Working Paper Series

Optimal Liquidity Regulation with Shadow Banking

WP 15-12R

Borys Grochulski
Federal Reserve Bank of Richmond

Yuzhe Zhang
Texas A&M University

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Optimal liquidity regulation with shadow banking*

Borys Grochulski† Yuzhe Zhang‡

March 25, 2016
Working Paper No. 15-12R

Abstract

We study the impact of shadow banking on optimal liquidity regulation in a Diamond-Dybvig maturity mismatch environment. A pecuniary externality arising out of the banks' access to private retrade renders competitive equilibrium inefficient. Shadow banking provides an outside option for banks, which adds a new constraint in the mechanism design problem that determines the optimal allocation. A tax on illiquid assets and a subsidy to the liquid asset similar to the payment of interest on reserves (IOR) constitute an optimal liquidity regulation policy in this economy. During expansions, when the return on illiquid assets is high, the threat of investors flocking out to shadow banking pins down optimal policy rates. These rates do not respond to business cycle fluctuations as long as the economy stays out of recession. In recessions, when the return on illiquid assets is low, optimal liquidity regulation policy becomes sensitive to the business cycle: both policy rates are reduced, with deeper discounts given in deeper recessions. In addition, when high aggregate demand for liquidity is anticipated, the IOR rate is reduced and, unless the shadow banking constraint binds, the tax rate on illiquid assets is increased.

Keywords: maturity mismatch, shadow banking, private retrade, pecuniary externality, liquidity regulation, interest on reserves
JEL codes: G21, G23, E58

1 Introduction

Beginning in the 1980s and leading up to 2007, a shadow banking system developed as a venue for origination and funding of illiquid bank assets outside of the realm of the government bank regulatory framework.[1] By 2007, the shadow banking sector had become about as large

*The authors are grateful to Douglas Diamond, Huberto Ennis, Alan Moreira, Guillaume Plantin, Kieran Walsh, and Ariel Zetlin-Jones for their helpful comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.
†Federal Reserve Bank of Richmond, borys.grochulski@rich.frb.org.
‡Texas A&M University, zhangeager@tamu.edu.
[1] Pozsar et al. (2012) define shadow banking as intermediation of credit through a wide range of securitization and secured funding techniques such as asset-backed commercial paper (ABCP), asset-backed securities (ABS),

as the traditional, regulated banking sector.[2] The view that shadow banking was a key factor contributing to the recent financial crisis, and to liquidity problems in particular, is shared by many academics and policymakers.[3] Optimal policy for liquidity regulation of banks should recognize the threat of capital moving to the shadow banking sector potentially rendering regulation ineffective and leading to inefficiently low levels of investment in liquid assets. Yet, this issue is understudied in the literature on optimal bank regulation.

In this paper, we study shadow banking as a costly but unregulated alternative to traditional banking. The question we ask is how the presence of this alternative affects optimal liquidity regulation of banks. As our benchmark, we take the pecuniary-externality-based theory of optimal liquidity regulation of Farhi et al. (2009). We extend this theory by adding the possibility of banks escaping regulation by moving assets to the unregulated shadow banking sector. We show that shadow banking tames regulation: optimal interventions are more modest when the banks' cost of engaging in shadow banking is lower.

In addition to considering shadow banking, we extend the theory of optimal liquidity regulation in the following two ways. First, we study how optimal regulations change with business cycle conditions, and how they respond to an increase in aggregate demand for liquidity. We find that optimal liquidity regulation policy does not respond to business cycle fluctuations outside of recessions. In recessions, optimal policy is relaxed, with larger slack granted in deeper recessions. When high demand for liquidity is anticipated, policy is tightened. Second, we present a new set of tools for implementation of liquidity regulation policy. We show how optimal policy interventions can be implemented via a proportional subsidy to liquid assets similar to payment of interest on reserves (IOR) and a flat-rate tax on illiquid assets.

Our model builds on the classic maturity mismatch problem studied in Diamond and Dybvig (1983), Holmstrom and Tirole (1998), Allen and Gale (2004), Farhi et al. (2009), and Farhi and Tirole (2012), among others. There are three dates, 0, 1, 2, and a population of ex ante identical banks, each with initial resources e.[4] Banks have the opportunity to invest in a long-term project at date 1 and, subject to an idiosyncratic shock, also at date 2. Each bank maximizes its continuation value, which is strictly increasing and concave in the scale of the long-term project funded by the bank. The long-term projects are bank-specific, i.e., non-traded. The idiosyncratic risk each bank faces is that its long-term investment opportunity may close early, i.e., at date 1.

This structure gives banks Diamond-Dybvig preferences over the timing of the funding of their long-term investments. A bank whose opportunity remains open beyond date 1 loses nothing by postponing investment until date 2, i.e., such a bank is perfectly patient at date 1.

---

collateralized debt obligations (CDOs), and repurchase agreements (repos). The growth of shadow banking beginning in 1980 has been widely documented, see for example Greenwood and Scharfstein (2013).

[2]As measured by the value of outstanding liabilities, see Figure 1 in Adrian and Ashcraft (2012).

[3]Brunnermeier (2009), Gorton and Metrick (2010), Financial Stability Board (2012), Bernanke (2012).

[4]With all agents ex ante identical in our model, we abstract from leverage.

A bank whose idiosyncratic opportunity closes at date 1 becomes extremely impatient at that date. Similar to Holmstrom and Tirole (1998) and Farhi and Tirole (2012), our model cuts out the withdrawal behavior of depositors and instead focuses directly on the banks' maturity mismatch problem.

There are two assets banks can use at date 0 to transfer resources forward in time: a liquid asset that matures at date 1 and yields a gross return of 1 (cash or central bank reserves), and an illiquid asset that matures at date 2 and yields a gross return of $R > 1$ at that date (bank loans, e.g., mortgages). After banks find out their patience type at date 1, a competitive market opens in which the illiquid asset is traded for date-1 resources (cash) at price p.[5] The illiquid asset, therefore, is completely illiquid technologically (i.e., cannot be physically turned into date-1 resources) but the presence of a market in which it can be sold gives it a degree of market liquidity. How liquid the asset is in the market sense depends on the equilibrium level of p. If $p < R$, the asset is not perfectly liquid, as it trades at date 1 at a liquidity discount. Diamond and Dybvig (1983) show that $p = 1 < R$ in a unique laissez-faire equilibrium.

Following Lorenzoni (2008) and Farhi et al. (2009), we assume that the date-1 market for the illiquid asset is private/anonymous, i.e., it cannot be interfered with by a regulator.[6] In the mechanism design problem defining the optimal choice of liquid and illiquid investment at date 0 and the optimal distribution of long-term investment between dates 1 and 2, the possibility of private retrade in this market makes the banks' incentive constraint tighter, as private retrade can enhance the banks' value of misrepresenting their patience type. Moreover, the value of this misrepresentation depends on the market price p, which in turn is determined by supply and demand in the retrade market. This dependence creates the so-called pecuniary externality: by taking p as given, an individual bank does not internalize the impact of its actions on the tightness of the incentive constraint faced by other banks. A planner solving the optimal mechanism design problem does internalize this impact. This discrepancy drives a wedge between the market and the planner's preferred allocation in this economy. Due to this wedge, the market allocation is inefficient, which gives rise to a role for regulation.[7]

With ex post retrade being private, the literature on pecuniary externalities studies regulations that are imposed on ex ante investment actions, which remain observable to the regulator.[8] We follow this approach herein and extend it in the following way.

We extend the theory of optimal bank regulation under pecuniary externalities by giving banks an outside option that allows them to escape regulations on ex ante investments. At

[5] In an equivalent formulation, banks could borrow against the illiquid asset instead of trading it at date 1. We assume assets are traded for the ease of exposition.

[6] One way to think about this assumption is that trading can be moved out of the jurisdiction (to an offshore location) and coordination of regulations across jurisdictions is impossible to achieve.

[7] As it is concerned with the efficiency of the overall allocation rather than with outcomes attained by any particular institution, the intervention we consider in this model falls into the category of macroprudential regulation.

[8] See Golosov and Tsyvinski (2007), Lorenzoni (2008), Farhi et al. (2009), Bianchi (2011), Di Tella (2014).

date 0, banks can choose to move resources to an unregulated shadow banking sector, where these resources can be invested in the liquid and the illiquid asset. The cost of this action is modeled as a fraction λ of resources invested in the illiquid asset. In the spirit of Jacklin (1987) and Kehoe and Levine (1993), shadow banks retain access to the private market for the illiquid asset at date 1.

The cost λ is a reduced-form way of modeling the shadow banks' lack of access to explicit (e.g., deposit insurance, discount window) and implicit (i.e., ex post support) government safety net that is available to formal banks.[9] Ceteris paribus, the lack of government backstops increases the shadow banks' cost of funding.[10] Since we abstract from leverage in this paper, we model the disadvantage of funding bank assets in the shadow sector directly as a markup λ on the cost of originating illiquid assets (loans) that shadow banks incur relative to formal banks. We characterize optimal liquidity regulations for any $\lambda \geq 0$.

The possibility of shadow banking adds a new constraint to the mechanism design problem defining (constrained-) optimal allocations in this economy: the ex ante value of remaining in the regulated banking sector must not be smaller than the value of becoming a shadow bank. Both these values depend on the equilibrium illiquid asset retrade price p, which in turn depends on the allocation itself. The mechanism-design problem is therefore nonstandard in that the agents' outside option value is endogenous to the mechanism.

To study this mechanism design problem, we first transform it into one in which the planner indirectly chooses the retrade price p while the rest of the allocation (i.e., the initial and final investment made by banks) is determined by the requirements of resource feasibility and incentive compatibility. It turns out that this transformation is particularly convenient for the analysis of the ex ante participation constraint. We show that the participation constraint boils down to an upper bound on the set of retrade prices p feasible for the planner. This upper bound is tighter when the resource cost of shadow banking, λ, is smaller.

Using the transformed mechanism design problem, we show that ex ante welfare attained in this economy (i.e., the banks' expected continuation value) is strictly increasing in the retrade price p, up to the first-best price $p^{\text{fb}} > 1$ that captures the (unconstrained-) optimal trade-off between liquidity insurance and the average return on investment. The objective of the planner, therefore, is to increase the retrade price p above the laissez-faire equilibrium price $p = 1$ as far toward the first-best price p^{fb} as possible without violating the banks' participation constraint.

If the cost of shadow banking λ is sufficiently high, the participation constraint does not bind and the constrained-optimal price p^* coincides with the first-best optimal price p^{fb}. If the

[9] Adrian and Ashcraft (2012) in fact define shadow banking as intermediation of credit outside of the government safety net.

[10] Using Fitch ratings data, Ueda and Weder di Mauro (2013) estimate the funding cost advantage of banks covered by the expectation of ex post government support at between 60 and 80 basis points. Explicit and priced safety net programs like deposit insurance and access to discount window can also reduce the banks' cost of funding if bank investors/depositors are risk averse.

cost of shadow banking is nil, the participation constraint renders all interventions infeasible, making the laissez-faire equilibrium price $p = 1$ constrained-optimal. In the intermediate range of the cost parameter λ, the constrained-optimal price p^* is below p^{fb} but above the laissez-faire equilibrium price $p = 1$. A scope for ex ante regulation therefore remains, but the magnitude of feasible policy interventions is restricted by the threat of shadow banking.

As in Farhi et al. (2009), the optimal policy intervention in our model can be implemented via a quantity restriction mandating a minimum proportion of liquid assets on the balance sheet of a bank at date 0. We review this result in Section 7. Our main focus, however, is on an alternative implementation of optimal liquidity regulation policy. We suppose that at date 0 the government/regulator imposes a proportional tax τ on origination of illiquid assets paired with a proportional subsidy i to investment in liquid assets. Such a subsidy is akin to payment of interest on liquid reserves held by banks at the central bank. We characterize the policy rate choices (i, τ) that implement the constrained-optimal allocation associated with the constrained-optimal retrade price p^*.

It is intuitive that a subsidy to investment in the liquid asset and a tax on origination of the illiquid asset tilt the asset allocation trade-off banks face ex ante in favor of the liquid asset. This tilt increases the supply of liquidity and decreases the supply of the illiquid asset in the retrade market at date 1, thus increasing the market-clearing price p. Both policy rates i and τ are uniquely determined by the optimal allocation. More precisely, they are determined by how much liquidity insurance the optimal allocation provides to banks. We show that the optimal IOR rate i is equal to the net return that impatient banks are able to earn at the optimal allocation.[11] This return is zero in the laissez-faire equilibrium. The tax rate τ is the corresponding discount (relative to the total return of R that patient banks earn in the laissez-faire equilibrium) in the total return that patient banks earn at the optimal allocation. Jointly, thus, i and τ implement an ex post transfer from patient to impatient banks, which from the ex ante point of view amounts to provision of liquidity insurance.

There is a substantial debate in policymaking circles on whether and how bank regulation policy should respond to the business cycle.[12] Our baseline model with no aggregate uncertainty is readily extended to allow for aggregate uncertainty that is resolved ex ante, i.e., before banks choose their liquid and illiquid initial investments. We study how the solution to the mechanism design problem and its implementation via the IOR rate i and a tax rate τ depend on the level of return on the illiquid asset R and, separately, on the fraction of impatient banks π. Since the return on bank assets is procyclical in U.S. data, we identify high R with times of economic expansion and low R with recessions.[13]

[11] This return is analogous to the first-period return on the deposit contract in Diamond and Dybvig (1983).

[12] See, for example, Financial Stability Forum (2009).

[13] Using FDIC data on all insured institutions, https://www.fdic.gov/bank/statistical/, it is easy to check that aggregate Return on Assets and Yield on Earning Assets are positively correlated with GDP growth. Detailed analysis of the data is available upon request.

We find that when the return rate R is high, the shadow banking constraint binds pinning down the constrained-optimal price p^* and, consequently, the optimal policy rates (i, τ). As long as this constraint continues to bind, i.e., for all R large enough, p^* and (i, τ) are independent of the particular realization of R. Therefore, optimal policy is insensitive to changes in R during economic expansions. When return R is low, i.e., in recessions, the banks' participation constraint does not bind and the constrained-optimal resale price matches the first-best optimal price, $p^* = p^{\text{fb}}$. The first-best price is sensitive to the realized return R. In particular, lower R implies lower p^{fb} and, thus, a deeper reduction in optimal policy rates (i, τ).

Finally, we consider how optimal policy rates respond to high anticipated liquidity demand measured by the fraction of impatient banks π. High demand for liquidity hinders the provision of liquidity insurance (the ex post transfer to an impatient bank becomes smaller), which implies that the IOR rate i is decreased. The optimal tax rate τ is increased to the extent consistent with not violating the banks' participation constraint.

Related literature We view this paper as connecting two literatures: the literature on optimal regulation with pecuniary externalities and the recent literature on shadow banking.

We extend the literature on pecuniary externalities by explicitly modeling the agents' ex ante outside option. In our case, this outside option represents shadow banking. The outside option imposes a new ex ante participation constraint in the mechanism design problem, which limits the strength of the pecuniary externality and reduces the scope for regulation.[14] Most directly, our analysis extends Farhi et al. (2009), which studies the impact of the retrade-based pecuniary externality on optimal liquidity regulation of banks without the possibility of shadow banking. Similarly, Stein (2012) studies optimal regulation of bank issuance of money-like riskless debt in a pecuniary-externality environment with a collateral constraint on banks and inefficient ex post fire sales of long-term assets into a market with downward-sloping demand. In contrast to our paper, he does not consider shadow banking as an option allowing banks to escape regulation. Other related papers, e.g., Golosov and Tsyvinski (2007), Lorenzoni (2008), Bianchi (2011), Di Tella (2014), study market failures due to pecuniary externalities in other applications. We conjecture that our result showing that ex ante outside options limit the strength of pecuniary externalities can be extended to these applications.

Several recent studies build positive models of shadow banking: Huang (2014), Moreira and Savov (2014), and Ordonez (2015). These papers study shadow banking as an unregulated, off-balance-sheet asset funding vehicle available to banks. Although these studies consider several interesting government interventions, their objective is primarily positive, so they do not characterize general, constrained-optimal regulation policy. Ordonez (2015), in particular, studies

[14] The ex ante participation constraint we study is different from the ex post participation constraint studied in Kehoe and Levine (1993). Our constraint limits the size of the pecuniary externality arising from private ex post retrade. In Kehoe and Levine (1993), the ex post participation constraint combined with spot labor markets gives rise to a pecuniary externality in the first place.

shadow banking as a value-enhancing response of banks to regulatory constraints on risk-taking that are too tight. In this paper, our objective is to examine the impact of shadow banking on optimal liquidity regulation. We treat shadow banking as an off-equilibrium phenomenon and solve for general, constrained-optimal liquidity regulation policy.

Plantin (2014) studies optimal capital regulation policy in a model with shadow banking. The role for regulation comes from the banks' failure to internalize the real adjustment costs in the production sector caused by the volatility in the final goods demand, which increases with the riskiness of bank deposits. Our model studies optimal regulation of liquidity, instead of capital, and uses an environment in which the role for regulation comes from a different friction (the pecuniary externality). However, as does Plantin (2014), we model shadow banking as unmonitored, ex post spot-market trade between banks and nonbanks. He allows for adverse selection in this market and finds that it can be beneficial, as it disrupts unmonitored trade and limits the size of shadow banking.[15]

We assume no uncertainty about asset quality in the secondary market for illiquid assets. In doing so, we follow the large literature on interbank markets that dates back to Bhattacharia and Gale (1986) and includes, among others, Allen and Gale (2004), Allen et al. (2009), Freixas et al. (2011), and Gale and Yorulmazer (2013). In contrast to most studies in this literature, however, we do not assume contract or market incompleteness. Instead, we solve a general mechanism design problem with resource, incentive, private retrade, and ex ante participation constraints.

The paper is organized as follows. Section 2 presents the baseline model without aggregate uncertainty. Section 3 discusses competitive equilibrium without intervention. Section 4 defines and solves the mechanism design problem describing constrained-efficient allocations. Section 5 studies the implementation with IOR and a tax τ. Section 6 extends the results to the case of ex ante aggregate uncertainty. Section 7 considers the alternative implementation with minimum liquidity requirements. Section 8 concludes. All proofs are relegated to the Appendix.

2 The model

We consider an economy populated by a continuum of ex ante identical investors. The economy extends over three dates $t = 0, 1, 2$ and has two assets available for investment at date 0: a liquid, low-yield asset s (cash or central bank reserves) and a technologically illiquid, high-yield asset x (such as mortgages or mortgage-backed securities). In addition, investors have idiosyncratic, long-term investment opportunities available at date 1 and, in some cases, also at date 2. Investors' objective is to maximize the continuation value $V(I)$, where I is the

[15] House and Masatlioglu (2015) study the effectiveness of equity injections and asset purchases in an interbank market with adverse selection. Bengui et al. (2015) assume illiquid assets to be completely nontradable and study public provision of liquidity through ex post bailouts of banks covered by an explicit government guarantee.

amount of capital invested in the long-term project.[16]

Each investor is endowed with $e > 0$ units of capital. At date 0, investors must decide whether to become a bank or a nonbank (a shadow bank). If an investor decides to do business as a bank, he is subject to government bank regulation. If an investor decides to do business as a nonbank, he is free from bank regulation, but faces an extra cost $\lambda \geq 0$ of originating the illiquid asset x.[17] In equilibrium, becoming a bank will make investors at least as well off as becoming a nonbank. In case of indifference, we assume investors become banks.[18] The possibility of becoming a nonbank will therefore serve as an out-of-equilibrium outside option for investors restricting the amount of government regulation that can be imposed on banks in equilibrium.

After investors decide to become banks at date 0, they make their initial investments s_0 and x_0 in, respectively, the liquid cash asset s (central bank reserves) and the illiquid asset x (bank loans). The cash asset pays a riskless return of 1 at date 1, and nothing at date 2. The illiquid asset pays nothing at date 1 and a riskless return of $R > 1$ at date 2. For now, we assume that R is a fixed constant. In Section 6, we discuss how the solution of our model depends on the level of R. Note that as of date 1 asset x is technologically illiquid, i.e., capital invested in x cannot be used to fund long-term investment I at date 1.

The opportunity to invest in I is subject to an idiosyncratic Diamond-Dybvig shock $\theta \in \{0, 1\}$ realized at date 1. If $\theta = 1$, the bank can invest in I at either date 1 or 2. If $\theta = 0$, the bank has access to the long-term investment I only at date 1. The shock θ is a liquidity shock: banks of type $\theta = 0$ need liquidity at date 1 in order to invest in I before the specific long-term investment opportunity available to them closes. We will refer to them as impatient banks. Banks with $\theta = 1$ will be called patient. We use these assumptions to model the classic maturity-mismatch problem. The illiquid investment x, on the one hand, produces excess return $R > 1$, but, on the other hand, exposes the bank to liquidity risk at date 1.

After banks find out their type θ, but before the opportunity to invest in I at date 1 closes, a competitive market for the illiquid asset opens, where banks can trade the illiquid asset x for cash at a market price p . The existence of this market allows impatient banks to avoid getting

[16]The scale of the long-term project is not modeled directly. If investors leverage their capital investment I and run the long-term project on a proportionally larger scale, the value $V(I)$ implicitly encompasses the costs and benefits of such actions.

[17]As discussed in the introduction, this cost is a reduced-form way of modeling the shadow banks' lack of access to the government safety net. We assume that acquiring liquid assets (cash) does not carry an extra cost for a nonbank. Such a cost would not make a difference in our analysis because, as shown in Section 4.3.2, nonbanks strictly prefer to hold only the illiquid asset and no cash in their portfolios, even without the cost of acquiring cash.

[18]This assumption is without loss of generality. Suppose some investors become nonbanks while achieving the same expected payoff as banks. This allocation can be replaced by another allocation in which all investors become banks and continue to achieve the same payoff. It is not hard to show that the new allocation could be made incentive-feasible, see Definition 1, because shadow banking is a weakly inefficient way of originating illiquid bank assets due to the deadweight cost $\lambda \geq 0$.

stuck with illiquid assets at date 1, as their holdings of x can always be sold, at the market price p. This price, however, can in equilibrium be lower than the asset's face value R.

In sum, given that the outside option of becoming a nonbank is not more attractive, all investors choose to become banks at date 0 and then choose a portfolio $(s_0, x_0) \in \mathbb{R}^2_+$ subject to

$$s_0 + x_0 + M(s_0, x_0) \leq e, \tag{1}$$

where s_0 is the amount invested in the cash asset, x_0 is the amount invested in the illiquid asset, and M represents the banks' costs of government regulations.[19] At date 1, after they find out their type θ, banks choose their net demands $n(\theta) \geq -x_0$ in the market for the illiquid asset, and the final long-term investments $I_1(\theta) \geq 0$ and $I_2(\theta) \geq 0$ subject to budget constraints

$$I_1(\theta) \leq s_0 - pn(\theta), \tag{2}$$

$$I_2(\theta) \leq (x_0 + n(\theta))R. \tag{3}$$

As evident from these budget constraints, the government does not impose any regulations/taxes on the secondary market for the illiquid asset. Following Lorenzoni (2008), Farhi et al. (2009), and others, we assume, here as well as in the mechanism design problem of the planner/government, that the secondary market for the long-term asset is outside of the reach of government regulation.[20]

The objective of a bank is to maximize the continuation value from the long-term investment I. For a bank of type θ who invests I_1 at date 1 and I_2 at date 2, the total continuation value is $V(I_1 + \theta I_2)$. The value function V is strictly increasing, strictly concave, and satisfies the following assumption.

Assumption 1 (Enough concavity) *$V'(I)I$ is strictly decreasing in I for all $I \geq e$.*

The ex ante expected payoff of a bank, therefore, is

$$\mathbb{E}[V(I_1 + \theta I_2)] \equiv \pi V(I_1(0)) + (1 - \pi)V(I_1(1) + I_2(1)), \tag{4}$$

where $0 < \pi < 1$ is the probability of $\theta = 0$. The bank's problem is to maximize (4) subject to budget constraints (1), (2), and (3).

If an investor decides to become a nonbank at date 0, his budget constraints are

$$s_0 + (1 + \lambda)x_0 \leq e, \tag{5}$$

[19] Note that these costs can be negative, e.g., when the government pays interest on reserves to the banks.

[20] The reason why the government cannot interfere with the secondary market could be anonymity of trade (i.e., any trades that banks and nonbanks execute in this market are not observable to the government). More generally, monitoring of transactions in this market may be very costly given the possibility of these transactions being moved to a different legal jurisdiction.

at date 0, and (2),(3) at date 1. Denote by $\tilde{V}_0(p, \lambda)$ the value of becoming a nonbank, i.e., the maximum of (4), subject to (5), (2), and (3).

We focus on the market provision of liquidity and financial firms' liquidity management, and not on other aspects of banking (leverage, fragility to runs, etc). Our banks face the trade-off between liquidity and return. Unless they invest 100 percent in the low-yield cash asset, they face a maturity mismatch problem in that their long-term investment opportunity I can close before the illiquid asset x matures. The secondary market for illiquid assets lets banks access liquidity, but only at a cost (because $p < R$).

3 Competitive equilibrium

In this section, we discuss competitive equilibrium in the laissez-faire (LF) economy, that is, the unregulated economy with $M = 0$. Since $\lambda \geq 0$, it is immediate that with $M = 0$ all investors prefer, at least weakly, to become banks. Diamond and Dybvig (1983) show that there exists a unique equilibrium in this LF economy.

Formally, banks' choices (s_0, x_0), $(n(\theta), I_1(\theta), I_2(\theta))$, and a price p are a competitive equilibrium if, taking p as given, the choices solve the banks' maximization problem, and the date-1 market for the illiquid asset x clears, i.e.,

$$\mathbb{E}[n] \equiv \pi n(0) + (1 - \pi) n(1) = 0. \tag{6}$$

Theorem 1 *In the economy with $M = 0$, there exists a unique equilibrium: $p = 1$, $(s_0, x_0) = (\pi e, (1 - \pi)e)$, $n(0) = -x_0$, $n(1) = s_0$, $I_1(0) = e$, $I_1(1) = 0$, $I_2(0) = 0$, $I_2(1) = Re$.*

The equilibrium price $p = 1$ is pinned down by an arbitrage-type argument comparing the liquid asset investment and a one-period investment in the illiquid asset at date 0. If p is not 1, one of these two investments dominates the other, and so the optimal investment choice at date 0 is not an interior one, which is inconsistent with clearing of the secondary market for the illiquid asset at date 1.

As shown in Jacklin (1987), this argument does not depend on the simple market structure we consider here (spot market at date 1 only). This argument carries over to a general market structure in which banks are allowed to enter and trade all conceivable state-contingent contracts. The intuition here is that it is not possible to contract around an arbitrage opportunity. Thus, $p = 1$ in any, even generally defined, competitive equilibrium.[21]

Note that in the LF equilibrium, the illiquid asset trades at date 1 at a discount relative to its fundamental value R. Due to inelastic supply of the illiquid asset (by the impatient types), a riskless payoff of R is bought by agents who are indifferent between resources at date 1 and

[21] See Section 3.1 in Farhi et al. (2009) for a formal proof of this result.

2 (the patient types) at an equilibrium price $p = 1 < R$. The difference between R and 1, thus, represents a liquidity discount at which the riskless one-period bond sells in equilibrium. Next, we ask if this discount is efficient. We define a general mechanism design problem, derive the constrained-optimal price p^*, and show that this price is less than R but more than 1, i.e., due to pecuniary externality, the equilibrium liquidity discount is too large and, therefore, the unregulated LF equilibrium allocation is inefficient.

4 Constrained-optimal allocations

In order to study the extent of pecuniary externality and the scope for government regulation in this economy, we characterize in this section constrained-optimal allocations. In the first subsection, we define the optimal mechanism design problem. In the second subsection, we reduce this problem to a simple, single-dimensional maximization problem. In the third subsection, we characterize its solution. In the fourth subsection, we comment on the pecuniary externality that exists in this environment with private retrade.

4.1 Mechanism design problem

In this economy, an allocation is $A = (s_0, x_0, s_1(\theta), x_1(\theta), n(\theta), I_1(\theta), I_2(\theta))$, where (s_0, x_0) are the amounts invested at date 0 in the two assets, $(s_1(\theta), x_1(\theta))$ is a state-contingent asset allocation at date 1, and $(n(\theta), I_1(\theta), I_2(\theta))$ are recommendations for the private actions that agents/banks are to take : n is the recommended trade in the private market for the illiquid asset, and I_t is recommended investment in the long-term projects at date t.

Definition 1 *Allocation A is incentive-feasible (IF) if*

(i) it is incentive compatible (IC), i.e., if there exists a price $p \geq 0$ such that a) for both θ

$$(\theta, n(\theta), I_1(\theta), I_2(\theta)) \in \arg\max_{\tilde{\theta}, \tilde{n}, \tilde{I}_1, \tilde{I}_2} \quad V(\tilde{I}_1 + \theta \tilde{I}_2) \tag{7}$$

$$s.t. \quad \tilde{I}_1 \leq s_1(\tilde{\theta}) - p\tilde{n}, \tag{8}$$

$$\tilde{I}_2 \leq (x_1(\tilde{\theta}) + \tilde{n})R, \tag{9}$$

and, b) the secondary market clears at p, i.e., $\mathbb{E}[n] = 0$.

(ii) it is resource feasible (RF), i.e.,

$$\begin{aligned} s_0 + x_0 &\leq e, \\ \pi s_1(0) + (1-\pi)s_1(1) &\leq s_0, \\ \pi x_1(0) + (1-\pi)x_1(1) &\leq x_0, \end{aligned}$$

(iii) it satisfies the ex ante participation constraint

$$\mathbb{E}[V(I_1 + \theta I_2)] \geq \tilde{V}_0(p, \lambda), \tag{10}$$

where

$$\begin{aligned} \tilde{V}_0(p,\lambda) \equiv \max_{\tilde{s}_0, \tilde{x}_0, n(\theta), \tilde{I}_1(\theta), \tilde{I}_2(\theta)} \quad & \mathbb{E}[V(\tilde{I}_1 + \theta \tilde{I}_2)] \\ s.t. \quad & \tilde{s}_0 + (1+\lambda)\tilde{x}_0 \leq e, \\ & \tilde{I}_1(\theta) \leq \tilde{s}_0 - p\tilde{n}(\theta), \\ & \tilde{I}_2(\theta) \leq (\tilde{x}_0 + \tilde{n}(\theta))R, \ \text{for } \theta = 0, 1. \end{aligned}$$

is the ex ante value of becoming a nonbank.

The incentive compatibility condition (7) requires that, taking the retrade price p as given, banks cannot improve their value by any joint deviation combining a misrepresentation of their type θ with trading in the private retrade market. This condition is the same as the notion of incentive compatibility with retrade used in Farhi et al. (2009) and other studies of pecuniary externalities.

The ex ante participation constraint (10) is new to this literature. In particular, because $\tilde{V}_0$ depends on p, the value of the banks' outside option is not taken parametrically in our model but rather is endogenous to the mechanism.

Social welfare is given by the ex ante expected value delivered to the representative bank:

$$\mathbb{E}[V(I_1 + \theta I_2)]. \tag{11}$$

The mechanism design problem is to find an IF allocation A that maximizes this objective among all IF allocations. Such an allocation will be referred to as a constrained-optimal allocation.

4.2 Reduction of the mechanism design problem

In this subsection, we show that the general mechanism design problem defined above can be reduced to one in which the planner chooses, indirectly, just the secondary-market price for the illiquid asset, p, and the rest of the allocation A is determined by the requirements of incentive-feasibility.

Lemma 1 *It is without loss of generality to restrict attention to allocations in which the planner recommends no trade in the private market, i.e., $n(\theta) = 0$ for $\theta = 0, 1$.*

This result is analogous to Cole and Kocherlakota (2001). It follows because using the state-contingent allocation of assets at date 1, $(s_1(\theta), x_1(\theta))$, the planner can replicate any trades that banks may want to execute in the private market without affecting the retrade price p (therefore also preserving the banks' ex ante participation constraint).

Given this lemma, in the remainder of this section we focus on allocations with zero recommended retrade. With $n(\theta) = 0$, budget constraints (8) and (9) imply

$$I_1(\theta) = s_1(\theta), \quad I_2(\theta) = Rx_1(\theta).$$

Also, for a given state-contingent allocation of assets at date 1, $(s_1(\theta), x_1(\theta))_{\theta \in \{0,1\}}$, RF constraints determine the requisite initial investment (s_0, x_0). Thus, under Lemma 1, the full allocation A is determined by the date-1 state-contingent allocation of assets $(s_1(\theta), x_1(\theta))_{\theta \in \{0,1\}}$, with ex ante social welfare (11) simplified to $\mathbb{E}[V(s_1 + \theta R x_1)]$.

Proposition 1 *An allocation is incentive compatible with price $p > 0$ if and only if*

(i) $s_1(0) + px_1(0) = s_1(1) + px_1(1)$,

(ii) $x_1(0) = 0$,

(iii) $p \begin{cases} \geq R, & \text{if } s_1(1) > 0; \\ \leq R, & \text{if } x_1(1) > 0. \end{cases}$

Condition (i) shows that, as in Allen (1985) and Cole and Kocherlakota (2001), incentive compatibility with retrade implies that the market value of assets allocated at date 1 to those who announce θ, i.e., $s_1(\theta) + px_1(\theta)$, must be the same for both announced types. Otherwise, all banks would report the realization of θ that receives the asset allocation $(s_1(\theta), x_1(\theta))$ with the higher market value. When the value of assets allocated to each announcement is the same, banks have no reason to misrepresent their type.

Condition (ii) is necessary for the recommendation of no private trade to be incentive compatible for the impatient types. Clearly, the impatient banks must receive no illiquid asset at date 1, for otherwise they would prefer to trade in the private market (sell $x_1(0)$ at any price).

Likewise, condition (iii) is necessary to guarantee that the patient types do not go to the private market. Indeed, if the allocation gives patient banks some liquid asset at date 1 and the banks are supposed to not trade in the private market, the return from buying the illiquid asset in that market must be weakly negative. If the allocation gives them a positive allocation of the illiquid asset, the return from selling it in the private market must be weakly negative.

Lemma 2 *At a constrained-optimal allocation a) the RF constraints hold as equality, b)* $s_1(1) = 0$*, and c)* $p \leq R$*.*

This lemma gathers three immediate necessary conditions for efficiency. In particular, condition b) says providing liquidity to the type that does not need it is not optimal, as liquidity is costly to provide. Condition c) says that since the patient types are to postpone their final investment to date 2, they need to earn a nonnegative return between dates 1 and 2.

4.2.1 Reduced mechanism design problem

Proposition 1 and Lemma 2 imply that we can further focus our analysis on a simple class of allocation mechanisms in which the planner indirectly chooses just the retrade price p, while the asset allocation $(s_1(\theta), x_1(\theta))_{\theta\in\{0,1\}}$, and thus the whole allocation A as well, is determined by the requirements of incentive-feasibility.

To see this, note first that with $x_1(0) = s_1(1) = 0$, the present value condition (i) in Proposition 1 reduces to $s_1(0) = px_1(1)$. This shows that for given $s_1(0)$ and $x_1(1)$, there exists a unique retrade price consistent with incentive compatibility, i.e., the price at which $s_1(0)$ and $x_1(1)$ have the same market value:

$$p = s_1(0)/x_1(1). \tag{12}$$

Second, with parts a) and b) of Lemma 2, the resource feasibility conditions imply

$$\pi s_1(0) + (1-\pi)x_1(1) = e. \tag{13}$$

Third, part c) of Lemma 2 requires $p \leq R$.

The mechanism design problem, thus, boils down to the choice of $s_1(0)$ and $x_1(1)$ subject to (12), (13), $p \leq R$, and the investors' ex ante participation constraint (10). Equivalently, we can think of the planner as choosing a price $p \leq R$ with $s_1(0)$ and $x_1(1)$ determined by (12) and (13).

The social welfare function that the planner maximizes (the banks' value) can be conveniently expressed in terms of just the retrade price p.

Lemma 3 *In the mechanism in which the planner indirectly chooses the retrade price $p \leq R$, the social welfare function is*

$$\mathbb{E}\left[V\left(\frac{p}{\pi p + 1 - \pi}e\left(1 - \theta + \theta\frac{R}{p}\right)\right)\right]. \tag{14}$$

The objective (14) shows the trade-off involved in the setting of the retrade price p. At date 1, all banks earn the return $\frac{p}{\pi p+1-\pi}$ on their initial investment e. Impatient banks invest $I_1 = \frac{p}{\pi p+1-\pi}e$ at that time. Patient banks are able to wait until the illiquid asset matures. They earn an additional return $\frac{R}{p}$ and invest $I_2 = I_1\frac{R}{p} = \frac{R}{\pi p+1-\pi}e$ at date 2. Higher p increases I_1, decreases I_2, and decreases the average return $\mathbb{E}[I_t] = \frac{p\pi+(1-\pi)R}{\pi p+1-\pi}$. The planner, therefore, faces a trade-off between return and insurance. Higher p provides more insurance at the cost of lower average return. By setting $p = R$, the planner can achieve full insurance with $I_1 = I_2 = \frac{R}{\pi R+1-\pi}e$, but the average return $\frac{R}{\pi R+1-\pi}$ is low. By setting the price $p = 0$, the return on e is maximal, R, (supported by $x_1(1) = e$ and $s_1(0) = 0$), but risk sharing is very poor, as $I_1 = 0$.

The indirect formulation of the mechanism design problem in which the planner chooses p is convenient because the banks' value of the outside option (the option of becoming a nonbank), $\tilde{V}_0(p, \lambda)$, depends on the allocation A (that is offered to banks) only through the illiquid asset resale price p associated with A.

Lemma 4 *In the mechanism in which the planner indirectly chooses the retrade price $p \leq R$, a nonbank's objective $\tilde{V}_0(p, \lambda)$ can be expressed as*

$$\mathbb{E}\left[V\left(\max\left\{1, \frac{p}{1+\lambda}\right\} e\left(1 - \theta + \theta\frac{R}{p}\right)\right)\right]. \tag{15}$$

We see in (15) that the structure of the payoff for a nonbank mirrors that of a bank, given in (14). All nonbanks earn the same initial return on their date-0 investment, and patient nonbanks earn the additional return $\frac{R}{p}$ by postponing their final investment to date 2. The initial return they earn, $\max\{1, \frac{p}{1+\lambda}\}$, is attained by investing all-cash at date 0 if $p < 1 + \lambda$ or putting all initial resources in the illiquid asset if $p > 1 + \lambda$. If $p = 1 + \lambda$, nonbanks are indifferent with respect to the asset allocation at date 0.

With these formulas for the payoffs to banks and nonbanks, we can express the banks' ex ante participation constraint in the following simple form.

Lemma 5 *In the mechanism in which the planner indirectly chooses the retrade price $p \leq R$, the participation constraint reduces to*

$$\frac{\lambda}{\pi} \geq p - 1 \geq 0. \tag{16}$$

The right inequality in (16) is the constraint imposed by the investor's option to become a nonbank and invest all-cash. The left inequality in (16) follows from the investor's option to become a nonbank and invest all-illiquid.

In sum, the mechanism design problem boils down to the choice of a retrade price $p \leq R$ that maximizes (14) subject to (16). This reduction of the full mechanism design problem leads to a simple characterization of all optimal allocations and their dependence on λ, which we provide in the next subsection.

4.3 Characterization of optima

4.3.1 Without the ex ante participation constraint

As a benchmark, let us first solve the planning problem with the RF and IC constraints but disregarding the ex ante participation constraint (16), i.e., as if banks did not have the option to become nonbanks.

Proposition 2 *There exists a unique maximizer p^{fb} for the objective function (14). Under Assumption 1, this maximizer satisfies*

$$1 < p^{\text{fb}} < R. \tag{17}$$

The price p^{fb} reflects the optimal trade-off between investment efficiency (the return on e) and insurance.[22] The optimal price falls between 1 and R because banks' relative risk aversion with respect to the liquidity shock, embedded in the concavity of V, is greater than 1 but less than infinity.[23] Under Assumption 1, V has enough concavity to lift the optimal price above 1 but keep it below R.

Note that, despite the fact that buyers in the private market for the illiquid asset (the patient types) are indifferent to the timing of their cash flow, the optimal allocation is consistent with the riskless payoff R being sold at date 1 at a price $p^{\text{fb}} < R$. This price is consistent with equilibrium in the private market due to infinite impatience of the impatient types, or cash-in-the-market pricing, as in Allen and Gale (1994).

Associated with the optimal price p^{fb} is the optimal allocation $(s_1^{\text{fb}}(0), x_1^{\text{fb}}(1))$ of the liquid and illiquid asset at date 1. From (12) and (13) we have

$$(s_1^{\text{fb}}(0), x_1^{\text{fb}}(1)) = \left(\frac{p^{\text{fb}} e}{\pi p^{\text{fb}} + 1 - \pi}, \frac{e}{\pi p^{\text{fb}} + 1 - \pi} \right). \tag{18}$$

The impatient banks attain the final investment $I_1^{\text{fb}}(0) = s_1^{\text{fb}}(0)$ and the patient banks attain $I_2^{\text{fb}}(1) = R x_1^{\text{fb}}(1) = \frac{R}{p^{\text{fb}}} s_1^{\text{fb}}(0) > s_1^{\text{fb}}(0)$. The inital investment in the two assets in this allocation is $s_0^{\text{fb}} = \pi s_1^{\text{fb}}(0)$ and $x_0^{\text{fb}} = (1 - \pi) x_1^{\text{fb}}(1)$.

4.3.2 With the ex ante participation constraint

Next, we reintroduce the participation constraint (16). Does the first-best optimum p^{fb} satisfy this condition? Clearly, (17) implies $p^{\text{fb}} - 1 \geq 0$, so the right-hand side of (16) is

[22] Although p^{fb} is defined as a solution to the second-best problem with private retrade, the notation reflects the fact that p^{fb} also solves the first-best planning problem—the problem in which all information is public, i.e., there are no IC constraints. Without the ex ante participation constraint (10), the IC constraint (7) does not bind, i.e., the first-best solution remains incentive compatible under private information about θ and private retrade in the secondary market for illiquid assets. See Farhi et al. (2009) for a full exposition of this result.

[23] It is easy to check that $p^{\text{fb}} = 1$ if V is logarithmic, and $p^{\text{fb}} = R$ if V is Leontieff.

satisfied. Whether $\frac{\lambda}{\pi} \geq p^{\text{fb}} - 1$ depends on the value of λ. If λ is high enough, p^{fb} will continue to be feasible even in the presence of the ex ante private investment constraint (16). In particular, let

$$\bar{\lambda} \equiv \pi(p^{\text{fb}} - 1). \tag{19}$$

This value is the threshold level of λ such that the ex ante participation constraint binds if $\lambda < \bar{\lambda}$ and does not bind if $\lambda \geq \bar{\lambda}$. With $\lambda < \bar{\lambda}$, thus, the first-best price p^{fb} is not incentive-feasible.

The next proposition solves for the constrained optimum with the participation constraint (16). The constrained-optimal retrade price will be denoted by p^*.

Proposition 3 *For all $\lambda \geq \bar{\lambda}$, the ex ante participation constraint does not bind and $p^* = p^{\text{fb}}$. For all $0 \leq \lambda < \bar{\lambda}$, the ex ante constraint binds and $p^* = \frac{\lambda}{\pi} + 1 < p^{\text{fb}}$. In sum,*

$$p^* = \min\left\{p^{\text{fb}}, \frac{\lambda}{\pi} + 1\right\}. \tag{20}$$

The proof follows from the fact that social welfare is increasing in p at all p smaller than p^{fb}. The constrained-optimal retrade price, therefore, is the largest price satisfying $\frac{\lambda}{\pi} \geq p - 1$ but not larger than p^{fb}.

The full constrained-optimal allocation, A^*, is determined by the price p^*. The allocation of the liquid and illiquid asset at date 1 is

$$(s_1^*(0), x_1^*(1)) = \left(\frac{p^* e}{\pi p^* + 1 - \pi}, \frac{e}{\pi p^* + 1 - \pi}\right), \tag{21}$$

and the final investment attained by the impatient and patient banks is, respectively,

$$I_1^*(0) = s_1^*(0) \text{ and } I_2^*(1) = R x_1^*(1) > s_1^*(0). \tag{22}$$

Substituting (20) into (21) and comparing with (18), we can also show how the date-1 allocation of assets in the constrained optimum relates to that in the first best:

$$s_1^*(0) \quad = \quad \min\left\{s_1^{\text{fb}}(0), \left(1 + \frac{1-\pi}{\pi}\frac{\lambda}{1+\lambda}\right)e\right\}, \tag{23}$$

$$x_1^*(1) \quad = \quad \max\left\{x_1^{\text{fb}}(1), \frac{e}{1+\lambda}\right\}. \tag{24}$$

The intuition behind Proposition 3 is as follows. If the ex ante constraint binds, i.e., if $\lambda < \bar{\lambda}$, the amount of insurance against the liquidity shock the planner can provide to banks is constrained by the threat of banks becoming nonbanks and employing the investment strategy described in Jacklin (1987). This strategy delivers the maximum value a nonbank can attain. In this strategy, the nonbank invests all of its resources in the illiquid assets. The nonbank subsequently sells these assets in the secondary market if it experiences a liquidity need at date 1, or holds onto them if it does not. Specifically, in this strategy the nonbank acquires $\frac{e}{1+\lambda}$ units of the illiquid asset x at date 0. If $\theta = 0$, it sells x and puts $\tilde{I}_1 = p\frac{e}{1+\lambda}$ in the long-term investment at date 1. If $\theta = 1$, it holds x to maturity and invests $\tilde{I}_2 = R\frac{e}{1+\lambda}$ at date 2.

The value of this strategy, $\mathbb{E}[V(\tilde{I}_1+\theta\tilde{I}_2)] = \mathbb{E}[V(\frac{pe}{1+\lambda}(1-\theta+\theta\frac{R}{p}))]$, decreases in the nonbanks' cost mark-up λ and increases in the private resale price p. The planner wants to increase p toward p^{fb} but is constrained by the banks' option of becoming nonbanks. Larger λ makes this option less attractive, which allows the planner to increase p more without triggering an exodus of assets to the nonbank sector.[24] If λ is large enough, i.e., $\bar{\lambda}$ or larger, the planner can lift p all the way up to p^{fb}.

It is easy to see how the constrained optimum depends on the quality of the investors' outside option, measured by the cost λ. If the outside option is sufficiently unattractive, i.e., $\lambda > \bar{\lambda}$, the participation constraint does not bind and the first best is attainable. Note that, as long as $\lambda > \bar{\lambda}$, it does not matter how high λ is. If the outside option is attractive enough to bind, however, the constrained optimum becomes worse the better the outside option is: lower λ implies lower p^* and lower final investment made at date 1, I_1^*. This means the amount of liquidity available at date 1 is smaller, and the maturity mismatch is worse. Banks get less insurance against the liquidity shock θ, and the illiquid asset is priced in the secondary market with a larger discount relative to its face value of R.

In the extreme case of $\lambda = 0$, the constrained-optimal price is $p^* = 1$, which coincides with the LF competitive equilibrium price and the optimal allocation of final investment (I_1^*, I_2^*) coincides with the competitive equilibrium allocation given in Theorem 1.

4.4 A pecuniary externality

With $\lambda = 0$, as we just saw, the ex ante participation constraint is so tight that the constrained-optimal price is $p^* = 1$, the constrained-optimal allocation coincides with the unregulated competitive equilibrium allocation, and thus the LF competitive equilibrium is constrained-efficient.

With $\lambda > 0$, however, there is a discrepancy between the competitive equilibrium and the constrained-optimal allocation. This market failure is due to what is often called a pecuniary externality. In our model, as in Farhi et al. (2009), the pecuniary externality is not due to incomplete markets, as in Geanakoplos and Polemarchakis (1986), but rather due to the fact that the resale price p enters the incentive compatibility constraint (7). If we extend the market structure from simple trade in the secondary market for the illiquid asset to the general one

[24] Equivalently, looking at (24) we can say that the planner wants to decrease $x_1(1)$ toward $x_1^{\text{fb}}(1)$ but is constrained by the possibility of investors obtaining $\tilde{x}_1(1) = \frac{e}{1+\lambda}$ as nonbanks. By making $\tilde{x}_1(1)$ smaller, higher λ decreases the value of this outside option and thus relaxes the constraint faced by the planner.

in which banks trade state-contingent claims—with one another or with a centralized counter-party—the pecuniary externality is still present and the competitive outcome is still inefficient. The reason is that in any competitive equilibrium, a firm—including a central counter-party—takes p as given, while the planner internalizes the impact of the primary allocation on the resale price. Formally, it is easy to extend the analysis in Farhi et al. (2009) to show that even the general Prescott-Townsend competitive equilibrium is inefficient in the present environment.

The strength of the pecuniary externality can be quantified by the size of the wedge between the constrained-optimal price p^* and the competitive equilibrium price, which is 1. Clearly, (20) implies that this wedge is strictly increasing in λ for all $\lambda < \bar{\lambda}$, and for $\lambda \geq \bar{\lambda}$ it remains constant, at its maximum level. The ex ante participation constraint, therefore, reduces the strength of the pecuniary externality. The tighter this constraint, i.e., the lower the cost λ, the weaker the externality.

5 Optimal liquidity regulation

Having characterized optimal allocations for all $\lambda \geq 0$, we now discuss implementation of these allocations as competitive equilibria with government regulation as defined in Section 3. If $\lambda = 0$, the constrained optimum coincides with the competitive equilibrium allocation and no intervention is needed. If $\lambda > 0$, however, the equilibrium allocation is inefficient due to the pecuniary externality. Next, we show how this market failure can be corrected with payment of interest on reserves (a subsidy to the liquid bank asset) and a proportional tax on the illiquid asset.

5.1 Competitive asset market equilibrium with IOR and a liquidity tax

Suppose the government pays interest on reserves (IOR) at the rate i and imposes a proportional (i.e., linear) liquidity tax τ on investment in the illiquid asset. The liquidity tax is imposed at date 0, which means that, in order to originate one dollar worth of the illiquid asset x, banks need to spend $1+\tau$ dollars. With IOR paid to the banks at the rate i, in order to have one dollar of liquidity at date 1, banks need to invest only $(1+i)^{-1}$ dollars in liquid reserves at date 0.

The government announces the policy rates (i, τ) before banks decide whether to remain a bank (i.e., subject to these policy instruments) or become a nonbank (not subject to government regulation but having to use the inferior illiquid asset origination technology with the extra cost λ). The government is committed to the preannounced rates. In equilibrium, all banks will at least weakly prefer to remain as banks rather than become nonbanks, and we assume that in the case of indifference they remain as banks. After this participation decision is made, banks invest and pay taxes at date 0, trade in the private market at date 1, and make their long-term investments I_t at dates $t = 1, 2$.

In the notation of Section 2, thus, the regulation/tax system we consider here is

$$M(s_0, x_0) = -\frac{i}{1+i} s_0 + \tau x_0, \tag{25}$$

under which the budget constraint in (1) now reads

$$\frac{s_0}{1+i} + (1+\tau)x_0 \leq e. \tag{26}$$

Competitive equilibrium is defined as in Section 3. Banks maximize (4) subject to the budget constraint (26) at date 0, and budget constraints (2) and (3) at date 1. The government faces a budget constraint $M(s_0, x_0) \geq 0$.

Next, we study banks' optimal investment choices subject to $M(s_0, x_0)$ given in (25). We characterize the solution to the banks' problem in the relevant range for the secondary illiquid asset market price p.

Lemma 6 *In the asset market economy with M given in (25), suppose $p \leq R$. Then, the investment portfolio (s_0, x_0) solving a bank's ex ante maximization problem has the bang-bang property: $s_0 = 0$ and $x_0 = \frac{e}{1+\tau}$ if $p > (1+\tau)(1+i)$; $s_0 = (1+i)e$ and $x_0 = 0$ if $p < (1+\tau)(1+i)$; and any (s_0, x_0) such that $\frac{s_0}{1+i} + (1+\tau)x_0 = e$ if $p = (1+\tau)(1+i)$. The ex ante expected value attained by a bank is*

$$\mathbb{E}\left[V\left(\max\left\{1+i, \frac{p}{1+\tau}\right\} e\left(1-\theta+\theta\frac{R}{p}\right)\right)\right]. \tag{27}$$

The bank's behavior with IOR i and tax rate τ has the same structure as in the laissez-faire equilibrium of Theorem 1. With access to the secondary market for the illiquid asset at date 1, which is open before the long-term investment I_1 has to be made, only the total value $s_0 + px_0$ of a bank's portfolio of liquid and illiquid assets matters to the bank. Indeed, given a portfolio (s_0, x_0), an impatient bank will sell the illiquid asset and put $I_1 = s_0 + px_0$ in the long-term investment at date 1. A patient bank will spend all its cash on the illiquid asset earning the additional return between date 1 and 2, which lets it invest $I_2 = (s_0 + px_0)R/p$ at date 2. In either case, the level of I_t the bank can afford is strictly monotone in $s_0 + px_0$, and whether or not the bank can earn the extra return R/p is out of its control.

The banks' problem of allocating e between the two assets at date 0, therefore, boils down to maximizing the date-1 value of its asset holdings, $s_0 + px_0$, subject to the budget constraint (26). In this maximization problem, both the objective and the constraint are linear in s_0 and x_0. The solution therefore has the bang-bang structure described in the above lemma.

As an immediate implication, note that banks will not choose an interior portfolio of cash and the illiquid asset unless the after-tax returns on these two investments are the same, i.e.,

$1+i=\frac{p}{1+\tau}$. The constrained-optimal allocation (associated with the retrade price p^*) requires a positive initial investment in both assets. Thus, for this allocation to be implementable, the after-tax returns on the two assets must be equal. We move to this discussion next.

5.2 Optimal IOR and liquidity tax

Theorem 2 *The optimal retrade price p^* and the associated optimal allocation A^* are an equilibrium in the economy with taxes (i,τ) if and only if*

$$1+i = \frac{s_1^*(0)}{e}, \tag{28}$$

$$1+\tau = \frac{e}{x_1^*(1)}, \tag{29}$$

$$(1+i)(1+\tau) = p^*. \tag{30}$$

Taxes (i,τ) make the optimal allocation and the optimal price p^* consistent with bank optimization by decreasing the value of the Jacklin deviation, which, as we saw in Theorem 1, forces the equilibrium price to be $p=1$ absent taxes. Indeed, without taxes and with the resale price $p^*>1$, a bank's value would be maximized by investing all resources e in the illiquid asset at date 0 and holding to maturity if patient, or selling at date 1 if impatient. This strategy would give the bank $x_1(1)=e>x_1^*(1)$ units of the illiquid asset at date 1 if patient, or $s_1(0)=p^*e=\frac{s_1^*(0)}{x_1^*(1)}e>s_1^*(0)$ units of the liquid asset if impatient, i.e., more than what the optimal allocation assigns, state-by-state.[25] Tax rates (i,τ) in (28) and (29) decrease the value of this investment strategy. With these tax rates, a deviating bank can acquire at most $\frac{e}{1+\tau}$ units of the illiquid asset at date 0. As (29) shows, this gives it $x_1(1)=\frac{e}{1+\tau}=x_1^*(1)$ units of the illiquid asset at date 1 if patient. If impatient, the bank can sell its illiquid portfolio and obtain $s_1(0)=p^*\frac{e}{1+\tau}=(1+i)(1+\tau)\frac{e}{1+\tau}=s_1^*(0)$ units of the liquid asset, where the last equality uses (28). Thus, the Jacklin strategy no longer delivers more assets than $(s_1^*(0),x_1^*(1))$, which makes the optimal allocation and the optimal price p^* consistent with equilibrium.

More generally, the rates i and τ in (28) and (29) change the structure of the return on initial resources e that a bank can earn regardless of its investment strategy. In the laissez-faire equilibrium of Theorem 1, with $p=1$, any initial allocation of e between the liquid and illiquid asset gives banks the same return of 1 between date 0 and date 1. Impatient banks must cash out at date 1 earning the total return of 1. Patient banks are able to earn the additional return R between dates 1 and 2, which gives them the total return of R. Because $(1+\tau)(1+i)=\frac{e}{x_1^*(1)}\frac{s_1^*(0)}{e}=p^*$, by Lemma 6, banks are also indifferent with respect to the initial asset allocation in the optimal equilibrium of Theorem 2. Any budget-feasible asset allocation produces the return of $1+i$ between date 0 and date 1, which is the total return earned by

[25] That $x_1^*(1)<e$ follows from $p^*>1$ and (21).

impatient banks. Patient banks earn an additional return R/p^* between dates 1 and 2, which gives them the total return of $(1+i)R/p^* = R/(1+\tau)$. As we see, the IOR rate i increases the total return earned by the impatient banks, from 1 to $1+i$, while the tax τ reduces the total return earned by the patient banks, from R to $R/(1+\tau)$. By doing so, i and τ improve ex ante liquidity insurance provided to all banks in equilibrium. In fact, they implement the maximal level of liquidity insurance consistent with incentives and the threat of shadow banking.[26]

These two rates, of course, are not independent. As we know from the planning problem, the whole optimal allocation is pinned down by the optimal retrade price p^*. Therefore, so are the optimal rates i and τ. Indeed, substituting (21) into (28) and (29) we get

$$\begin{aligned} i &= \frac{(1-\pi)(p^*-1)}{\pi(p^*-1)+1}, \\ \tau &= \pi(p^*-1). \end{aligned}$$

These expressions make it immediately clear that $i > 0$ and $\tau > 0$ unless $p^* = 1$, which we recall is the case only if $\lambda = 0$. Both i and τ increase in p^*, i.e., larger distortions are needed to implement a larger wedge $p^* - 1$.

Finally, we can use (19) and (20) to express the optimal rates as

$$i = \frac{1-\pi}{\pi}\frac{\min\{\lambda,\bar{\lambda}\}}{1+\min\{\lambda,\bar{\lambda}\}}, \tag{31}$$

$$\tau = \min\{\lambda,\bar{\lambda}\}. \tag{32}$$

As we see, optimal intervention is stronger when the threat of shadow banking is lower, i.e., when the cost of shadow banking λ is higher. Optimal policy rates become flat (no longer increasing) in λ for $\lambda > \bar{\lambda}$. This, of course, is because the first-best optimum is attained with any $\lambda \geq \bar{\lambda}$, and the optimal price $p^* = p^{\text{fb}}$ no longer increases in λ, which is evident in (20).

5.3 Robustness to regulatory arbitrage

Tax rates (i,τ) given in (31) and (32) are the unique proportional tax rates implementing the constrained-optimum as a competitive equilibrium subject to the threat of shadow banking. In our model, the ex ante choice to become a bank or a shadow bank is discrete, i.e., institutions cannot be both (or commingle banking and shadow banking activities). In this section, we point out that under the proportional tax system (i,τ), this restriction is innocuous, i.e., the equilibrium is robust to commingling of banking and shadow banking.

To see this, note first that if $\lambda > \bar{\lambda}$, then (32) implies $\tau = \bar{\lambda} < \lambda$. This means that a) nonbanks prefer to be banks as banks' cost of originating illiquid assets is lower and they

[26] Note also that positive policy rates i and τ flatten the after-tax yield curve.

can also earn IOR on liquid assets, and b) banks have no incentive to try to earn IOR on their holdings of liquid assets while moving their illiquid assets to shadow banking (perhaps by setting up an off-balance-sheet Structured Investment Vehicle), as the cost λ is higher than the tax τ. If $\lambda \leq \bar{\lambda}$, (32) implies that $\tau = \lambda$. Nonbanks cannot earn IOR, but they attain the same ex ante value as banks. This is because their optimal investment strategy is the Jacklin all-illiquid deviation, and under (i, τ) the banks are indifferent between all-illiquid, all-liquid, or an interior portfolio. Thus, banks have no incentive to move illiquid assets to the shadow banking sector because $\tau = \lambda$. Nonbanks could not benefit from having access to IOR because in order to collect it they would need to reduce their holdings of illiquid assets, which would decrease the "Jacklin premium" they can earn on the all-illiquid portfolio.

In addition to showing that equilibrium is robust to this kind of regulatory arbitrage, the above observations also make it clear that IOR is the only reason why banks are willing to hold liquidity (i.e., positive s_0) in equilibrium. With $\tau = \lambda$ and $i = 0$, banks would be sufficiently deterred from becoming nonbanks. But as banks, they would strictly prefer to hold only the illiquid asset if its resale price is p^*. This is because in the absence of i the return on the liquid asset is 1 and the return on the illiquid asset is $p^* \frac{1}{1+\tau} > 1$. The subsidy $i > 0$ to the liquid asset is just large enough to make banks indifferent to holding liquidity, which allows for positive liquid investment ex ante and supports the optimal price $p^* > 1$ as an equilibrium outcome.

Moreover, the robustness of the equilibrium to regulatory arbitrage depends crucially on the linearity of the subsidy-tax system (i, τ). This point can be easily seen in the following two examples.

First, consider a regulatory system in which there is no tax on illiquid assets, but instead the IOR paid to banks is funded with a lump-sum tax T levied on banks. With these instruments, a bank's budget constraint at date 0 is $(1+i)^{-1} s_0 + x_0 \leq e - T$. It is easy to verify that under the assumption of no commingling of banking and shadow banking this system can implement the constrained-optimal allocation with the IOR rate $i = p^* - 1 > 0$ and the lump sum tax $T = \pi e \frac{p^*-1}{\pi p^* + 1 - \pi} > 0$. In this system, however, the IOR rate is high (it covers the whole wedge $p^* - 1 > 0$), which gives investors an incentive to engage in regulatory arbitrage in the following way. By setting up as nonbanks, investors can avoid the tax T. By depositing their resources with a bank, they can earn the high IOR. The value of this strategy dominates the equilibrium value, as nonbanks, without paying T, still enjoy the same benefits as banks. This equilibrium, therefore, is not robust to regulatory arbitrage, but rather it does depend on the assumption of no commingling.

Similarly, consider a system in which there is no IOR but instead there is a lump sum subsidy S to banks funded by a proportional tax τ on illiquid bank assets. A bank's budget constraint is $s_0 + (1+\tau) x_0 = e + S$. It is easy to verify that this system can implement the constrained-optimal allocation with $\tau = p^* - 1$ and $S = (1-\pi) e \frac{p^*-1}{\pi p^* + 1 - \pi} > 0$ if no commingling is allowed. The investors, however, have an incentive to engage in regulatory arbitrage by setting up as banks, in order to collect S, while at the same time holding their illiquid assets in the nonbank sector, which effectively lets them pay the cost λ on the investment in illiquid assets instead of tax τ. Outside of the cases with λ higher than $\bar{\lambda}/\pi$, this cost is lower than the tax τ, which makes this strategy preferable to the on-equilibrium strategy. Therefore this equilibrium, too, is nonrobust to regulatory arbitrage.

6 Aggregate uncertainty

Thus far, we have assumed purely idiosyncratic shocks. In this section, we consider two extensions adding aggregate uncertainty. First, we study how the optimal tax rate τ and the IOR rate i depend on the economy-wide level of return on the illiquid assets, R. Then, we study how the optimal policy rates depend on the aggregate level of demand for liquidity. These extensions are essentially comparative statics, because we assume that aggregate uncertainty is resolved ex ante, before initial investment in the liquid and illiquid asset is made. Our analysis in this section, thus, shows how optimal policies depend on an exogenous aggregate state of the economy rather than how they respond to an ex post aggregate shock.

6.1 Liquidity regulation over the business cycle

Consider the following extension of the model. Fix the cost parameter $\lambda > 0$ and suppose the return on the illiquid asset R is a random variable drawn from some continuous distribution with support $[\underline{R}, \bar{R}]$, where $1 < \underline{R} < \bar{R}$. The realization of R becomes publicly known at date 0 before any decisions are made. In this section, we discuss how optimal policy rates (i, τ) depend on R.

The first-best retrade price p^{fb} now depends on the aggregate state of the economy, i.e., it depends on the realized value of R. Let us use the notation $p^{\text{fb}}(R)$ to denote this dependence.

Lemma 7 $p^{\text{fb}}(R)$ *is strictly increasing.*

The intuition for the above result is as follows. First-best optimality requires that ex post, i.e., at date 1, the marginal value of the liquid asset and the illiquid asset be the same. Under Assumption 1, higher R reduces the marginal value of the illiquid asset (as V' drops faster than R increases). To match this reduction in the marginal value, the quantity of the liquid asset must be increased. Thus, $s_1(0)$ is higher and, by the ex ante resource constraint, $x_1(1)$ is lower when R is higher, which implies that the retrade-market-clearing price $p = s_1(0)/x_1(1)$ is higher at the first best, i.e., p^{fb} is strictly increasing in R.

Definition (19) implies that the threshold $\bar{\lambda}$ below which the participation constraint binds is strictly increasing in p^{fb}. The above lemma thus implies that $\bar{\lambda}$ is strictly increasing in R.

We will denote this relation by $\bar{\lambda}(R)$. Given the fixed value of the cost parameter $\lambda > 0$, the participation constraint can bind in some states R and not in others depending on whether or not $\bar{\lambda}(R)$ is larger than λ. Assume $\underline{R}$ is close enough to 1 so that $\bar{\lambda}(\underline{R}) < \lambda$. Assume also $\bar{R}$ is high enough so that $\bar{\lambda}(\bar{R}) > \lambda$. Under these assumptions, by continuity and strict monotonicity of $\bar{\lambda}(R)$, there is a unique threshold $\hat{R} \in (\underline{R}, \bar{R})$ such that $\bar{\lambda}(\hat{R}) = \lambda$.

The solution to the mechanism design problem in Proposition 3 and the characterization of the optimal policy rates in Theorem 2 apply in each aggregate state R. In particular, the optimal retrade price is

$$p^*(R) = \min\left\{p^{\text{fb}}(R), \frac{\lambda}{\pi} + 1\right\}.$$

Lemma 7 immediately implies that the constrained-optimal price $p^*(R)$ increases with R for all $R < \hat{R}$. All $R \geq \hat{R}$, however, share the same constraint-optimal price: $p^*(R) = \frac{\lambda}{\pi} + 1$. Equation (21) then implies that the allocation of the liquid and illiquid asset at date 1 is also independent of R in this range. By (22), the final investment I_1^* of the impatient banks is the same while the final investment I_2^* of the patient banks increases one-for-one with R, at all $R \geq \hat{R}$.

The optimal policy rates (i, τ) are as follows. In all states $R \geq \hat{R}$, λ is lower than the threshold $\bar{\lambda}(R)$, the ex ante participation constraint binds, and (31) and (32) reduce to, respectively, $i = \frac{1-\pi}{\pi}\frac{\lambda}{1+\lambda}$ and $\tau = \lambda$, which are independent of R. The optimal policy rates are thus the same in all states $R \geq \hat{R}$. If we identify economic expansions with times when $R \geq \hat{R}$, the constrained-optimal policy rates (i, τ) remain constant as long as the economy is in an expansion. In particular, in a repeated version of the model, the policy rates would not respond to fluctuations in R as long as R remains above $\hat{R}$.

Note that above $\hat{R}$, higher return on illiquid asset R does give a higher expected value to the banks (because I_2 increases), but insurance against the liquidity shock is poorer (as I_1 does not increase in R). This is because higher R increases the value of the Jacklin all-illiquid investment strategy, which increases the value of becoming a nonbank. Only when the retrade price is kept constant (i.e., not increasing in R), the nonbanks' value does not increase faster than the banks' value, and the ex ante participation constraint is preserved.

In states $R < \hat{R}$, the threshold $\bar{\lambda}(R)$ becomes lower than λ, the participation constraint does not bind, and (31) and (32) reduce to $i = \frac{1-\pi}{\pi}\frac{\bar{\lambda}(R)}{1+\bar{\lambda}(R)}$ and $\tau = \bar{\lambda}(R)$. Since $\bar{\lambda}(R)$ varies with R, the optimal policy rates are now sensitive to the realization of R. In particular, they are strictly increasing in R. For example, if R is very low (close to $\underline{R}$), the participation constraint is very slack, but also the optimal price $p^*(R)$ is very close to 1, i.e., the pecuniary externality is weak and the policy rates i and τ needed to implement $p^*(R)$ are low.

In sum, the optimal IOR rate i and the illiquid asset tax rate τ remain constant in expansions, defined as $R \geq \hat{R}$. The optimal policy rates are determined by the binding investor participation constraint and do not respond to fluctuations in R within this range. In recessions, defined as $R < \hat{R}$, the two policy instruments become sensitive to R as the participation constraint is slack and the optimal policy rates must match the required wedge $p^*(R)-1$, which varies positively with R. In particular, in recessions both τ and i are below their expansion levels. In recessions, thus, it is optimal to reduce the policy rates (i, τ) with deeper reductions in deeper recessions.

6.2 Optimal regulation with an aggregate liquidity shock

Let us come back to a fixed return on the illiquid asset, R, and instead consider uncertainty over the aggregate demand for liquidity. We model high aggregate demand for liquidity as an aggregate state in which a higher fraction of banks become impatient at date 1. We analyze how the solution of the model and the optimal policy rates (i, τ) change in this state. As before, all aggregate uncertainty is resolved at date 0.

Formally, we consider two ex ante aggregate states with different probabilities of an investor being impatient (thus also with different ex post fraction of impatient investors) in each state. Let $\pi \in \{\underline{\pi}, \bar{\pi}\}$ with $0 < \underline{\pi} < \bar{\pi} < 1$. We will refer to $\underline{\pi}$ as the baseline state and $\bar{\pi}$ as the aggregate liquidity shock state.

We start by examining how the first-best allocation depends on π.

Lemma 8 *$x_1^{\mathrm{fb}}(1)$ and $s_1^{\mathrm{fb}}(0)$ are strictly decreasing in π, x_0^{fb} is strictly decreasing and s_0^{fb} is strictly increasing in π.*

As we see, the initial investment in the liquid asset is larger in state $\bar{\pi}$ but the ex post allocation of the liquid asset to each impatient bank is smaller. In a sense, the economy is poorer in the state $\bar{\pi}$ as high anticipated demand for liquidity leads to a decrease in investment in the illiquid, high-yield asset at date 0, decreasing the average return on the initial resources e.

Let us now consider the participation constraint. Two cases are possible. Case one: the participation constraint binds in the baseline, low-liquidity-demand state, which in terms of the definitions from the previous subsection corresponds to the situation in which the aggregate liquidity shock can occur during an expansion. Case two: the participation constraint does not bind in the baseline state, which corresponds to the situation in which the aggregate liquidity shock can hit during a recession.

Lemma 8 and equation (24) immediately imply the following.

Corollary 1 *If the participation constraint binds in state $\underline{\pi}$, then it also binds in state $\bar{\pi} > \underline{\pi}$.*

We now can discuss how the constrained-optimal IOR rate i and the tax rate τ depend on the aggregate liquidity shock. If the participation constraint binds in the baseline state, then, by Corollary 1, it binds in both aggregate liquidity states. Equation (24) thus implies that $x_1^*(1)$ is the same in the two aggregate liquidity states, equal to $\frac{e}{1+\lambda}$. From (29) we then obtain that the tax rate τ is the same in the two states as well, equal to λ. If the participation constraint is slack in state $\underline{\pi}$, then $x_1^*(1)$ is lower and, by (29), the tax rate τ is higher in state $\bar{\pi}$. Lemma 8 and equation (23) imply that $s_1^*(0)$ is lower in state $\bar{\pi}$ independently of whether the participation constraint binds in the baseline state $\underline{\pi}$ or not. By (28), this means that the optimal IOR rate i is always lower in the aggregate state with high anticipated liquidity demand.

In sum, optimal policy rates respond to the aggregate liquidity demand shock as follows. The IOR rate i is decreased when high liquidity demand is anticipated. The tax rate τ is increased if high liquidity demand occurs in a recession (where the investors' participation constraint is slack therefore τ can be increased) or is unchanged in an expansion (where the participation constraint already binds and τ cannot be increased anymore).

7 Optimal regulation via a minimum liquidity requirement

The foregoing analysis is focused on the IOR/tax implementation of the optimal allocation. In this section, we briefly discuss the corresponding implementation of optimal liquidity regulation via a quantity restriction.

In a model without the possibility of shadow banking, i.e, without the ex ante participation constraint, Farhi et al. (2009) show how the first-best optimum p^{fb} can be implemented with a minimum liquidity requirement of the following form

$$\frac{s_0}{e} \geq \iota, \tag{33}$$

where ι is a policy parameter.[27] Thus, the quantity regulation of Farhi et al. (2009) takes M to be identically zero in the budget constraint (1) but instead imposes (33) as an additional constraint in a bank's maximization problem at date 0. Farhi et al. (2009) show that with

$$\iota = \frac{s_0^{\mathrm{fb}}}{e}$$

the market equilibrium allocation coincides with the first-best optimal allocation associated with the first-best retrade price p^{fb}.

This result translates directly into our model, with the possibility of shadow banking restricting the implementable level of ι. In particular, it is not hard to check that the minimum liquidity requirement (33) implements the constrained-optimal allocation associated with p^* if

[27] Kucinskas (2015) studies a liquidity requirement of this form in a model with mutual funds.

the liquidity floor parameter is set as follows:

$$\iota = \frac{s_0^*}{e} = \min\left\{\frac{s_0^{\text{fb}}}{e}, \frac{\pi+\lambda}{1+\lambda}\right\}. \tag{34}$$

When the participation constraint does not bind, $\iota = \frac{s_0^{\text{fb}}}{e}$ as in Farhi et al. (2009). When it binds, however, ι cannot be as set as high as $\frac{s_0^{\text{fb}}}{e}$. The possibility of shadow banking in this case necessitates that the liquidity requirement be loosened. In particular, the participation constraint binds in expansions, i.e., when R is higher than $\hat{R}$. In expansions, therefore, the minimum liquidity requirement is not pinned down by the optimal trade-off between return and liquidity insurance, as in Farhi et al. (2009). It is pinned down by the binding participation constraint at the level $\frac{\pi+\lambda}{1+\lambda}$. In our model, thus, the optimal liquidity requirement is in expansions pinned down by a different set of forces than in Farhi et al. (2009). In particular, the requirement is not sensitive to the economy-wide rate of return R on illiquid assets.

In recessions, $s_0^* = s_0^{\text{fb}}$, and the liquidity requirement is sensitive to R. The proof of Lemma 7 in the Appendix shows that s_0^{fb} is strictly increasing in R. Thus, the liquidity requirement is relaxed further (the floor ι is lower) in deeper recessions (when R is lower).

Finally, (34) and Lemma 8 show how the optimal liquidity floor depends on the aggregate fraction of impatient banks π. Since both s_0^{fb} and $\frac{\pi+\lambda}{1+\lambda}$ are strictly increasing in π, the liquidity requirement is always tightened when high aggregate demand for liquidity is anticipated.

8 Conclusion

In this paper, we extend the pecuniary-externality-based theory of optimal liquidity regulation of banks by allowing for the possibility of shadow banking. We provide a highly tractable extension of the standard banking model with private retrade. In our model, the outside option of shadow banking boils down to a new constraint on the set of implementable asset prices. We view our analysis as making three contributions.

First, we show that the possibility of shadow banking tames regulation. In the extreme case of costless shadow banking ($\lambda = 0$), unregulated competitive equilibrium is efficient, i.e., optimal liquidity regulation is nil.

Second, we present a novel set of tools for implementation of the optimal liquidity policy: a flat-rate tax on illiquid assets combined with interest on reserves, i.e., a subsidy to liquid assets. In 2008, the need to support market liquidity was the justification given for accelerating Congress's authorization for the Federal Reserve to pay IOR to depository institutions.[28] Our analysis provides a normative rationale for IOR consistent with this justification. Indeed, the payment of IOR is in our model a part of the policy implementing the optimal level of liquidity

[28] See The President's Working Group on Financial Markets (2008).

in equilibrium.

Third, we conduct comparative statics exercises to show how optimal policy, both in the form of tariff (i, τ) and quota ι, responds to two kinds of aggregate uncertainty: changes in the rate return on illiquid assets, and changes in the aggregate level of liquidity demand caused by a heightened fraction of impatient banks. Interestingly, the optimal policy responds to fluctuations in the rate of return on illiquid assets only when this rate is low. The model thus suggests that bank liquidity regulation policy should be constant in expansions and relaxed in recessions.

It is possible to extend our approach and include shadow banking not as an off-equilibrium possibility constraining equilibrium outcomes but rather as a part of the equilibrium outcome. One way in which active shadow banking can be introduced to the model is to allow for heterogeneity among banks in the cost of shadow banking λ. We conjecture that our results characterizing optimal liquidity regulation policy continue to hold in this extension of the model. In particular, we conjecture that policy rates are more sensitive to the rate of return on illiquid assets when this rate is low.

Similarly, our model can be extended to endogenize the cost of shadow banking λ. One way to do it could be to introduce a friction in the secondary market for illiquid assets that puts nonbanks at a disadvantage relative to banks.[29] Another could be a reduced-form approach allowing the planner to spend resources ex ante to increase λ. We conjecture that our results are robust to these extensions. Optimal policy would still be implemented with IOR and a tax on illiquid assets with tighter policy needed in expansions. In particular, the planner's incentive to spend resources on increasing λ would be absent in recessions, while this margin would become active in expansions.

Appendix

Proof of Theorem 1

If $p < 1$, then banks' optimal choices are $(s_0, x_0) = (e, 0)$, $(n(0), I_1(0), I_2(0)) = (0, e, 0)$, and $(n(1), I_1(1), I_2(1)) = \left(\frac{e}{p}, 0, R\frac{e}{p}\right)$. Thus, $\mathbb{E}[n] = (1-\pi)\frac{e}{p} > 0$, i.e., the market-clearing condition (6) is violated, and so there is no equilibrium with $p < 1$. If $p > 1$, then banks' optimal choices are $(s_0, x_0) = (0, e)$, $(n(0), I_1(0), I_2(0)) = (-e, pe, 0)$, and $(n(1), I_1(1), I_2(1)) = (-e, pe, 0)$ for $p > R$ or $(n(1), I_1(1), I_2(1)) = (0, 0, Re)$ for $1 < p \leq R$. In either case, the market-clearing condition (6) is violated: with $p > R$, we have $\mathbb{E}[n] = -e < 0$ (both types want to sell); with $1 < p \leq R$, we have $\mathbb{E}[n] = -\pi e < 0$ (the impatient type wants to sell). Thus, there is no equilibrium with $p > 1$.

With $p = 1$, banks are indifferent among all pairs (s_0, x_0) such that $s_0 + x_0 = e$. Given any

[29] E.g., Diamond (1997) considers limited market access; Plantin (2014) considers adverse selection.

such pair (s_0, x_0) and $p = 1$, the optimal choices at date 1 are $(n(0), I_1(0), I_2(0)) = (-x_0, e, 0)$ and $(n(1), I_1(1), I_2(1)) = (s_0, 0, Re)$. The market-clearing condition $-\pi x_0 +$ $(1-\pi)s_0 = 0$ is satisfied if and only if $(s_0, x_0) = (\pi e, (1-\pi)e)$, which gives us a unique equilibrium. QED

Proof of Lemma 1

If A is incentive-feasible with a retrade price p, then redefining $s_1(\theta)$ to be $\hat{s}_1(\theta) = s_1(\theta) - pn(\theta)$, $x_1(\theta)$ to be $\hat{x}_1(\theta) = x_1(\theta) + n(\theta)$, and $\hat{n}(\theta) = 0$ delivers a new allocation that is resource feasible and incentive compatible with the same price p, but has zero trade. These two allocations yield the same final investment plan $(I_1(\theta), I_2(\theta))$, hence they achieve the same level of welfare for banks. Because p is the same at both allocations, the value of becoming a nonbank is not increased. Hence, the ex ante participation constraint is preserved at the new allocation with no trade. QED

Proof of Proposition 1

Necessity:

(i) By contradiction, if $s_1(0)+px_1(0) < s_1(1)+px_1(1)$, then type 0 would misreport to achieve a higher utility in the retrade market. Similarly, if $s_1(0) + px_1(0) > s_1(1) + px_1(1)$, then type 1 would misreport.

(ii) With $x_1(0) > 0$, the impatient type 0 could achieve a higher value than $V(s_1(0)+0Rx_1(0))$ (the value under no-trade) by selling $x_1(0)$ in the retrade market.

(iii) Because type 1 is indifferent between I_1 and I_2, he would choose to invest when it is relatively cheaper. If he chooses $s_1(1) > 0$, it must be the case that I_1 is weakly cheaper, that is, $\frac{R}{p} \leq 1$. Analogously, $x_1(1) > 0$ implies $\frac{R}{p} \geq 1$.

Sufficiency: The equal present value condition (i) implies truth-telling. The impatient type 0 will not want to retrade because of (ii). Likewise, the patient type 1 will not want to retrade at p because of (iii). QED

Proof of Lemma 2

(i) By contradiction, suppose one of the three RF constraints is slack at allocation A. There exists some $\delta > 1$ such that the allocation A^δ defined by $s_1^\delta(\theta) = s_1(\theta)\delta$, $x_1^\delta(\theta) = x_1(\theta)\delta$, $s_0^\delta = (\pi s_1(0) + (1-\pi)s_1(1))\delta$, and $x_0^\delta = (\pi x_1(0) + (1-\pi)x_1(1))\delta$ is still resource feasible. It follows from Proposition 1 that the new allocation A^δ remains incentive compatible under retrade price p. Like A, allocation A^δ satisfies the ex ante participation constraint because the retrade price p is the same, i.e., the deviator's value remains unchanged. Because A^δ achieves higher social welfare, A cannot be optimal.

(ii) We show that an incentive-feasible allocation A in which $s_1(1) > 0$ is dominated by another incentive-feasible allocation A' in which $s'_1(1) = 0$ and $p \leq R$. Note first that since A is IC, by part (iii) in Proposition 1, $s_1(1) > 0$ implies $p \geq R$. We construct A' separately for the case $p > R$ and for the case $p = R$.

(a) If $p > R$, then, by (iii) in Proposition 1, $x_1(1) = 0$ at A. By (ii) in Proposition 1, $x_1(0) = 0$, so $x_0 = 0$ at A. By (i) in Proposition 1, $s_1(0) = s_1(1) = e$ at A. The ex ante value delivered at allocation A is thus $V(e)$. This value is less than $\pi V(e) + (1-\pi)V(Re)$, which is delivered by allocation A' in which $p' = 1$ and

$$\begin{aligned} s'_1(1) &= x'_1(0) = 0, \\ s'_1(0) &= x'_1(1) = e. \end{aligned}$$

Allocation A' satisfies the ex ante participation constraint because if a nonbank chooses $\tilde{s}_0$ and $\tilde{x}_0 = \frac{e-\tilde{s}_0}{1+\lambda}$, his total income at date 1 is $\tilde{s}_0 + p'\tilde{x}_0 = \frac{e+\lambda\tilde{s}_0}{1+\lambda} \leq e$, and so he receives (weakly) less date-1 income, and thus also the final value, than what he can get as a bank at allocation A'.

(b) If $p = R$, then, by (i) and (ii) in Proposition 1, $s_1(0) = s_1(1) + Rx_1(1)$. Construct an allocation A' by $p' = R$ (i.e., the same as at A), $s'_1(1) = 0$, and

$$\begin{aligned} s'_1(0) &= s_1(0) + \epsilon s_1(1), \\ x'_1(1) &= x_1(1) + \frac{\epsilon+1}{R} s_1(1), \end{aligned}$$

where $\epsilon \equiv \frac{(1-\pi)(R-1)}{\pi R+1-\pi}$. It follows from $s_1(0) = s_1(1) + Rx_1(1)$ that $s'_1(0) = s'_1(1) + Rx'_1(1)$, so A' remains incentive compatible. The resource constraint is satisfied by A' because the definition of ϵ implies

$$\pi s_1(0) + (1-\pi)(s_1(1) + x_1(1)) = \pi s'_1(0) + (1-\pi)(s'_1(1) + x'_1(1)).$$

Allocation A' also satisfies the ex ante participation constraint, as the deviator's value remains unchanged under $p' = R = p$.

(iii) By (iii) in Proposition 1, $x_1(1) > 0$ implies $p \leq R$.

QED

Proof of Lemma 3

It follows from $s_1(0) = px_1(1)$ and the resource constraint that $s_1(0) = \frac{pe}{\pi p+1-\pi}$ and $x_1(1) = \frac{e}{\pi p+1-\pi}$. Therefore, type 0's value is $V(s_1(0)) = V\left(\frac{pe}{\pi p+1-\pi}\right)$, while type 1's value is $V(x_1(1)R) = V\left(\frac{pe}{\pi p+1-\pi}\frac{R}{p}\right)$. QED

Proof of Lemma 4

Fix a portfolio $(\tilde{s}_0, \tilde{x}_0)$ chosen by a nonbank at date 0. At date 1, an impatient nonbank will sell $\tilde{x}_0$ and invest $\tilde{I}_1 = \tilde{s}_0 + p\tilde{x}_0$. Because $p \leq R$, a patient nonbank will hold onto $\tilde{x}_0$ and buy $\frac{\tilde{s}_0}{p}$ additional units of the illiquid asset, which gives it the final investment $\tilde{I}_2 = (\frac{\tilde{s}_0}{p} + \tilde{x}_0)R = (\tilde{s}_0 + p\tilde{x}_0)\frac{R}{p}$ at date 2. In both cases, thus, the final investment attained by the nonbank depends on $(\tilde{s}_0, \tilde{x}_0)$ only though $\tilde{s}_0 + p\tilde{x}_0$, the value of the portfolio at date 1. The nonbank's portfolio choice at date 0, thus, is equivalent to maximizing this value subject to the budget constraint (5). This is a linear problem with the bang-bang solution $\tilde{s}_0 + p\tilde{x}_0 = e\max\left\{1, \frac{p}{1+\lambda}\right\}$. In sum, the value attained by the impatient nonbank is $V(\tilde{I}_1) = V\left(e\max\left\{1, \frac{p}{1+\lambda}\right\}\right)$, and the value attained by the patient nonbank is $V(\tilde{I}_2) = V\left(e\max\left\{1, \frac{p}{1+\lambda}\right\}\frac{R}{p}\right)$. QED

Proof of Lemma 5

Given Lemma 3 and Lemma 4, the private investment constraint can be written as

$$f\left(\frac{p}{\pi p+1-\pi}\right) \geq f\left(\max\left\{1, \frac{p}{1+\lambda}\right\}\right),$$

where f is a strictly increasing function defined as

$$f(t) = \mathbb{E}\left[V\left(te\left(1-\theta+\theta\frac{R}{p}\right)\right)\right]. \tag{35}$$

Applying f^{-1} to the above inequality, we have

$$\frac{p}{\pi p+1-\pi} \geq 1 \quad \text{and} \quad \frac{p}{\pi p+1-\pi} \geq \frac{p}{1+\lambda},$$

which gives us (16). QED

Proof of Proposition 2

The derivative of the objective function in (14) is $\frac{\pi(1-\pi)}{(\pi p+1-\pi)^2}G(p)$, where

$$G(p) \equiv V'\left(\frac{pe}{\pi p+1-\pi}\right)e - V'\left(\frac{Re}{\pi p+1-\pi}\right)Re.$$

At $p = 1$, it follows from Assumption 1 that $G(1) = V'(e)\,e - V'(Re)\,Re > 0$. At $p = R$, $G(R) = V'\left(\frac{Re}{\pi R+1-\pi}\right)(1-R)e < 0$. Intermediate Value Theorem states that there exists $p^{\text{fb}} \in (1, R)$ such that $G(p^{\text{fb}})$ (and thus also the derivative of the objective function) is zero.

Because $V'(\cdot)$ is decreasing, $\frac{pe}{\pi p+1-\pi}$ is increasing in p, and $\frac{Re}{\pi p+1-\pi}$ is decreasing in p, we know that $G(p)$ is decreasing in p. Therefore p^{fb} is unique. QED

Proof of Proposition 3

The derivative of the social welfare function is positive at all $p \in [1, p^{\text{fb}})$ but is negative at $p > p^{\text{fb}}$. If $\lambda \geq \bar{\lambda}$, because p^{fb} is feasible, the constrained-optimal price $p^* = p^{\text{fb}}$. If $\lambda < \bar{\lambda}$, because the social welfare function increases at all $p \in [1, 1 + \frac{\lambda}{\pi}]$, the constrained-optimal price $p^* = 1 + \frac{\lambda}{\pi}$. Thus, in (16) the constraint $p \geq 1$ never binds and the constraint $p \leq 1 + \frac{\lambda}{\pi}$ binds if and only if $\lambda < \bar{\lambda}$. QED

Proof of Lemma 6

The ex post budget constraints (2) and (3) can be collapsed to a single constraint

$$I_1 + \frac{p}{R} I_2 \leq s_0 + p x_0. \tag{36}$$

If $\theta = 0$ is realized, the bank will choose $I_1 = s_0 + p x_0$ and $I_2 = 0$. Because $p \leq R$, with $\theta = 1$ the bank will choose $I_1 = 0$ and $I_2 = (s_0 + p x_0) \frac{R}{p}$. Substituting these values into the bank's objective function gives us

$$V\left((s_0 + p x_0) \left(1 - \theta + \theta \frac{R}{p} \right) \right) \tag{37}$$

as the value that type θ bank achieves ex post. The ex ante problem of a bank can be thus written as maximization of $f(\frac{s_0 + p x_0}{e})$ subject to the budget constraint (26), where f is the strictly increasing function defined in (35). This problem is equivalent to maximization of $s_0 + p x_0$ subject to the budget constraint (26). This is a linear problem with the bang-bang solution structure given in the statement of the lemma. Substituting this solution into (37), we get (27). QED

Proof of Theorem 2

Assume i and τ are as in (28) and (29). Then (12) implies

$$(1+i)(1+\tau) = \frac{s_1^*(0)}{x_1^*(1)} = p^*.$$

By Lemma 6, it is individually optimal for each bank to invest ex ante in any portfolio such that $\frac{s_0}{1+i} + (1+\tau) x_0 = e$. In particular, the portfolio (s_0^*, x_0^*) does satisfy this condition because

$s_0^* = \pi s_1^*(0)$ and $x_0^* = (1-\pi)x_1^*(1)$ and so

$$\frac{s_0^*}{1+i} + (1+\tau)x_0^* = \frac{\pi s_1^*(0)}{\frac{s_1^*(0)}{e}} + \frac{e}{x_1^*(1)}(1-\pi)x_1^*(1) = e.$$

Market clearing and government budget balance follow from the fact that the optimal allocation A^* is an incentive-feasible allocation with price p^*. In particular, the market-clearing condition at date 1, $\pi p^* x_0^* = (1-\pi)s_0^*$, follows from the fact that A^* and p^* satisfy (12). Thus, we have an equilibrium.

Conversely, suppose p^* and the allocation A^* are an equilibrium in the economy with IOR and the tax on illiquid assets, under some rates (i, τ). Because A^* is interior, Lemma 6 implies that (i, τ) must satisfy

$$(1+i)(1+\tau) = p^*. \tag{38}$$

Because A^* satisfies the banks' budget constraint at date 0, we have $\frac{s_0^*}{1+i} + (1+\tau)x_0^* = e$. Multiplying this by $(1+i)$ and using (38), we get

$$s_0^* + p^* x_0^* = (1+i)e. \tag{39}$$

Because A^* satisfies the banks' budget constraints at date 1, we have $s_0^* + p^* x_0^* = s_1^*(0)$, and $s_0^* + p^* x_0^* = p^* x_1^*(1)$. Using (39), the first of these conditions implies (28). Using (38) and (39), the second one implies (29). QED

Proof of Lemma 7

In each aggregate state R, the first-best allocation satisfies

$$V'(s_1^{\mathrm{fb}}(0)) = V'(R x_1^{\mathrm{fb}}(1))R. \tag{40}$$

We will show that $x_1^{\mathrm{fb}}(1)$ is strictly decreasing in R. By contradiction, suppose $\underline{R} \leq R < R^1 \leq \bar{R}$ and $x_1^{\mathrm{fb}}(1)(R) \leq x_1^{\mathrm{fb}}(1)(R^1)$. The resource constraint (13) implies $s_1^{\mathrm{fb}}(0)(R) \geq s_1^{\mathrm{fb}}(0)(R^1)$. Hence,

$$\begin{aligned} V'(R x_1^{\mathrm{fb}}(1)(R))R &= V'(s_1^{\mathrm{fb}}(0)(R)) \\ &\leq V'(s_1^{\mathrm{fb}}(0)(R^1)) \\ &= V'(R^1 x_1^{\mathrm{fb}}(1)(R^1))R^1 \\ &\leq V'(R^1 x_1^{\mathrm{fb}}(1)(R))R^1, \end{aligned}$$

but Assumption 1 implies that $V'(Rx)R$ is strictly decreasing in R, which gives us a contradiction. Hence $x_1^{\mathrm{fb}}(1)$ is strictly decreasing in R. The resource constraint (13) thus implies that

$s_1^{\text{fb}}(0)$ is strictly increasing in R. These imply that $p^{\text{fb}} = \frac{s_1^{\text{fb}}(0)}{x_1^{\text{fb}}(1)}$ is strictly increasing in R. QED

Proof of Lemma 8

Equation (40), which holds in each aggregate state π, implies that $x_1^{\text{fb}}(1)$ and $s_1^{\text{fb}}(0)$ are co-monotone, i.e., change in the same direction as π changes. Suppose $x_1^{\text{fb}}(1)$ and $s_1^{\text{fb}}(0)$ are weakly increasing in π, i.e., $x_1^{\text{fb}}(1)(\underline{\pi}) \leq x_1^{\text{fb}}(1)(\bar{\pi})$ and $s_1^{\text{fb}}(0)(\underline{\pi}) \leq s_1^{\text{fb}}(0)(\bar{\pi})$. This leads to the following contradiction

$$\begin{aligned} e &= \underline{\pi} s_1^{\text{fb}}(0)(\underline{\pi}) + (1-\underline{\pi}) x_1^{\text{fb}}(1)(\underline{\pi}) \\ &< \bar{\pi} s_1^{\text{fb}}(0)(\underline{\pi}) + (1-\bar{\pi}) x_1^{\text{fb}}(1)(\underline{\pi}) \\ &\leq \bar{\pi} s_1^{\text{fb}}(0)(\bar{\pi}) + (1-\bar{\pi}) x_1^{\text{fb}}(1)(\bar{\pi}) \\ &= e, \end{aligned}$$

where the first inequality follows from $p^{\text{fb}}(\underline{\pi}) > 1$. The RF conditions in Definition 1 imply $(1-\pi)x_1^{\text{fb}}(1) = x_0^{\text{fb}}$. Since both $(1-\pi)$ and $x_1^{\text{fb}}(1)$ are strictly decreasing in π, so is x_0^{fb}. By the RF condition at date 0, s_0^{fb} is strictly increasing in π. QED

References

Adrian, T. and A. B. Ashcraft (2012). Shadow banking: A review of the literature. *FRBNY Staff Report* (580).

Allen, F. (1985). Repeated principal-agent relationships with lending and borrowing. *Economics Letters 17*, 27–31.

Allen, F., E. Carletti, and D. Gale (2009). Interbank market liquidity and central bank intervention. *Journal of Monetary Economics 56*(5), 639 – 652. Carnegie-Rochester Conference Series on Public Policy: Distress in Credit Markets: Theory, Empirics, and Policy. November 14-15, 2008.

Allen, F. and D. Gale (1994). Limited market participation and volatility of asset prices. *The American Economic Review 84*(4), 933–955.

Allen, F. and D. Gale (2004). Financial intermediaries and markets. *Econometrica 72*(4), 1023–1061.

Bengui, J., J. Bianchi, and L. Coulibaly (2015). Shadow banking and public liquidity. *Working Paper*.

Bernanke, B. (2012, April). Fostering financial stability. Speech at the 2012 Federal Reserve Bank of Atlanta Financial Markets Conference, Stone Mountain, Georgia.

Bhattacharia, S. and D. Gale (1986). Preference shocks, liquidity, and central bank policy. In W. Barnett and K. Singleton (Eds.), *New Approaches to Monetary Economics*, pp. 69–88. Cambridge University Press.

Bianchi, J. (2011). Overborrowing and systemic externalities in the business cycle. *American Economic Review 101*(7), 3400–3426.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives 23*(1), 77–100.

Cole, H. L. and N. R. Kocherlakota (2001). Efficient allocations with hidden income and hidden storage. *The Review of Economic Studies 68*, 523–542.

Di Tella, S. (2014, July). Optimal financial regulation and the concentration of aggregate risk. Stanford GSB Working Paper.

Diamond, D. W. (1997). Liquidity, banks, and markets. *Journal of Political Economy 105*(5), 928–956.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Farhi, E., M. Golosov, and A. Tsyvinski (2009). A theory of liquidity and regulation of financial intermediation. *The Review of Economic Studies 76*(3), 973–992.

Farhi, E. and J. Tirole (2012). Collective moral hazard, maturity mismatch, and systemic bailouts. *The American Economic Review 102*(1), pp. 60–93.

Financial Stability Board (2012). Strengthening oversight and regulation of shadow banking.

Financial Stability Forum (2009). Reducing procyclicality arising from the bank capital framework. *Working Paper*.

Freixas, X., A. Martin, and D. Skeie (2011). Bank liquidity, interbank markets, and monetary policy. *Review of Financial Studies 24*(8), 2656–2692.

Gale, D. and T. Yorulmazer (2013). Liquidity hoarding. *Theoretical Economics 8*(2), 291–324.

Geanakoplos, J. and H. Polemarchakis (1986). Existence, regularity, and constrained suboptimality of competitive allocations when the asset market is incomplete. In W. Heller, R. Starr, and D. Starrett (Eds.), *Essays in Honor of Kenneth Arrow*, Volume 3, pp. 65–95. Cambridge University Press.

Golosov, M. and A. Tsyvinski (2007). Optimal taxation with endogenous insurance markets. *Quarterly Journal of Economics 122*(2), 487–534.

Gorton, G. and A. Metrick (2010). Regulating the shadow banking system [with comments and discussion]. *Brookings Papers on Economic Activity Fall*.

Greenwood, R. and D. Scharfstein (2013). The growth of finance. *Journal of Economic Perspectives 27*(2), 3–28.

Holmstrom, B. and J. Tirole (1998). Private and public supply of liquidity. *Journal of Political Economy 106*(1), 1–40.

House, C. L. and Y. Masatlioglu (2015). Managing markets for toxic assets. *Journal of Monetary Economics 70*, 84 – 99.

Huang, J. (2014). Banking and shadow banking. *Princeton University Working Paper*.

Jacklin, C. J. (1987). Demand deposits, trading restrictions, and risk sharing. In E. C. Prescott and N. Wallace (Eds.), *Contractual Arrangements for Intertemporal Trade*, pp. 26–47. Minneapolis, MN: University of Minnesota Press.

Kehoe, T. J. and D. K. Levine (1993). Debt-constrained asset markets. *The Review of Economic Studies 60*(4), 865–888.

Kucinskas, S. (2015). Liquidity creation without banks. *Working Paper VU University Amsterdam*.

Lorenzoni, G. (2008). Inefficient credit booms. *The Review of Economic Studies 75*(3), 809–833.

Moreira, A. and A. Savov (2014). The macroeconomics of shadow banking. *NBER Working Paper* (20355).

Ordonez, G. (2015, January). Sustainable shadow banking. *University of Pennsylvania Working Paper*.

Plantin, G. (2014). Shadow banking and bank capital regulation. *Review of Financial Studies*, 146–175.

Pozsar, Z., T. Adrian, A. Ashcraft, and H. Boesky (2012). Shadow banking. *Federal Reserve Bank of New York Staff Reports* (458), 1 –35.

Stein, J. C. (2012). Monetary policy as financial stability regulation. *The Quarterly Journal of Economics 127*(1), 57–95.

The President's Working Group on Financial Markets (2008, October). Statement HP-1177. http://www.treasury.gov/press-center/press-releases/Pages/hp1177.aspx.

Ueda, K. and B. Weder di Mauro (2013). Quantifying structural subsidy values for systemically important financial institutions. *Journal of Banking and Finance 37*(10), 3830 – 3842.