents relocating may receive less consumption than agents not relocating.

When a discount window is introduced in the environment, outcomes are sensitive to the interest rate used for rediscounting. Smith (2002) studies the case when the discount window provides loans at a rate that is equal to the nominal interest rate in the economy. Because banks are indifferent between holding money and taking loans at the discount window, price level indeterminacy becomes a feature of the equilibrium. However, Antinolfi and Keister (2006) provide a more general analysis of discount window policies and show that when the discount window rate is a "penalty rate" there is a unique steady state equilibrium in the economy.

Antinolfi and Keister (2006) also show that it is optimal to make the penalty on the discount window interest rate as small as possible. The logic follows from ideas discussed already by Williamson (1998). Basically, banks take loans at the discount window to provide better insurance to depositors. If those loans are provided at a penalty rate, then banks economize on them and reduce insurance. It is optimal to minimize the resulting misallocation of risk by reducing the cost of insurance as much as possible (without going as far as to make banks indifferent between holding liquid assets and taking discount window loans, which would result in equilibrium indeterminacy).

Smith's setup is also suitable for studying the interaction between monetary policy and discount window policy. Here, monetary policy is understood as the rate of growth of money, which in steady state translates directly into the inflation rate. Interestingly, by changing the real return on money, monetary policy can induce banks to invest more or less in the productive technology – as money competes with productive investment in the bank's portfolio allocation problem.

Smith shows that in this economy, in the absence of a discount window, it is not optimal to follow the Friedman rule (a policy of targeting the nominal interest rate to be zero), since such a policy tends to drive

productive investment to levels that are suboptimally low. As a result, in the optimum, the rate of return on productive investment is higher than the return on money, and banks do not fully insure the relocation shock. This situation, then, opens the door to operating a discount window.

Once the discount window is in place, banks have less reason to hold liquidity, and instead they choose to invest more. In fact, Antinolfi and Keister (2006) show that by combining a discount window with a monetary policy that closely approximates the Friedman rule, the economy can get arbitrarily close to the first-best allocation. In the resulting equilibrium, banks invest most of the proceeds from deposits in productive investment and borrow from the discount window to deal with the relocation (liquidity) shock.

Limited Commitment and Aggregate Liquidity

Holmstrom and Tirole (1998) set out to specify a micro-founded general equilibrium model where shortages of aggregate liquidity can happen. By comparison, in Freeman's (1996) model shortages of liquidity happen because some of the liquidity available to the economy is held, at certain crucial points in time, in the wrong hands. In Holmstrom and Tirole (1998), instead, *aggregate liquidity* (all considered) is insufficient to allow the economy to reach the optimal allocation of resources.

Generating an aggregate shortage of liquidity is not an easy task – particularly when the objective is to keep the argument, and all its details, fully specified. A natural reaction to informal descriptions of situations where there is a shortage of liquidity (or collateral) is to ask: Why wouldn't the price of the liquid assets adjust to resolve the shortage? Holmstrom and Tirole address this and other related issues explicitly and are able to provide a formal equilibrium model that delivers a shortage of aggregate liquidity. Equipped with such a laboratory, then, Holmstrom and Tirole address the question of government intervention without the drawback of having ruled out, for unexplained reasons, alternative arrangements that could, in principle, improve the situation. Given the nature of Holmstrom and Tirole's contribution, a meaningful discussion requires a relatively detailed description of the specifics of their model. We turn to this description next.

Consider a setup in which there are a large number of firms that invest in a productive technology that produces random returns. In an interim period, between the time when the initial investment happens and when the returns are realized, each firm requires an extra investment to keep production running. The amount of extra investment is

also a random variable, interpreted as a liquidity shock. After liquidity shocks are observed, the firms' managers may undertake costly effort in order to improve the probability of success of the investment (i.e., increase the probability that returns are high). This effort is not observable, so the manager of the firm must receive part of the return as compensation to provide him with incentives to make the appropriate amount of effort. The direct implication of this moral hazard problem is the *limited pledgability* of future cash flows.

At the time of the liquidity shock, firms need to obtain funding from external sources. If a firm does not obtain extra funding, production is discontinued and the (potential) future return is lost. Given this, the firm's borrowing capacity is given by the total expected future return on investment, net of the compensation to the manager.¹³

In the first-best allocation (that is, when manager effort is observable), all firms with a liquidity shock lower than the expected value of future returns receive funding. In the second-best optimum (that is, when manager effort is not observable), not all those firms may receive funding. Yet, it is the case that in the second-best optimum some firms with liquidity shocks higher than the expected value of future returns *net of manager compensation* do receive funding. This outcome cannot be supported in a laissez-faire market arrangement – in such a case, only firms that have *net* future expected returns higher than the liquidity shock will receive funding to continue the project.¹⁴

In the model, an assumption of lack of commitment limits the ability of the private market to provide adequate liquidity to firms. In particular, deep-pocketed investors cannot sell uncollateralized liquidity insurance to firms because those investors can default with impunity *ex post* – when the time to make good on insurance claims comes. The only way to provide insurance credibly is to use productive assets as backing for the resulting promises. Alternatively, firms could hold claims to those assets directly and sell them to investors when the liquidity needs arise.

Note, however, that the total available value of claims on productive assets depends on the value of future returns, which is (to the extent that funding is provided) independent of interim liquidity needs of

¹³ Firms are able to obtain external financing at the time of the initial investment because *expected* returns are positive. In many contingencies, when the liquidity shock is not too large, the initial investment delivers a high return. In other states of the world, though, when the liquidity shock is high, the *ex-post* return on that investment becomes very low, diluted by the issuance of new claims used to deal with the liquidity shock. While both contingencies are possible, expected returns at the initial investment stage are assumed to always be positive.

¹⁴ Holmstrom and Tirole consider partial liquidation, but it does not solve the problem because constant returns to scale make partial liquidation ineffective.

firms. Hence, those claims may not be sufficient to back the amount of insurance required to achieve the optimum. In that sense, the economy may experience an aggregate shortage of liquidity. To quote Holmstrom and Tirole (1998, p. 15): “Consumers cannot sell claims on (or borrow against) their future endowments because they can default with impunity. Only promises that are backed up by marketable assets (claims on firms) can be made. This is a key assumption. Without it, there would be no shortage of liquid instruments, nor any role for government intervention.”

Holmstrom and Tirole consider two cases: one where the liquidity shocks are independent across firms and the other when the liquidity shocks are correlated, creating aggregate shocks. When the liquidity shocks are independent, there is a private arrangement that can achieve the second-best optimum. Basically, firms form coalitions that resemble financial intermediaries. These intermediaries give each firm a committed line of credit that can be used if the firm cannot obtain enough funds in the market to accommodate those liquidity shocks that deserve funding according to the second-best allocation. In other words, financial intermediaries provide sufficient liquidity insurance consistent with the optimum and that insurance amounts to ex-post (i.e., after the liquidity shocks are realized) cross-subsidization across firms.

At this point, it is worth discussing briefly why intermediaries are needed. Recall that due to the lack of commitment, all claims from intermediaries need to be backed by claims on productive assets. Now suppose that instead of forming an intermediary, firms only trade claims on the future return of their investment. One possibility would be that firms, aside from making the initial productive investment, also dedicate some of their initial resources to acquire assets that they can later sell if necessary when the liquidity shocks are realized. Since the only store of value in the economy is the stock of claims issued by productive firms, the question becomes whether the total value of those claims is enough to implement the efficient (second-best) allocation. Holmstrom and Tirole show that this is not the case.

This shortage of liquidity is purely a matter of misallocation. In fact, at the level of the aggregate economy there are enough claims to potentially fund all the needed liquidity. However, since firms buy those claims before knowing their liquidity shocks, some of the claims end up in the hands of firms that do not need them – that is, those firms with low liquidity shocks. More succinctly, the fact that, ex post, liquidity is inappropriately distributed across firms is what makes it insufficient (as in Freeman [1996]). Forming an intermediary allows for a better ex-post allocation of liquidity and, in this way, improves outcomes. In practice, the way this happens is that some of the credit

lines are not fully drawn upon, permitting the liquidity to be more exclusively dedicated to satisfy the demands of firms that need to draw heavily on their credit lines.

With independent liquidity shocks, introducing intermediaries is “enough” and there is no role for government intervention. The case when the liquidity shocks are correlated, instead, provides a (potential) justification for government-provided liquidity. The basic idea is that the government can issue (noncontingent) claims on future tax revenue that firms then can hold (as a store of value) and potentially sell if and when they experience a large enough liquidity shock. If taxation is distortionary, the optimum may require that the bonds issued by the government sell at a premium (a liquidity premium). Then, given that premium, firms will adjust their demand for liquidity, which in turn determines the size of the liquidity shocks that they can withstand. Essentially, since holding government bonds is expensive, firms adjust their decisions in order to economize their reliance on those bonds.

With *noncontingent* government bonds, the best-attainable allocation (appropriately defined) may involve partial liquidation of investment (or the liquidation of some, but not all, firms with a given value of the liquidity shock). The implementation of this optimum is non-trivial: as Holmstrom and Tirole explain, one possibility would be to have some of the firms issuing both equity and short-term debt (with a specific, and somewhat unrealistic, covenant).

The government, though, can actually improve the allocation by issuing *state-contingent* bonds that pay a positive amount only when extra aggregate liquidity is needed. The rationale for this is simple: noncontingent bonds will provide excess liquidity in most states of the world. Since this liquidity is expensive to create – as it involves distortionary taxation – it is optimal to minimize the production of excess liquidity. Holmstrom and Tirole use this result to motivate possible state-contingent policies (monetary and fiscal) that are aimed at managing the provision of aggregate liquidity.

In a companion discussion of aggregate liquidity shortages, Holmstrom and Tirole (1996) explicitly consider discount window lending as a possible (state-contingent) policy that could be used to achieve the socially optimal allocation without resorting to the more uncommon state-contingent bonds. In principle, one interpretation would be that the counterpart of the necessary premium on government bonds is to have a penalty rate at the discount window. This interpretation, then, provides a possible justification for an *optimal penalty* rate at the discount window on the basis that government production of

liquidity involves distortionary taxation.¹⁵ Note finally that the discount window would be used only in situations when there is an aggregate liquidity shock and “insufficient” private claims. This implies that the discount window would be particularly active at times that are often considered crisis-like situations.

The model by Holmstrom and Tirole (1996) highlights the close connection between monetary and fiscal policy when liquidity demand refers to the need to access riskless claims issued by the government. It suggests that in many cases the LLR function could be handled directly by the fiscal authorities, in particular if government bonds and reserves serve an equivalent role for solving the issue at hand.

Deposit Insurance, Bank-Failure Resolution, and Bankers’ Incentives

Freixas, Parigi, and Rochet (2004) present a model where bank depositors are fully covered by deposit insurance and it is the government (i.e., the party providing insurance to depositors) that needs to design the appropriate framework to manage bankers’ incentives. The optimality of deposit insurance is not addressed in the paper. In fact, depositors are assumed to be deep-pocketed, risk-neutral individuals. The paper can be seen, then, as an effort toward understanding how to organize the banking system, and the relevant government interventions, given that a decision has been made to provide deposit insurance.¹⁶

There are three relevant periods in the model. In the initial period, bankers take deposits and complement those funds with their own capital, which is assumed to be a fixed amount. With those funds, bankers make risky investments. In the interim period, bankers find out the state of their finances. Finally, in the last period, payoffs are realized.

Three situations are possible in the interim period: the banker may be solvent or insolvent, and if the banker is solvent, it may or may not experience a liquidity shock in the form of deposit withdrawals. Furthermore, if the banker is insolvent, it could in principle engage in “gambling for resurrection” by borrowing some funds and investing

¹⁵ The traditional Bagehot doctrine on discount window lending involves a penalty rate as well. This penalty is often motivated as a way to control moral hazard. However, the implications of a penalty rate on incentives can be subtle. Castiglionesi and Wagner (2012), for example, demonstrate that under some conditions a penalty rate may actually increase moral hazard.

¹⁶ Repullo (2000) studies the conflict of interest between a deposit insurance agency and a central bank confronting the decision to lend to a bank in need of liquidity. See also Kahn and Santos (2005).

them in a way that gives the banker a small probability of recovering from insolvency.

When the bank is solvent, investment has a positive probability of paying out in the last period. Investment by insolvent banks, instead, is sure to not produce any payoff in the future, unless the bank gambles for resurrection and succeeds. Bankers' incentives play a role in the initial and interim periods.

First, in the initial period, bankers can exert some costly effort to increase the probability that the bank will be solvent. In turn, if the bank is solvent, then the banker can exert some effort in the interim period to increase the probability that investment will be successful. In both cases, exerting effort is socially desirable, and the job of the government is to design a system that compensates bankers so as to induce them to do what is best for society.

Increasing bankers' compensation reduces the resources left to pay depositors (since bank capital is assumed fixed). Hence, the only way to accommodate higher banker compensation is to reduce total deposits and, hence, total investment. Since investment is productive, reducing the level of investment is costly for society. In other words, there is a trade-off between compensating bankers and the level of total investment undertaken by banks. The incentive design problem then involves compensating bankers with the minimal amount that would still induce them to exert effort. This incentive problem is similar to the one analyzed by Holmstrom and Tirole (1998).

In some situations, providing incentives to bankers in the interim period involves paying them enough that it is optimal for them to also exert effort in reducing the probability of insolvency in the initial period. In those cases, illiquid banks can resort to the interbank market for funds without compromising the optimality of the allocation.¹⁷ The more interesting cases occur when extra incentives are needed to reduce the bank's probability of insolvency.

In principle, an unverified insolvent bank can pretend to be an illiquid bank, take a loan in the interim period (of the same amount as illiquid banks), and use those funds to gamble for resurrection. To avoid this situation, insolvent bankers need to be compensated to agree to identify themselves as insolvent (and not engage in socially wasteful gambling for resurrection). This may require that, upon failure, shareholder value is not fully wiped out. In this sense, bank-resolution policies are an important component of the incentive scheme.

¹⁷ Strictly speaking, the optimal allocation is implementable only if interbank loans are not subject to repayment risk. This requires that the size of the loan be small relative to the lower bound on the return of investment.

Given that insolvent bankers receive a positive compensation after declaring bankruptcy, the efficient allocation requires that banks borrowing funds in the interim period pay a premium for those funds. There are two reasons for this: first, by charging a penalty rate, the government reduces total shareholder compensation and manages the trade-off between incentive provision and total investment. Second, the penalty rate is a way to induce sorting: once funds are offered at a premium, insolvent banks have no incentives to borrow, while illiquid banks still do.¹⁸

One way to implement the optimal allocation is to have interbank loans be junior to claims of the deposit insurance fund and have discount window loans be senior to both. Under such a situation, the central bank has the ability to fine-tune the pricing of discount window loans, establishing the appropriate penalty rate consistent with optimality, which can still be below the alternative rate that banks would need to pay in the interbank market (where loans are uncollateralized, junior claims).

To close this discussion, it is worth pointing out that in the Freixas-Parigi-Rochet model this kind of arrangement where discount window loans (at a penalty rate) can be part of the optimal way to organize the banking system in the presence of deposit insurance depends on several particular conditions on parameters. In many other situations, the discount window has no clear role in the sense that it cannot improve on what can be achieved with only an interbank market and, in those cases, it may not be possible to decentralize the optimal allocation.

Unique-Equilibrium Coordination Failures

Rochet and Vives (2004) study a banking problem where the assumed banking arrangement may create a coordination failure. While the model has the flavor of the Diamond-Dybvig model, there are some significant differences. To start, contrary to Diamond and Dybvig (1983), this paper does not focus on understanding the constrained-optimal allocations without any exogenously imposed institutional constraints. In fact, some features of the banking arrangement are taken as given without providing explicit micro-foundations. The emphasis, instead, is in understanding the implications of those features once they are in

¹⁸ This is an alternative theoretical justification from the one provided by Holmstrom and Tirole (1996) for charging a penalty interest rate on loans at the discount window. To the extent that this penalty rate is more about managing bankers' incentives, it is closer in interpretation to Bagehot's doctrine.

place.¹⁹ The authors discuss throughout the paper possible avenues to approach the micro-foundations question. Their overriding objective, however, is to try and keep the framework as simple as possible to be able to employ the global-games methodology (Morris and Shin 2003) that pins down equilibrium even in the presence of coordination failures.

Banks in the model have some capital and receive some deposits from investors. With those resources, the bank can invest in risky assets or in reserves. Investment takes two periods to mature, at which time it delivers some returns. In an interim period, before investment matures, depositors are entitled to withdraw their deposits from the bank. There are no liquidity shocks (there are no impatient agents of the Diamond-Dybvig type). Instead, depositors receive idiosyncratic signals about the future return of investment and, for each depositor, if her signal is (sufficiently) bad, then she decides to withdraw her money from the bank early (in the interim period).

When some depositors withdraw early, the bank can use the reserves to pay those depositors. If withdrawals are higher than reserves, the bank sells the investment in the market at a discount. The discount is exogenously assumed, and it stands for possible fire sales (or other sources of liquidity premia). The authors discuss how adverse selection could motivate fire sales, but this aspect is not explicitly modeled. Finally, if the bank cannot repay the promised amount to all depositors (early and late withdrawers), it fails. The bank is not allowed to adjust payments to depositors; that is, payouts are noncontingent, unless of course the bank fails.

Depositors' preferences are not explicitly spelled out. Rather, depositors follow what the authors call a *behavioral rule*: each depositor wants to withdraw if her individual assessment of the probability that the bank will fail is high enough. The authors postulate that this is a reasonable rule to capture the behavior of fund managers investing in, say, jumbo CDs at banks. The authors argue that this interpretation is more in line with "modern" versions of bank runs (where withdrawals by wholesale-funding sources play a prominent role).

There are situations in the model when the bank is solvent but may still fail because of the need to accommodate early withdrawals by liquidating assets at a discount. The idea is that when there are fire

¹⁹ Broadly speaking, this is the approach also taken by Freixas, Parigi, and Rochet (2004) and, to a lesser extent, by Bhattacharya and Gale (1987) and Williamson (1998). As we discussed before, this approach is open to the criticism that any attempts at endogenizing the institutional arrangements may require new assumptions that could have important implications for (and potentially undo) the results discussed by these authors.

sales, if enough agents withdraw early, then the bank liquidates assets (i.e., invested resources) at a discount and there are fewer resources available to pay other depositors. This makes failure more likely and feeds back to the number of withdrawals, increasing it as more depositors conclude that the bank will fail given their private signal of the value of the return on investment.²⁰

The discount window is assumed to have an informational advantage originated in the central bank's supervisory powers. This information allows the discount window to recognize the "true" value of the assets of the banks, not the one implied by the fire sales. In this way, using the discount window, the central bank can avoid any early liquidation of assets and, hence, the failure of solvent banks.

The central bank is assumed to also have access to funding at no extra cost. In other words, the central bank can access resources without having to increase distortionary taxation. These extra resources available to the central bank are not explicitly modeled, and it is not clear what would be optimal if these resources were explicitly taken into account from the start.

In the model, discount window credit should not be provided at a penalty rate. Rochet and Vives discuss many of the factors not present in the model that would suggest that a penalty rate may be appropriate – for example, if the central bank has better information relative to the private sector, but not perfect information. There is also a discussion of the possibility that the private sector could provide lines of credit to banks and then actively monitor them (as suggested by Goodfriend and Lacker [1999]). The authors point out that this could be an appropriate approach when there are no central-bank advantages in supervisory knowledge and financial capacity.

Late in the paper (in Section 7), the authors sketch a justification for the deposit contract that allows agents to withdraw in the intermediate period. As in Freixas, Parigi, and Rochet (2004), the idea is that the bank manager needs to exert effort to improve the distribution of possible investment returns, but that effort is costly and not verifiable. A way to give incentives to bank managers is to allow depositors to withdraw early. The bank manager is assumed to specially dislike bank failure in the intermediate period, and depositors can "discipline" the manager by threatening to withdraw early (Calomiris and Kahn 1991).

²⁰ Depositors follow a threshold rule: if the signal is below a threshold, then the depositor withdraws early. The threshold gets determined in equilibrium and depends on the strategy of other depositors (it is a fixed point) because if more depositors withdraw, the bank is, for a given return on investment, more likely to fail (as more withdrawals mean more early liquidation at discounted values).

A couple of interesting situations may arise in this case. First, it is no longer efficient for the central bank to intervene in a way that rules out all possible early bank failures. Some early bank failures are necessary to provide incentives to bank managers. Still, the equilibrium without intervention may result in too many early bank failures. Even if the bank is insolvent, there are cases where it is efficient for the central bank to provide credit in the intermediate period to avoid early liquidation of assets at discounted (fire sale) values. The bank will still fail in the second (final) period, but losses would be lower (even after the loan from the central bank is fully paid back).

Perhaps more interesting is the fact that there are some cases where the central bank needs to intervene and close down a bank in the intermediate period even though the bank is solvent. The authors interpret this as a “prompt corrective action” rule. The idea is that sometimes, to give bank managers incentives to exert effort, the bank needs to be liquidated early even if the bank would be solvent (assuming that the central bank provides appropriate discount window liquidity). Hence, a central bank that has a discount window open to all banks needs to complement that policy with a “prompt corrective action” policy when the incentive problem of bank managers is severe enough.

Inalienable Human Capital and Banking

In a series of papers published in the early 2000s, Diamond and Rajan developed a comprehensive theory of banking. Initially, they studied what banks do and the optimal structure of banking contracts and of banks’ balance sheets. In a second stage, they extended the model to address systemic banking crisis (Diamond and Rajan 2005). This extended model is the one used here as a basis for the discussion.

There are three types of agents in the economy: investors, bankers, and entrepreneurs. Entrepreneurs have projects that need funding. Investors have funds that could be used to fund those projects. Funding is in short supply, though, so only a portion of the projects can actually get funding. All projects pay the same return after some period of time. While ex ante all projects are identical, ex post some projects take longer to mature. In other words, a proportion of the projects pay their return early (“early projects”), and the rest (“late projects”) pay some time later.

Investors need to consume early (that is, at the time when the early projects pay out their return). Bankers and entrepreneurs, on the other hand, can consume late.

There is an information friction that complicates the funding of projects. Entrepreneurs cannot commit ex ante to run their project

after receiving funding, and their human capital is essential for running the project successfully. One way to think about this lack of commitment problem is that the courts system cannot force entrepreneurs to dedicate their *inalienable* human capital to the process of running a project (Hart and Moore 1994). As a result, after a project receives funding, the entrepreneur can threaten to withdraw his human capital and induce a renegotiation of the terms of the loan – an instance of the well-known hold-up problem.

Bankers have a technological advantage over investors. In particular, bankers are able to learn about the project and run it if necessary. The return that the project delivers when run by a banker is a fraction of what the entrepreneur can get, but it is not zero. Hence, the banker can fund the entrepreneur up to the amount that the banker would be able to get when running the project himself. If the entrepreneur tries to renegotiate, the banker takes over the project and runs it himself. Knowing this, the entrepreneur does not attempt to renegotiate. Here, it is important that the loans to entrepreneurs are callable (i.e., that the bank can ask for repayment at any time and take over the project if the repayment does not happen).²¹

Since only investors, not bankers, have the resources to fund the projects, channeling funds to entrepreneurs requires that investors make deposits at the banks and those banks make loans to the entrepreneurs. As a result, a second hold-up problem arises: the banker, after receiving deposits from investors, could try to renegotiate the contract by threatening to not collect from the entrepreneurs.²² One way to solve this second hold-up problem is to create a collective action problem among the bank's depositors. In particular, the bank can offer uninsured demand deposits that are paid out on a first come, first serve basis and obtain deposits from a large number of investors. Courts can enforce deposit contracts as long as the bank has funds. Under these conditions, there is an equilibrium in which investors/depositors run whenever the bank attempts to renegotiate down the payments associated with the deposit contract. If a depositor thinks that other depositors will run when threatened with renegotiation, then it is in her best interest to run as well, even if in the end depositors, as a group, obtain less than if the run would not have occurred.

²¹ Diamond and Rajan (2005) have a discussion of the empirical relevance of callable bonds in their paper and further point out that a significant proportion of outstanding commercial and industrial loans in the U.S. are of very short maturity (which makes them essentially callable).

²² No one else but the bank that initially funded an entrepreneur has the ability to collect from the entrepreneur, so the loans are illiquid from the perspective of the bank. That is, the bank cannot sell its loans in the market.

Similar to Rochet and Vives (2004), here runs on the bank are a way to provide appropriate incentives to bankers. Entrepreneurs cannot receive funding directly from investors due to an extreme hold-up problem. The possibility of runs, in turn, controls the hold-up problem between bankers and investors (depositors). This allows funds to flow from investors to entrepreneurs, through bankers, in a way that otherwise would not be possible.

Bankers can also restructure projects at any time before the projects mature. A restructured project yields some resources immediately and some resources in the future. The payouts from a restructured project can be collected by anyone (there is no hold-up problem in that case). However, restructuring projects is costly in the sense that a restructured project yields fewer resources than the initial investment.

Each bank is subject to an idiosyncratic shock that determines the fraction of projects in its portfolio that are maturing early. This shock is crucial for bank solvency. Projects that mature early provide resources to pay initial investors – all of them needing to consume early. So, a high fraction of projects maturing early makes the bank more likely to be solvent. The bank can also access a market for liquidity and try to obtain resources by borrowing against the return from the late projects. How much liquidity the bank can obtain depends on the market interest rate. If the bank cannot raise enough liquidity to pay all initial investors, then it is deemed insolvent.

The market for liquidity at the time when initial investors need to consume is a key market in the model. The interest rate in that market plays a role in determining bank solvency and project restructuring. In turn, demand and supply of liquidity in that market depends on depositors' and banks' decisions. Let us now briefly discuss how demand and supply of liquidity in that market get determined and how they depend on the interest rate.

Project restructuring impacts both supply and demand of liquidity. Both solvent and insolvent banks may engage in project restructuring. The decision of solvent banks to restructure late projects depends on the market interest rate. If the interest rate is low, then a solvent bank will choose to continue all projects. For intermediate values of the interest rate, a solvent bank will choose to restructure only enough projects to pay back initial investors. Finally, if the interest rate is high, a solvent bank will choose to restructure all of its late projects. The reason behind this pattern of behavior is simple. To continue funding late projects, a bank needs to attract new deposits. If the interest rate is high, deposits are costly. If, instead, the bank just restructures late projects, it obtains immediate liquidity.

Bank solvency also depends on the market interest rate. In principle, a bank can become insolvent just because the interest rate in the market is too high. Since borrowing is backed by the future discounted value of late-project returns, when the market interest rate is very high the bank cannot borrow as much. As a result, the bank has access to less liquidity and may not be able to pay initial investors in full.

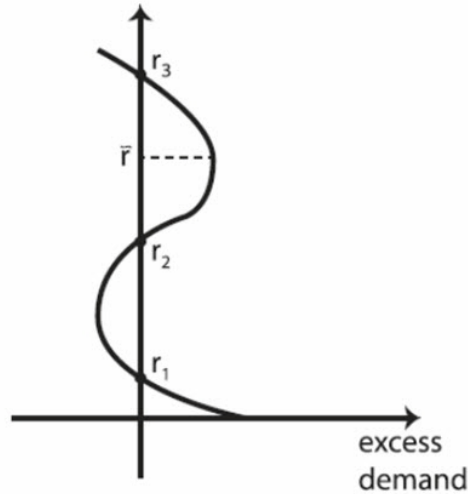
The demand and supply of liquidity in the market also depend on bank solvency. One basic source of liquidity in the market is the income of entrepreneurs with projects that mature early and have a loan from a solvent bank. Due to limited pledgability and lack of commitment, these entrepreneurs' income is higher than what they need to repay the bank. Since entrepreneurs do not need to consume early, they can reinvest these extra resources by lending to solvent banks in need of funding.

Solvent banks with late projects in their portfolios need to obtain extra liquidity in the market to pay back initial investors without having to restructure those projects. The banks can, then, use the future return on those late projects to guarantee their ability to repay new loans when they become due.

Bank liquidation impacts the demand and supply of liquidity in complex ways. Restructured projects tend to increase the demand for liquidity in the market. The reason is that restructured projects generate pledgeable future income that banks would want to use in order to borrow extra liquidity. Bank liquidations, by forcing project restructuring, also increase the demand for liquidity in the market. Furthermore, bank runs trigger restructuring of early projects, which reduces the supply of liquidity because entrepreneurs receive less income to reinvest.

The multiple effects and interactions between the market interest rate and banks' decisions imply that the excess demand function for liquidity may be nonmonotonic. Conditional on a given number of bank failures, increases in the interest rate tend to lower the excess demand for liquidity. However, changes in the number of bank failures can change this relationship. Both solvent and insolvent banks demand liquidity from the market. When movements in the interest rate push banks out of the solvent group and into the insolvent group, the impact on total demand (and supply) of liquidity can result in a segment of the excess demand function with positive slope (see Figure 1).²³

²³ An increase in the interest rate can increase the excess demand for liquidity when the proportion of early projects is high in banks that switch from solvent to insolvent due to the interest rate increase. When banks with a high proportion of early projects become insolvent, the resulting restructuring of those early projects creates an extra demand for liquidity in the market.

Figure 1 Excess Demand for Liquidity

We have discussed so far how projects get funded and why banks are needed in that process. We also discussed the determination of demand, supply, and interest rates in the market for liquidity. A crucial factor driving policy interventions in the model is the possibility of bank insolvency. The timing of the arrival of information is important for this issue. In particular, information about banks' idiosyncratic shocks arrives *before* early projects pay their return. Initial investors observe these shocks and have rational expectations about market interest rates. At that time, then, they can calculate whether a bank will be solvent given the expected interest rates. If the bank is insolvent, initial depositors run on the bank. In response, the bank restructures all projects, even those that are expected to mature early. This surge to generate liquidity is socially very inefficient. In fact, note that liquidity is not really needed at that point. Depositors are demanding liquidity even before anyone needs to consume. This nonfundamental demand for liquidity is a direct consequence of the self-fulfilling run on demand deposits.²⁴

²⁴ Because the bank is insolvent, depositors expect a (necessary) renegotiation of their deposit contracts in the near future. Since depositors also expect that other de-

The payoffs associated with bank deposits are not allowed to be contingent on the realization of the idiosyncratic shocks. This contract incompleteness is particularly consequential when a high proportion of late projects put a bank in an insolvency position. In equilibrium, depositors at such a bank anticipate renegotiation and, hence, decide to run. This run is not essential for disciplining the banker. The banker, regardless of behavior, just does not have enough resources to pay every initial investor in full. Adjusting down payments would be more efficient than liquidating the bank and restructuring all projects. This is an important source of inefficiency in the model.

Diamond and Rajan are forthright about the importance of this assumption. For example, they tell us that they “do not allow contracts to be directly contingent on the state,” which they assume “is observable but not verifiable” and explain that their “analysis is positive – to show what happens when there is an ex-post solvency problem or an aggregate liquidity shortage given the use of demand deposit contracts.” To be sure, they confirm that “if there was no uncertainty about the ex-post state of nature, or if there were complete markets, no such conditions would arise.”²⁵

Diamond and Rajan are particularly interested in periods of crisis. One way to think about banking crises in this environment is to consider the case when the average fraction of early projects across banks in the economy is a random variable. When most banks have lots of early projects, those banks are solvent and bank runs do not occur. However, when the proportion of early projects is low for many banks, some of these banks become insolvent and experience runs. Depending on how bank failures and the consequent restructuring of projects impact market liquidity, situations that can be regarded as inefficient banking crises can develop in equilibrium.

Diamond and Rajan discuss examples of this kind of inefficient crisis. In particular, they describe a situation where the banking system, in principle, could satisfy the liquidity needs of investors by

positors will react to the threat of renegotiation by withdrawing their deposits, the expectation of that renegotiation immediately triggers the run.

²⁵ The deposit contracts considered by Diamond and Rajan are *demand* deposits in the sense that withdrawal can happen at will and at par, at any time. The bank does not offer, for example, term deposits for which withdrawal cannot happen until the early projects have matured. In principle, the “callable” feature of deposit contracts can be formally justified within the structure of the model. Since the banker can try to renegotiate the payment to initial investors at any time (by threatening to not collect from entrepreneurs) the deposit contracts have to be “runnable” at all times (to discipline the banker). If deposits were term deposits, then, before the deposit matures, the bank could try to renegotiate the claims of initial investors and these investors would not be able to run on the bank because the courts would not enforce repayment until the term of the deposit expires.

restructuring only late projects. However, in equilibrium all banks experience a run and all projects, including early maturing ones, are restructured. Diamond and Rajan call this situation a *systemic meltdown*.

To gain some intuition over why a meltdown can happen, notice that solvent banks do not restructure late projects when the interest rate is low, and hence the supply of liquidity in the market is also low. With an excess demand for liquidity, the interest rate tends to increase, which can make more banks insolvent. As insolvency produces more restructuring of early projects, this reduces even further the supply of liquidity in the market – increasing excess demand, which then can be increasing in interest rates. If the interest rate keeps rising, eventually all banks are subject to a run and all projects get restructured. This process generates a form of financial “contagion” that spreads via changes in market interest rates.²⁶

As Diamond and Rajan explain, the timing of arrival of information is crucial for this “contagion” outcome. In the model, information about the aggregate state (the distribution of shocks to banks) arrives earlier than the time when liquidity is produced and consumed, and depositors redeem their claims immediately after the arrival of this information, in anticipation of future liquidity shortages. The resulting wave of withdrawals forces banks to restructure projects – even early projects that would otherwise be a source of liquidity in the market.

The inefficiency of equilibrium outcomes opens the door to potentially beneficial policy interventions. Diamond and Rajan define two benchmark policy interventions: a *pure liquidity infusion* and a *pure recapitalization*. A pure liquidity infusion involves providing loans to banks at the prevailing interest rate. In a pure recapitalization, banks receive a transfer (a subsidy) in the form of claims that can be traded in the market to obtain liquidity and avoid failure. The policy authority has the ability to tax some agents in order to fund the interventions. Diamond and Rajan argue that, in their model, any financial market intervention can be viewed as a combination of these two pure forms of intervention.

Of particular interest for the subject of this article – i.e., discount window policy – is the case of a pure liquidity infusion. To fund the loans, the policy authority taxes investors after they had a chance to withdraw from their bank. Under certain conditions, this intervention can increase market liquidity and lower the interest rate so that fewer

²⁶ Freixas, Parigi, and Rochet (2000) and Allen and Gale (2000) are classic papers studying contagion originating in more standard types of spillovers in banking – such as when one bank’s failure directly creates losses on other banks’ loan portfolios.

banks are insolvent and fewer inefficient runs happen in equilibrium. This policy intervention is clearly not a Pareto improvement since some investors are taxed and end up consuming less. Still, the policy may produce a welfare improvement by reducing inefficiencies associated with the unnecessary liquidation of banking assets.

The nonmonotonicity of the excess demand function in the market for liquidity implies that, under certain conditions, the model also has multiple equilibria. Figure 1 illustrates such a situation (see Lemma 3 in Diamond and Rajan [2005]). The three possible equilibrium interest rates are r_1 , r_2 , and r_3 . When the market interest rate is above \tilde{r} , only self-sufficient banks (which do not need to obtain liquidity in the market to pay back investors) are solvent. In the equilibrium where the interest rate is r_3 , the initial investors expect a high interest rate that makes all but the self-sufficient banks insolvent. As a result, all banks that are not self-sufficient suffer runs, and the liquidity in the market confirms the high expected interest rate r_3 .

To avoid a highly inefficient situation like the one associated with interest rate r_3 , a policy authority (a central bank) can put in place a discount window that offers loans to banks at an interest rate between r_1 and r_2 . Under such policy, the only equilibrium in the market would be the more efficient situation associated with interest rate r_1 . In fact, discount window activity would actually not be observed in equilibrium. The discount window acts just as a mechanism to coordinate agents' expectations. It is important to recognize here, though, that such policy only works if agents believe that the central bank would have access to sufficient tax revenues were actual discount window lending to become necessary. In other words, policy credibility is essential, even if never tested.

Widespread Pessimism and Flight-to-Quality Episodes

Caballero and Krishnamurthy (2008) study an economy with a continuum of agents who may experience a liquidity shock (an urgent need to consume). The shocks are correlated across agents: in particular, the economy may experience one or two “waves” of shocks. If the economy experiences one wave of shocks, then half the agents in the economy receive the shock. Who in the population receives the shock is random. If the economy experiences two waves of shocks, then those agents who did not receive a shock in the first wave receive a shock in the second wave. Agents know the probabilities of the economy experiencing none, one, or two waves of shocks, but they do not know if they will receive a

shock in the first wave or in the second wave if the economy experiences two waves.

There is no asymmetric information in the economy, and there are complete markets. Hence, agents can enter insurance contracts (or buy and sell contingent claims) in order to insure the liquidity shocks. In the optimal allocation, agents buy more insurance for the eventuality of receiving a liquidity shock in the first wave than in the second wave. This is the case because the first wave is more likely (the economy can experience a second wave only if it has already experienced a first wave).

While in the Caballero and Krishnamurthy economy there is no sequential service constraint (Diamond and Dybvig 1983), there is sequentiality in the revelation of the state: first agents find out that there has been a first wave and, only after making/receiving the payments corresponding to that contingency, agents find out if there is a second wave.

Note that after entering an insurance agreement, if an agent gets a shock in the second wave, we could say that he is ex-post unlucky. That is, the eventuality that was less likely and the one for which the agent bought less insurance has actually happened. Caballero and Krishnamurthy consider the possibility that agents could become overly pessimistic about this contingency. If such is the case, this effectively increases the probability that the agents assign to the event of being affected by the second wave of shocks. As a result, then, agents increase the amount of insurance that they buy to cover that contingency. If this bias is large enough, agents basically insure the second-wave shock as much as they insure the first-wave shock (note that if they do, then there is no longer a sense in which an agent who receives the shock in the second wave is unlucky relative to an agent who receives a shock in the first wave – they both consume the same).

Caballero and Krishnamurthy use the model to think about a situation where there is a *flight-to-quality* event. In normal times, agents in the economy use an unbiased estimate of their likelihood of being in the second wave of liquidity shocks. However, when something unusual and unexpected happens (something that does not directly affect them but makes them pessimistic about their prospects), agents become overly pessimistic about their own situation (not the situation of the aggregate economy) and act as if they were more likely (than what they really are) to experience a second-wave shock. As a result, agents start buying more insurance for the second-wave shock and lower the amount of insurance they buy for the first-wave shock. The direct implication is that there ends up being much less liquidity during the (much more likely) first wave of shocks.

Caballero and Krishnamurthy study optimal intervention by a central bank in their environment. The central bank does not have more information than agents, but since the central bank cares primarily about “aggregates” and not about particular agents, it is not exposed to the pessimistic biases of individual agents.²⁷ During a flight-to-quality episode, the central bank could improve outcomes by (somehow) inducing agents to insure less against the second-wave shock and more against the first-wave shock.

Caballero and Krishnamurthy consider the case when the central bank has access to resources (“collateral”) that private agents do not have. They argue this is similar to the assumption in Holmstrom and Tirole (1998) where the central bank can exploit the power it has to impose future taxes. The point they want to make is that the benefit of having and using those resources during a flight-to-quality episode is higher than the direct cost of obtaining those resources. The policy intervention calls for the central bank to promise agents that it will provide them with resources if the second-wave shock hits. Agents anticipate the central-bank contingent-transfer and reduce the amount of insurance they obtain against the second-wave shock. At the same time, they increase the amount of insurance they obtain for the first-wave shock (the more likely shock). Note that the intervention is only in the rare event that two waves of shocks hit the economy. In that sense, the paper provides justification for a “last resort” intervention policy. Note, however, that the policy influences outcomes not just when they happen but in all eventualities, because it changes agents’ ex-ante decisions and improves the insurance arrangements (which would otherwise be inappropriate due to biases that agents have in assessing the probability of extreme, unfavorable individual events).

It is interesting to note (as the authors do) that the LLR policy is aimed at correcting decisions by private agents to over-insure some shocks and under-insure others. For this reason, moral hazard is not a big problem. It is the case that agents reduce private insurance for the shocks that are insured by the central bank, but agents also increase the amount of private insurance on the shocks that were previously under-insured. There is a certain degree of complementarity between the public insurance of the second-wave shock and the private insurance of the first-wave shock (when one increases, the other increases too). The provision of public insurance for some (unlikely)

²⁷ A key feature that allows the central bank to intervene efficiently without having any a priori advantage over agents (in information or perceptions) is that the source of the problem is not that agents are overly concerned about aggregate shocks. They are overly concerned only about the impact of those aggregate shocks on their individual outcomes. So, when the central bank aggregates outcomes, the biases disappear.

shocks helps to correct the under-insurance of other (more likely) shocks, which are not publicly insured. Broadly speaking, the central-bank backstop in this model does what central-bank backstops generally do – it reduces agents’ incentives to insure themselves against the event that the central bank is backstopping them on. Usually, the weakening of incentives results in inefficiencies (due to moral hazard). Here, however, given that agents were initially over-insuring the risk in question, the effect of the backstop is to correct a distortion and improve efficiency.²⁸

It seems likely that an agent with deep pockets and an unbiased assessment on the probability of different shocks to individuals (or with the ability to diversify/aggregate/pool risk across individuals) would be able to sell insurance at a profit. The paper is clear in explaining that if there are sufficient resources in the economy, then the optimal allocation involves full insurance of all shocks and misperceptions are irrelevant for allocations. Only when liquidity is limited (that is, when resources available in the short run are limited) agents’ misperceptions create a problem. Still, the paper does not consider a situation with heterogeneous agents, some with plentiful resources and some with limited resources, such that some beneficial trade could occur. It seems likely that in those circumstances, the misperceptions of some agents could create extraordinary benefits for other agents (those with access to liquidity).

In the model, the central bank is just an agent who has access to resources via taxation and is able to redistribute those resources to correct inappropriate private insurance arrangements. But the central bank is benevolent, so it is not exploiting agents’ misperceptions.²⁹ The extent to which a private deep-pocketed and self-interested agent could improve the overall situation is, hence, not clear.

There is no formal explanation for what triggers the switch of agents to a pessimistic state. Caballero and Krishnamurthy provide an informal discussion of the situations that are most likely to trigger such perception shifts. New (unanticipated, unknown) events and

²⁸ Intervention could still create moral hazard if agents have to incur costs to become better informed about the economy and the nature of the shocks. This may be especially important if such information could make them less prone to misperceptions that make them overly pessimistic.

²⁹ Note that defining a benevolent central bank is not without complications. Basically, the central bank ends up being paternalistic in the sense that it is using probabilities to assess outcomes that are different from the priors used by the agents. If the agents take their priors as fixed features of their preferences, then using different probabilities is not consistent with choosing agents’ most-preferred allocations. The authors discuss this at length in the paper and provide several interpretations that justify their use of a paternalistic central bank.

innovations play a prominent role in their discussion. Based on these informal discussions, they conclude that interventions that are aimed at dealing with new and unknown situations and that create and coordinate understanding of such situations (such as facilitating discussions among major market participants, in the spirit of the intervention by the New York Fed during the collapse of the hedge fund Long-Term Capital Management in 1998) could be beneficial.³⁰

Limited Enforcement in the Interbank Market

Gertler and Kiyotaki (2010) explore a dynamic macroeconomic model with a financial intermediation sector and frictions in the ability of the financial sector to obtain external funding. In the model, there are a large number of firms and a large number of islands (locations, sectors). While labor can move freely across firms and islands, capital cannot. Furthermore, each period the firms in some islands (but not in all) receive an opportunity to invest. Firms that are able to invest generate extra demand for liquidity for the financial intermediaries (banks) of their island. In consequence, some banks in the economy value liquidity more than others, and an interbank market for funds can emerge to exploit the gains from trade.

Financial intermediaries are assumed to be necessary to channel funds from depositors (and potentially other banks) to productive firms. There are a large number of households with intertemporal preferences over consumption and leisure, and their behavior can be characterized using the problem of a representative household. Gertler and Kiyotaki make further technical assumptions to facilitate aggregation in the production side of the economy and keep the model tractable. Without any financial frictions, the economy actually reduces to a (relatively standard) Real Business Cycle model.

Financial intermediaries take deposits from households to then fund firms' capital and investment. The friction in the intermediation process is a version of the agency problem studied by Kiyotaki and Moore (1997), which endogenously imposes a limit on the ability of banks to obtain external funding. The basic assumption is that after a bank obtains funds and acquires claims on productive firms, the bank can divert a fraction of those claims for its private benefit and default on its creditors. To avoid default, creditors are willing to fund only a portion

³⁰ A topic that remains largely unexplored in the theoretical literature is the role of the discount window during system disruptions (Lacker 2004; and Ennis and Price 2015). A notable exception is Martin (2009).

of the total claims held by a bank. Bank net worth is needed for the rest.³¹

The evolution of bank net worth depends on past investment by the bank and the cost of its funding. Gertler and Kiyotaki assume an exogenous exit rate from banking so that the accumulation of net worth over time does not fully resolve the agency problem. Banks can attract deposits from any island in the economy, but they can only buy productive claims from firms in their island. Furthermore, banks access the deposit market before knowing which islands will receive the investment-opportunity shock. By the time the shocks are realized, the deposit market is closed. These assumptions induce a level of market segmentation that is crucial for the outcomes of the model.

The market for productive claims in an island has firms on the supply side and banks on the demand side. The demand for claims by banks is downward sloping since, for a given level of bank net worth (and hence funding), higher prices imply that banks can buy fewer claims. Increases in bank net worth, in turn, shift upward the demand curve for claims. In this way, the demand side of the market for productive claims is reminiscent of the “cash-in-the-market” mechanism in Freeman (1996).

The supply of claims also depends on the market price for claims. If claims sell at a high price, then investment is more profitable; firms that have an opportunity to invest, invest more; and hence there are more claims to be sold in the market. The equilibrium price of claims clears the market every period.

When bank net worth is not too high and the agency problem is operational, the marginal return from buying an extra productive claim is higher than the cost of the extra deposits necessary to fund that claim. Financial intermediaries are credit-constrained.

Since, in the deposit market, banks do not yet know if they are on an island where firms will have an opportunity to invest, they all raise the same amount of deposits. After the investment opportunities realize, however, banks can interact in an interbank market. Gertler and Kiyotaki consider the case when the same agency frictions that apply to deposits also apply to borrowings in the interbank market and the case when the frictions in the interbank market are less intense or not existent at all.

If the interbank market is frictionless, then the economy functions as if banks would not face any idiosyncratic liquidity shocks. Banks in

³¹ There are no frictions in the relationship between banks and firms. Banks, instead of making loans to firms, buy claims on the future cash flow associated to capital investment.

investing islands borrow from banks in noninvesting islands. It is still the case that aggregate bank lending is constrained by aggregate bank net worth (the agency problem limits deposit funding), but there are no extra inefficiencies that arise from the combination of market segmentation and idiosyncratic shocks. In particular, the price of productive claims is the same in all islands.

When the interbank market is subject to frictions, on the other hand, the financial intermediation system cannot fully circumvent market segmentation via interbank trading. The supply of productive claims is higher in investing islands and, in consequence, the price of those claims is lower. This means that it is more profitable to acquire a claim in an investing island. But, funds from noninvesting islands may not flow to the investing islands if there is not enough net worth to support the extra funding (without compromising incentives). In equilibrium, then, the price of claims in investing islands is relatively low and, as a result, total investment is inefficiently low. The financial friction affects the real allocation of resources in the economy and its general macroeconomic performance.

Interbank borrowing impacts the incentive problem faced by banks in the same way that deposits do. Increasing interbank borrowing does not help increase the total amount of funds available to banks. Given borrowing constraints that are binding, more interbank borrowing has to be compensated with a decrease in deposits to keep bank leverage consistent with incentives.

Gertler and Kiyotaki discuss possible interventions that could be used to address the effects of interbank-market frictions on the real allocation of resources. They consider three types of policies: direct lending to firms, a discount window, and bank-equity injections. For this article, their discussion of the discount window is most germane.

If the central bank does not have an advantage over the private sector in its ability to enforce repayment, then discount window lending cannot improve outcomes. If, instead, the central bank has an enforcement advantage, then lending to banks in investing islands can increase total investment and improve economic outcomes in the economy. In fact, this enforcement advantage can make the discount window so attractive as to completely displace any private interbank trading. If both the discount window and the interbank market are to remain active, then the central bank needs to charge a penalty rate for loans at the discount window.

The discount window enforcement advantage makes it also a better alternative to deposits and it could displace deposits as a source of funding for banks, as well. To rule out this rather extreme (and unrealistic) outcome, Gertler and Kiyotaki impose a limit on the ability of the

central bank to efficiently evaluate borrowers and enforce repayment at the discount window. This limit translates formally into a capacity constraint that implies that the enforcement advantage of the central bank applies only up to a given maximum amount of discount window lending.

The financial frictions in the model are operative at all times, and the central bank intervention, to the extent that it can improve the situation, can do so also at all times. Gertler and Kiyotaki are particularly interested, however, in understanding the specific ways in which financial frictions may impact macroeconomic outcomes during crises. The way they model crises is by introducing a shock to the quality of capital. In effect, a negative shock to capital quality reduces the value of intermediaries' asset portfolios. Leverage amplifies the initial effect, reducing significantly the demand for productive claims in the economy. The fall in demand drives prices of claims down (a form of fire sales that arises due to the model's cash-in-the-market pricing), which feeds back into the balance sheet of banks, reducing their ability to finance capital even further. Moreover, the drop in current profits reduces the accumulation of bank net worth and tends to create more protracted crisis-like episodes. In other words, financial frictions can amplify and propagate the underlying shocks that drive crises in the model.

Using a quantitative example, Gertler and Kiyotaki discuss how policy can reduce the impact of a crisis shock in the economy. While they only consider direct lending as a policy response, they contend that discount window lending would have a similar power to dampen the macroeconomic implications of those shocks.

Adverse Selection

Philippon and Skreta (2012) study a model of financial contracting in the spirit of Myers and Majluf (1984), where private information and adverse selection generate suboptimally low levels of investment. The model has a large number of firms that need funding for a productive investment project. There is also a large set of risk-neutral investors with deep pockets. Firms also own "legacy" assets of different quality that influence the ability of firms to repay debt in the future. The quality of the legacy assets is private information, generating a distribution of different levels of repayment risks across firms.

In the absence of intervention, the interest rate on loans in the market reflects the average repayment risk of the set of firms asking for loans. Firms with low repayment risk end up facing a less attractive deal in the market and hence find investment less beneficial. Firms that decide not to invest do not seek funding in the market. In equilibrium,

only those firms with repayment risk above an endogenous threshold will undertake the investment projects and be active in the credit market. This is the case even though all firms' investment projects have a positive expected net cash flow. In other words, under perfect information, it would be optimal that they all invest.

Philippon and Skreta study optimal government interventions in this setup using a mechanism design approach. Interventions are optimal if they achieve a level of investment at minimum cost for the government. They show that, to increase investment and move the economy closer to efficiency, the government needs to make direct loans to firms at a lower rate than the one prevailing in a *laissez-faire* situation. The government program attracts firms with relatively low probability of repayment. As a result, the composition of the pool borrowing from private investors improves, allowing the private market interest rate to be lower and making the program consistent with an active private credit market.

The government lending program can be considered a version of the discount window. Because in some equilibrium situations there is selection in the participation decision of firms, with firms borrowing from the discount window having high repayment risk, the model can produce (equilibrium) discount window stigma (Courtois and Ennis 2010).³²

An important contribution of Philippon and Skreta is to show that in their framework direct lending is the best way to design a government-intervention program – in the sense that it minimizes the cost of the intervention for a given level of targeted investment. In this way, the paper provides strong support for the idea that, in certain situations, using the discount window to make low-interest-rate loans to firms (banks) can enhance efficiency in the economy, particularly in periods when adverse selection seems to be the main friction thwarting the appropriate functioning of private credit markets.³³

³² Ennis (2017) studies in detail the implications for discount window stigma of the Philippon-Skreta model.

³³ In a very recent paper, Gorton and Ordoñez (2016) also study an economy with private information where a discount window can have an efficiency enhancing role. Interestingly, stigma plays a role in Gorton and Ordoñez's model as well, but instead of hampering the ability of the central bank to provide appropriate liquidity to banks, stigma gives incentives to banks not to reveal their borrowing activities and, in this way, increases the effectiveness of the government program.

Over-the-Counter Trading in the Interbank Market

Ashcraft and Duffie (2007) document that the intraday allocation and pricing of funds in the U.S. interbank market tend to reflect the decentralized nature of transactions in that market. Furthermore, their stylized facts are consistent with the predictions coming out from search-based theories of over-the-counter (OTC) financial markets, which have recently received significant attention in the literature (Duffie, Gârleanu and Pedersen 2005).

Afonso and Lagos (2015) study intraday interbank OTC borrowing and lending with a focus on fund intermediation – that is, situations where a bank borrows from another bank in anticipation of lending those funds to yet another bank during the same trading session. They compare the implications of the model with various indicators of activity in the U.S. federal funds market and conclude that the model does a good job of capturing those features. The discount window plays a relatively passive role in Afonso and Lagos’ model with banks tapping the window at the end of the trading session if their balances are below a required value.

Ennis and Weinberg (2013) also study a model with bilateral bargaining and search frictions in the interbank market.³⁴ In the model, though, banks only get one chance to interact in the OTC interbank market, and hence no intermediation of the type highlighted by Afonso and Lagos takes place in equilibrium. Ennis and Weinberg (2013) focus on the issue of stigma at the discount window. They assume that banks can transact frictionlessly with the central bank in a way that is reminiscent of Williamson’s (1998) assumption that the central bank (and only the central bank) can trade in all locations and, in that way, circumvent the assumed market segmentation.

In the Ennis-Weinberg model, banks own assets of heterogeneous quality that determine their loan-repayment risk. In effect, banks sell assets to investors in order to repay interbank loans. An investor may not be able to observe the quality of the asset held by banks but can use information on the activities of banks in the interbank market to try to infer the quality of assets. When a bank with a low-quality asset is trying to borrow in the interbank market, it may not be able to obtain a loan if its counterparty can evaluate the asset and determine that it is

³⁴ See Bech and Klee (2011) and Acharya, Gromb, and Yorulmazer (2012) for two other models of the interbank market where bargaining plays an important role. Acharya, Gromb, and Yorulmazer highlight how the discount window influences the outside option of the borrowing side of the bargaining game and, in that way, the outcome of the negotiations.

low quality. Banks that do not obtain funding in the interbank market may access the discount window. Under some conditions, banks with low-quality assets are more likely to be in such a situation. As a result, the pool of banks borrowing at the window is biased toward banks with low-quality assets, and borrowing at the discount window becomes an endogenous negative signal of the quality of assets held by banks. In equilibrium, some banks can become “reluctant” to borrow from the discount window and may prefer to borrow from the interbank market at interest rates higher than the discount window rate just to avoid being stigmatized in the asset market.

Ennis and Weinberg are mainly concerned with the positive implications of the model and particularly with respect to discount window stigma. There is less work done on the normative aspects of discount window lending in this type of model. However, there is a very active literature addressing the general issues related to OTC trading in financial markets. The lessons for discount window policy that could come out from that body of work have not yet been fully developed, but based on some recent contributions such as, for example, Lagos, Rocheteau, and Weill (2011), it seems to be a promising avenue for further research.

2. CONCLUSIONS

In this essay, I have reviewed a strand of the economic literature dedicated to gaining a better understanding of the role of the discount window as the instrument of a LLR policy. My main focus has been on general equilibrium rationalizations of the policy using explicit, formal economic models. While this covers an important part of the existing literature, it is by no means comprehensive – I covered only papers where the discount window was explicitly discussed. In general, the discussion in this theoretical literature is held at a relatively abstract level, relying on simplified formal descriptions of financial interactions, without capturing the peculiarities so often present in practice.

There is a parallel literature discussing more practical considerations related to discount window policy without resorting to formal economic models for framing the main arguments (see, for example, Carlson, Duygan-Bump, and Nelson [2015] and the classic “little” book by Friedman [1960]). The relevance of this more applied literature can be easily recognized. I contend that, even for the practitioner, there are valuable insights emerging from the more theoretical literature described in this article. It seems likely that familiarity with this theoretical literature is also much less common in policy circles. By minimizing the focus on technical issues, one objective of this essay was to try

and bring down barriers between the practice and the theory behind central-bank liquidity provision.

Formal arguments, if well-structured, are either complete explanations of ideas or explicit about the areas of incompleteness (that is, where ad-hoc assumptions are being employed for lack of a good explanation or just as a shortcut). By reviewing the formal models available in the literature, one is able to get a better sense of the issues that are well-understood and the issues that are still largely unexplained or plainly unexplored.

To close the article, let me provide a brief summary of the main ideas addressed in the existing models. At their core, the models need to formalize a concept of liquidity. Different models do this in different ways. In many cases, agents (households and/or firms) in the economy confront an urgent need to access extra resources. The Diamond and Dybvig model is, of course, a canonical example of this approach. Other examples include the case when firms need interim extra funding to continue running their project, or when a subset of agents in the economy are moving to a different location and only some commodities (assets or goods) are transportable.

Idiosyncratic shocks across agents often motivate the formation of banks that act as coalitions to pool the risk associated with those shocks. In other models, banks are useful just because they have a technological advantage (in monetary loans, for example) relative to individual investors. Some models have no explicit institution resembling a bank – individual agents directly interact with the discount window.

In general, to have the discount window playing a valuable role in an economy with banks, it must be the case that those banks are organized in a way that keeps them exposed to residual uncertainty, even after pooling individual agents' exposures, and, furthermore, that there are barriers impeding the reallocation of resources through markets. In many cases, the limits to market functioning originate in segmentation and the resulting impossibility for certain agents to engage in potentially beneficial trade. In other cases, private information or limited commitment undermines some agents' capacity to trade. Legal and institutional constraints also play a role in some of the models.

The combination of bank-level liquidity shocks and market frictions creates liquidity shortfalls that result in a misallocation of resources. Sometimes the misallocation has to do with uneven consumption across households, and in other cases it is caused by the early liquidation of productive investment. Another source of misallocation is the possibility of having positive net present value projects that go unexploited.

It is interesting to highlight that, in some of the models, liquidity rationing results from the fact that the price of the liquid assets is

pinned down by a different set of factors than those associated with liquidity demand. So, for example, if the price of liquid assets is given by the future discounted value of its associated cash flows, but those liquid assets are needed in an interim period for liquidity purposes, then rationing and “scarcity” may happen. A similar situation arises when money simultaneously plays the role of the available liquid asset, aside from its usual role as a store of value. Holmstrom and Tirole’s (1998) and Freeman’s (1996) models, respectively, are good illustrations of this general but rather subtle idea.

Market frictions are often not enough to make discount window interventions beneficial – in many of the models, the government also has an advantage over private agents in its ability to overcome physical impediments to trade (spatial or otherwise) or in its ability to tax individuals in the future. Indeed, in some of the models, the government (via the discount window or otherwise) can generate the needed extra liquidity by issuing claims on future taxes and committing to fulfill them in the future. In other cases, the discount window is simply assumed to be better able (than private agents) to redistribute liquidity across agents at a given point in time due to, for example, its relative ubiquity.

To counterbalance the advantage attributed in the models to the discount window as a channel to allocate liquidity, some of the models contemplate the threat of moral hazard that comes with interventions and the provision of liquidity insurance by the central bank. While the moral hazard implications of central-bank lending are well-recognized in policy circles, the subject is (perhaps surprisingly) not very thoroughly studied in the more formal and technical literature reviewed here.

As should be clear even from this brief closing summary of the main ideas, there are a lot of elements that need to be present to create the conditions for the discount window to be a valuable institution in an economy. For this reason, in general, the models so far developed are relatively abstract and, at the same time, complex. Despite that, my contention was that many practical insights can come out from a detailed study of those models. This review hopefully serves as a concise introduction and potentially a useful guide to those interested in pursuing such an undertaking.

REFERENCES

- Acharya, Viral V., Denis Gromb, and Tanju Yorulmazer. 2012. “Imperfect Competition in the Interbank Market for Liquidity as a Rationale for Central Banking.” *American Economic Journal: Macroeconomics* 4 (April): 184–217.
- Acharya, Viral V., and Tanju Yorulmazer. 2008. “Cash-in-the-Market Pricing and Optimal Resolution of Bank Failures.” *Review of Financial Studies* 21 (November): 2705–42.
- Afonso, Gara, and Ricardo Lagos. 2015. “Trade Dynamics in the Market for Federal Funds.” *Econometrica* 83 (January): 263–313.
- Allen, Franklin, Elena Carletti, and Douglas Gale. 2009. “Interbank Market Liquidity and Central Bank Intervention.” *Journal of Monetary Economics* 56 (July): 639–52.
- Allen, Franklin, and Douglas Gale. 2000. “Financial Contagion.” *Journal of Political Economy* 108 (February): 1–33.
- Antinolfi, Gaetano, Elisabeth Huybens, and Todd Keister. 2001. “Monetary Stability and Liquidity Crises: The Role of the Lender of Last Resort.” *Journal of Economic Theory* 99 (July): 187–219.
- Antinolfi, Gaetano, and Todd Keister. 2006. “Discount Window Policy, Banking Crises, and Indeterminacy of Equilibrium.” *Macroeconomic Dynamics* 10 (February): 1–19.
- Ashcraft, Adam B., and Darrell Duffie. 2007. “Systemic Illiquidity in the Federal Funds Market.” *American Economic Review Papers and Proceedings* 97 (May): 221–5.
- Bech, Morten L., and Elizabeth Klee. 2011. “The Mechanics of a Graceful Exit: Interest on Reserves and Segmentation in the Federal Funds Market.” *Journal of Monetary Economics* 58 (July): 415–31.
- Bhattacharya, Sudipto, and Douglas Gale. 1987. “Preference Shocks, Liquidity and Central Bank Policy.” In *New Approaches to Monetary Economics*, edited by W.A. Barnett and K. J. Singleton. Cambridge: Cambridge University Press, 69–88.
- Board of Governors of the Federal Reserve System. 1968. “Reappraisal of the Federal Reserve Discount Mechanism: Report of a System Committee.” Steering Committee for the Fundamental Reappraisal of the Discount Mechanism (July 15).

- Caballero, Ricardo J., and Arvind Krishnamurthy. 2008. "Collective Risk Management in a Flight to Quality Episode." *Journal of Finance* 63 (October): 2195–2230.
- Calomiris, Charles, and Charles Kahn. 1991. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review* 81 (June): 497–513.
- Carlson, Mark A., Burcu Duygan-Bump, and William R. Nelson. 2015. "Why Do We Need Both Liquidity Regulations and a Lender of Last Resort? A Perspective from Federal Reserve Lending During the 2007-09 US Financial Crisis." Bank for International Settlements Working Paper 493 (March).
- Castiglionesi, Fabio, and Wolf Wagner. 2012. "Turning Bagehot on his Head: Lending at Penalty Rates when Banks Can Become Insolvent." *Journal of Money, Credit and Banking* 44 (February): 201–19.
- Champ, Bruce, Bruce D. Smith, and Stephen D. Williamson. 1996. "Currency Elasticity and Banking Panics: Theory and Evidence." *Canadian Journal of Economics* 29 (November): 828–64.
- Courtois, Renee, and Huberto M. Ennis. 2010. "Is There Stigma Associated with Discount Window Borrowing?" Federal Reserve Bank of Richmond *Economic Brief* 10-05.
- Diamond, Douglas W., and Philip Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (June): 401–19.
- Diamond, Douglas W., and Raghuram G. Rajan. 2005. "Liquidity Shortages and Banking Crises." *Journal of Finance* 60 (April): 615–47.
- Duffie, Darrell. 2010. "Presidential Address: Asset Price Dynamics with Slow-Moving Capital." *Journal of Finance* 65 (August): 1237–67.
- Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen. 2005. "Over-the-Counter Markets." *Econometrica* 73 (November): 1815–47.
- Ennis, Huberto M. 2017. "Interventions in Markets with Adverse Selection: Implications for Discount Window Stigma." Federal Reserve Bank of Richmond Working Paper 17-01 (January).
- Ennis, Huberto M., and Todd Keister. 2008. "Understanding Monetary Policy Implementation." Federal Reserve Bank of Richmond *Economic Quarterly* 94 (Summer): 235–63.

- Ennis, Huberto M., and Todd Keister. 2010. "On the Fundamental Reasons for Bank Fragility." Federal Reserve Bank of Richmond *Economic Quarterly* 96 (First Quarter): 33–58.
- Ennis, Huberto M., and David A. Price. 2015. "Discount Window Lending: Policy Trade-offs and the 1985 BoNY Computer Failure." Federal Reserve Bank of Richmond *Economic Brief* 15-05.
- Ennis, Huberto M., and John A. Weinberg. 2013. "Over-the-Counter Loans, Adverse Selection, and Stigma in the Interbank Market." *Review of Economic Dynamics* 16 (October): 601–16.
- Ennis, Huberto M., and John A. Weinberg. 2016. "The Role of Central Bank Lending in the Conduct of Monetary Policy." Federal Reserve Bank of Richmond *Economic Brief* 16-12.
- Flannery, Mark J. 1996. "Financial Crises, Payment System Problems, and Discount Window Lending." *Journal of Money, Credit and Banking* 28 (November, part 2): 804–24.
- Freeman, Scott. 1996. "The Payments System, Liquidity, and Rediscounting." *American Economic Review* 86 (December): 1126–38.
- Freeman, Scott. 1999. "Rediscounting Under Aggregate Risk." *Journal of Monetary Economics* 43 (February): 197–216.
- Freixas, Xavier, Bruno M. Parigi, and Jean-Charles Rochet. 2000. "Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank." *Journal of Money, Credit and Banking* 32 (August, part 2): 611–38.
- Freixas, Xavier, Bruno M. Parigi, Jean-Charles Rochet. 2004. "The Lender of Last Resort: A Twenty-First Century Approach." *Journal of the European Economic Association* 2 (December): 1085–1115.
- Friedman, Milton. 1960. *A Program for Monetary Stability*. New York: Fordham University Press.
- Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." In *Handbook of Monetary Economics 3A*, edited by B.M. Friedman and M. Woodford. Amsterdam: Elsevier, 547–99.
- Goodfriend, Marvin, and Robert G. King. 1988. "Financial Deregulation, Monetary Policy, and Central Banking." Federal Reserve Bank of Richmond *Economic Review* 74 (May/June): 3–33.

- Goodfriend, Marvin, and Jeffrey M. Lacker. 1999. "Limited Commitment and Central Bank Lending." Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Fall): 1–27.
- Gorton, Gary, and Guillermo Ordoñez. 2016. "Fighting Crises." National Bureau of Economic Research Working Paper 22787 (October).
- Hart, Oliver, and John Moore. 1994. "A Theory of Debt Based on the Inalienability of Human Capital." *Quarterly Journal of Economics* 109 (November): 841–79.
- Holmstrom, Bengt R., and Jean Tirole. 1996. "Modeling Aggregate Liquidity." *American Economic Review Papers and Proceedings* 86 (May): 187–91.
- Holmstrom, Bengt R., and Jean Tirole. 1998. "Private and Public Supply of Liquidity." *Journal of Political Economy* 106 (February): 1–40.
- Kahn, Charles, and Joao Santos. 2005. "Allocating Bank Regulatory Powers: Lender of Last Resort, Deposit Insurance and Supervision." *European Economic Review* 49 (November): 2107–36.
- Keister, Todd. 2016. "Bailouts and Financial Fragility." *Review of Economic Studies* 83 (April): 704–36.
- Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (April): 211–48.
- Lacker, Jeffrey M. 2004. "Payment System Disruptions and the Federal Reserve Following September 11, 2001." *Journal of Monetary Economics* 51 (July): 935–65.
- Lagos, Ricardo, Guillaume Rocheteau, and Pierre-Olivier Weill. 2011. "Crises and Liquidity in Over-the-Counter Markets." *Journal of Economic Theory* 146 (November): 2169–205.
- Martin, Antoine. 2006. "Liquidity Provision vs. Deposit Insurance: Preventing Bank Panics without Moral Hazard." *Economic Theory* 28 (May): 197–211.
- Martin, Antoine. 2009. "Reconciling Bagehot and the Fed's Response to September 11." *Journal of Money, Credit and Banking* 41 (March/April): 397–415.
- Mills, David C. Jr. 2006. "Alternative Central Bank Credit Policies for Liquidity Provision in a Model of Payments." *Journal of Monetary Economics* 53 (October): 1593–1611.

- Morris, Stephen, and Hyun Song Shin. 2003. "Global Games: Theory and Applications." In *Advances in Economics and Econometrics. Proceedings of the Eighth World Congress of the Econometric Society*, edited by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge: Cambridge University Press, 56–114.
- Myers, Stewart C., and Nicholas S. Majluf. 1984. "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have." *Journal of Financial Economics* 13 (June): 187–221.
- Philippon, Thomas, and Vasiliki Skreta. 2012. "Optimal Interventions in Markets with Adverse Selection." *American Economic Review* 102 (February): 1–28.
- Ratnovski, Lev. 2009. "Bank Liquidity Regulation and the Lender of Last Resort." *Journal of Financial Intermediation* 18 (October): 541–58.
- Repullo, Rafael. 2000. "Who Should Act as Lender of Last Resort? An Incomplete Contracts Model." *Journal of Money, Credit and Banking* 32 (August): 580–605.
- Rochet, Jean-Charles, and Xavier Vives. 2004. "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association* 2 (December): 1116–47.
- Smith, Bruce D. 2002. "Monetary Policy, Banking Crises, and the Friedman Rule." *American Economic Review Papers and Proceedings* 92 (May): 128–34.
- Williamson, Stephen D. 1998. "Discount Window Lending and Deposit Insurance." *Review of Economic Dynamics* 1 (January): 246–75.