Regulatory Arbitrage and Systemic Liquidity Crises

Stephan Luck†  Paul Schempp‡

Job market paper, September 22, 2015

After the 07-09 financial crisis, contractual linkages between commercial banks and the shadow banking sector have been identified as a source of fragility and targeted by recent regulatory reforms. In this paper, we present a novel channel through which panics in the shadow banking sector can adversely affect commercial banks despite absence of any contractual linkages between the two sectors. We develop a banking model in which commercial banks are regulated and covered by a safety net, while shadow banking circumvents regulatory requirements and thus comes with the prospect of panic-based runs. Given that both sectors use the same sources for wholesale funding, a run on shadow banks also deteriorates the funding conditions of regulated banks by inducing fire sales. Commercial banks may thus become illiquid and insolvent even though there is no risk of a classic bank run. The equilibrium size of the shadow banking sector is too large as agents neither internalize their impact on the equilibrium fire-sale prices nor on the costs the deterioration of funding conditions imposes on the safety net. We discuss implications for macroprudential regulation, particularly restrictions on wholesale funding, and indicate limitations of the Basel III liquidity regulation.

Key works: regulatory arbitrage; shadow banks; bank runs; systemic risk; fire sales; liquidity regulation

JEL codes: G21, G23, G28

∗We are very thankful for Martin Hellwig’s extensive advice and support. We are also thankful to Markus Brunnermeier, Jean-Edouard Colliard, Olivier Darmouni, Peter Englund, Maryam Farboodi, Valentin Haddad, Hendrik Hakenes, Nobuhiro Kiyotaki, Stephen Morris, Anatoli Segura, as well as participants of the summer school in economic theory of the econometric society in Tokyo and the student research workshop in Princeton. Financial support by the Alexander von Humboldt Foundation and the Max Planck Society is gratefully acknowledged.

†luck@coll.mpg.de, University of Bonn, Princeton University and Max Planck Institute for Research on Collective Goods.

‡schempp@coll.mpg.de, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, D-53113 Bonn, Germany. Phone: ++49 (0) 228 91416-65
1. Introduction

Regulatory arbitrage and the growth of shadow banking has been identified as one of the main ingredients to the 2007-09 financial crisis (Brunnermeier, 2009; FCIC, 2011). In particular, explicit or implicit contractual linkages between the shadow banking sector and commercial banks have been identified as a source of fragility (Acharya and Matthew, 2009; Acharya et al., 2013). Hence, post-crisis reforms have targeted the contractual channels through which the turmoil in the shadow banking sector has affected the commercial banking sector (compare, e.g., Section 619 of the Dodd-Frank Act, referred to as the “Volcker Rule”, Report of the Vickers Commission, Liikanen Report). Naturally, the question arises whether the implemented and proposed reforms are effective in reducing financial fragility of the banking sector?

This paper sheds lights on this issue by discussing a new theoretical channel for how regulatory arbitrage in banking may contribute to overall financial fragility despite absence of contractual linkages between regulated and unregulated banking. We show that panic-based runs in the shadow banking sector adversely affect the commercial banking sector via a deterioration of funding conditions. The main mechanism is that panic-based runs in the shadow banking sector induce fire sales. A binding cash-in-the-market constraint then also leads to a deterioration of funding conditions in the market for secured wholesale funding of regulated banks. Contagion via the deterioration of funding conditions implies that runs outside the commercial banking sector may lead to illiquidity and insolvency of commercial banks even when there are no runs inside the banking sector. The deposit insurance may need to live up to its promise and require funding even when insured institutions are not subject to panics. Thus, we argue that non-regulated banking activities may affect the commercial banking sector via channels beyond those that have been targeted in the post-crisis reforms.

Moreover, we show that the extent of regulatory arbitrage is excessive and the equilibrium size of the shadow banking sector too large. The underlying mechanism is a pecuniary externality that operates through fire-sale prices. When deciding to hold deposits at a shadow bank instead of at a regular bank, agents do not internalize that they reduce the fire-sale price. Low fire sale prices in turn have two negative effects: On the one hand, they make shadow banks less attractive, and on the other hand increase the funding needs of the deposit insurance scheme via a deterioration of funding conditions for regulated banks. Under the premise, that the extent of regulatory arbitrage cannot be controlled, the question arises whether restrictions on wholesale funding may be a desirable tool for macroprudential regulation of regular banks. In our setup, such a policy would shield the regulated banking sector from adverse consequences originating
outside the sector. However, it would also lead to allocative inefficiencies that imply further growth of the shadow banking sector and thus stronger fire sales in case of a crises. Liquidity regulation may thus backfire: even though the banking sector becomes stable, overall financial stability may erode. The overall welfare implications depend crucially on how the social planner weighs the cost of the deposit insurance scheme.

Our findings contribute to the understanding on how regulatory arbitrage and recent regulatory reforms contribute to financial stability. Previous to the 2007-09 financial crisis, many commercial banks had set up off-balance sheet conduits to finance long-term real investment by issuing short-term debt (Pozsar et al., 2013). In the summer of 2007, increased delinquency rates on subprime mortgages ultimately led to the collapse of the conduits’ main source of funding: the market for asset-backed commercial papers (ABCP) (see, e.g., Kacperczyk and Schnabl, 2009; Covitz et al., 2013). Many commercial banks had explicitly or implicitly sponsored these conduits and the collapse forced banks to take the conduits’ assets and liabilities on their balance-sheets, thus creating severe solvency issues.

From an ex-post perspective, it appears that off-balance sheet banking had to a large extent been conducted to circumvent existing capital regulation (see, e.g., Acharya et al., 2013). In this context, the adverse implications of explicit as well as implicit contractual linkages between regulated and non-regulated banks have been identified as a particularly important source of instability (Segura, 2014). Consequentially, the overwhelming regulatory response has been to close the obvious loopholes in regulation by outright prohibition of contractual links between depository institutions and other parts of the financial system.

The regulatory measures have been implemented under the premise that a prohibition of explicit or implicit contractual linkages between commercial banking and other types of banking can shield the former from turmoil originating in the latter. In particular, regulation has focused on prohibiting sponsor support, as well as on the separation of traditional banking and other activities, such as proprietary trading and market making.

---

1 Asset-backed commercial paper conduits were set up to finance mortgage-backed securities (MBS) and asset-backed securities (ABS) by issuing ABCP or medium-term notes (MTN), and were granted explicit credit or liquidity guarantees (credit or liquidity enhancements), or implicit guarantees as in the case of structured investment vehicles (SIV).

2 To some observers, this had already been clear prior to the crisis; see Jones (2000).

3 The reform proposals include a ring-fencing of depository institutions and systemic activities (Report of the Vickers Commission), separation between different risky activities (Liikanen Group), and prohibition of securities trading in commercial banks (Section 619 of the Dodd-Frank Act, referred to as the “Volcker Rule”).

4 E.g., the Financial Services Act of 2013, which was based in the Vickers Commissions Report, suggests
At the same time, the question has been raised whether depository institutions should be allowed to fund some of their activities by using wholesale funding. The Basel III liquidity regulation proposes to restrict the holding of potentially illiquid assets and the reliance on unstable funding sources by introducing the liquidity coverage ratio (LCR) and the net stable funding ratio (NSFR).

We argue that prohibiting contractual linkages is not sufficient to shield the regulated banking sector from financial fragility. Naturally, commercial banks and shadow banks hold similar assets and rely on similar sources of funding. Thus, a turmoil in the shadow banking sector is likely to have an impact on the regulated banking sector, even in the absence of any contractual linkages. If the extent of regulatory arbitrage is not under the control of the regulator, other macroprudential tools such as a strict liquidity regulation may be a viable option to improve the attained allocation. It can be debated whether the liquidity regulation as proposed in Basel III is sufficiently strict to effectively shield commercial banks. However, our model indicates that in order to justify such a regulation, one has to trade off the efficiency of commercial banks and the size of the shadow banking sector against the cost of financing the deposit insurance scheme.

Our model builds on – and at the same time nests – the canonical banking model of maturity and liquidity transformation by Diamond and Dybvig (1983). We enrich the setup along two dimensions: On the one hand, we introduce a new type of agents, called “investors”. These agents are only present from the interim period onwards and can provide funds to banks in exchange for claims on future cash-flows, as in Luck and Schempp (2014b). This makes it optimal for intermediation to substitute storage with what we refer to as “secured wholesale funding”. On the other hand, we introduce a shirking technology that allows a disciplining role of short-term debt, reminiscent of Calomiris and Kahn (1991) and Diamond and Rajan (2001). First, we show that the disciplining effect of short-term debt allows intermediaries to implement the first-best allocation in our setup (Proposition 1). Intermediaries refrain from shirking as they cannot enjoy private benefits if their depositors collectively withdraw. We also show that while short-term debt may be disciplining, it is also necessarily associated with the possibility of panic-based runs (Proposition 2), as discussed by Admati and Hellwig (2013). A regulator that decides to provide a deposit insurance to eliminate panic-based runs would thus undermine the disciplining effect of short-term debt. This makes it necessary to complement a safety net with capital requirements, which induces diligent

---

3
behavior via a textbook skin-in-the-game mechanism à la Tirole (2010).

Assuming that financing bank activities with equity is costly, incentives to circumvent regulation may arise. Intermediaries may place themselves outside the regulatory perimeter in the shadow banking sector. Institutions in this sector are not covered by the deposit insurance and have no access to the central bank’s discount lending, implying that they can be subject to panic-based runs. We emphasize a new theoretical channel through which these runs may be contagious, affecting the regulated banking sector: A systemic run on shadow banks induces fire sales with cash-in-the-market pricing à la Allen and Gale (1994). In particular, a binding cash-in-the-market constraint implies that wholesale funding conditions also deteriorate in a fire sale (Proposition 3). As regulated commercial banks optimally rely on secured wholesale funding, a fire sale creates real costs for them. Depending on the size of the shadow banking sector, regulated banks may ultimately become insolvent even though there are no classic bank runs. Insolvency of regulated bank in turn implies that the the deposit insurance may require actual funding.

We derive the equilibrium composition of the economy (Proposition 4) by introducing sunspot runs as in Cooper and Ross (1998) and Gertler and Kiyotaki (2015). Consumers trade-off stable low interests offered by commercial banks against higher but risky interests offered by shadow banks. While holding an account at a bank comes with the regulatory cost resulting from the capital requirement, short-term claims are insured. In contrast, a shadow bank deposit is not subject to the regulatory cost and can thus promise a higher interest. However, the downside of a deposit at a shadow bank is the prospect of a run, potentially inducing a costly fire sale. The equilibrium size of the two sectors is determined by consumers being indifferent between holding a regular bank deposit and a shadow bank deposit. As consumers are atomistic they do not internalize their impact on the equilibrium fire sale prices, similar to Lorenzoni (2008). Consequentially, in equilibrium, the shadow banking sector is too large such that fire sale prices are too low (Proposition 5). Due to the low fire sale prices, expected utility of a shadow bank deposit is lower than would be socially optimal. Moreover, welfare is further decreased as the low equilibrium fire sale price implies a deterioration of funding for commercial banks in case of a run. Consequentially, the expected cost of the deposit insurance scheme are higher in equilibrium than what would be socially optimal.

Under the premise that the extent of regulatory arbitrage cannot be controlled by regulator, we ask whether wholesale funding restrictions may improve the attained allocation. We show that wholesale funding restrictions shield the regulated banking sector

\[^5\text{We follow the definition of shadow banking in Luck and Schempp (2014a).}\]
from any adverse consequences of runs in the shadow banking sector. The expected costs for the safety net scheme are thus reduced to zero. However, the allocative inefficiency makes the regulated banking sector less attractive and increases the shadow banking sector (Proposition 6). This in turn implies that expected fire sales are even lower. Whether the attained allocation is welfare improving depends on how much weight the social planner puts on the cost of funding the deposit insurance scheme.

We further discuss the role of direct contractual linkages between the two sectors in the form of liquidity guarantees. We show that liquidity guarantees are optimal from the perspective of a single intermediary, but they exacerbate the adverse consequences of runs in the shadow banking sector (Proposition 7). In this case, intermediaries do not internalize their effect on fire sales when providing liquidity guarantees. This shows that the prohibition of contractual linkages may indeed be desirable, but not necessarily sufficient.

Finally, we also briefly analyze how government interventions in the form of a Lender of Last Resort (LoLR) and a Market Maker of Last Resort (MMLR) may affect the stability of the financial system. We find that while these interventions can shield the regulated banks, they necessarily also benefit the shadow banks.

We proceed as follows: Section 2 and Section 3 give the models’ setup as well as the first-best allocation, and show how intermediaries can implement the first-best allocation, but also point to the fact that this implementation is fragile in the sense that panic-based runs may occur. Section 4 shows how a deposit insurance combined with a capital requirement can eliminate the fragility. Section 5 gives the two important results of the paper: it shows how regulatory arbitrage reintroduces fragility even in the absence of contractual linkages and shows why the equilibrium size of the shadow banking is too large from a welfare-perspective. Finally, Sections 6 to 8 show the effects of wholesale funding restrictions, liquidity guarantees and central bank interventions before Section 9 and Section 10 conclude with a review of the literature and a discussion.

2. Setup

Consider an economy that goes through a sequence of three dates, \( t \in \{0, 1, 2\} \). There is a single good that can be used for consumption as well as for investment. The economy is populated by three types of agents: consumers, intermediaries, and investors.

Technologies
Altogether, there are three technologies available for investment (see a summary of the payoff structure in Table 1). There is a short technology (“storage”) available in \( t = 0, 1, \ldots \).
transforming one unit invested in $t$ into one unit in $t + 1$. Moreover, there are two illiquid technologies available for investment in $t = 0$: a “productive technology” and an unproductive “shirking technology”. Both technologies are technologically illiquid, i.e., for one unit invested they produce a return in $t = 1$ only if they are physically liquidated, and the physical liquidation rate of the technologies is assumed to be $\ell \to 0$. Note that the technologies (or claims on the technologies’ future returns) may nonetheless be sold at higher values at a secondary market that will be specified below.

The return properties of the illiquid technologies in $t = 2$ are as follows: One unit invested in the productive technology yields a safe return of $R$ units in $t = 2$. One unit invested in the shirking technology yields a safe return of $R_{\text{shirk}} < 1$ in $t = 2$. However, this technology yields a private benefit $B > 0$ in $t = 2$ which is available only to the agent who owns investment at this point in time, i.e., it is non-transferable and non-contractible. Moreover, it only accrues if the technology is not physically liquidated in the interim period.

<table>
<thead>
<tr>
<th></th>
<th>$t = 0$</th>
<th>$t = 1$</th>
<th>$t = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage in $t = 0$</td>
<td>-1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Storage in $t = 1$</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Productive technology</td>
<td>-1</td>
<td>$\ell \to 0$</td>
<td>$R$</td>
</tr>
<tr>
<td>Shirking technology</td>
<td>-1</td>
<td>$\ell \to 0$</td>
<td>$R_{\text{shirk}} + B$</td>
</tr>
</tbody>
</table>

*Table 1: Payoff structure of technologies*

We assume that $R_{\text{shirk}} + B \leq 1$. This implies that the shirking technology is inefficient, although it generates a private benefit. As will become clear later, the possibility of investing in this technology and financing the investment by short-term debt will give rise to moral hazard. This moral hazard will lead to the necessity of capital regulation once a deposit insurance undermines the disciplining effect of short-term debt.

**Consumers**

There is a continuum of consumers with mass one. Initially, consumers face idiosyncratic uncertainty with regard to their preferred date of consumption, and they may lend their endowment to intermediaries to invest on their behalf.

Each consumer is endowed with 1 unit of the good in $t = 0$. There are two types

---

*We assume that consumers cannot invest in technologies directly in the initial stage and trade technologies in the interim period. They can only lend their funds to intermediaries. In the Section we argue briefly why we can focus on a banking solution directly, i.e., why a banking solution dominates a financial markets solution.*
of consumers, *patient* and *impatient* consumers: a fraction $\pi$ is impatient and derives utility only from consumption in $t = 1$, $u(c_1)$, and a fraction $1 - \pi$ is patient and derives utility only from consumption in $t = 2$, $u(c_2)$. We restrict attention to CRRA utility, i.e., the period-utility function has the form $u(c_t) = \frac{1}{1-\eta} c_t^{1-\eta}$, with $\eta > 1$.

Initially, consumers do not know their type; their probability of being impatient is identical and independent, so all consumers have the same prior $\pi$ initially. In period one, each consumer privately learns his type, this can be considered as a liquidity shock.

A consumption profile $(c_1, c_2)$ denotes an allocation where an impatient consumer receives $c_1$ and a patient consumer receives $c_2$. As of period 0, such a consumption profile induces an expected utility of

$$U(c_1, c_2) = \pi u(c_1) + (1 - \pi) u(c_2) = \frac{1}{1-\eta} \left[ \pi c_1^{1-\eta} + (1 - \pi) c_2^{1-\eta} \right]. \tag{1} \{\text{eq:utility}\}$$

Notice that the attributes *patient* and *impatient* characterize the consumer’s exogenous type, which determines his preference. In contrast, the attributes *late* and *early* will characterize the timing of actual consumption, and in the case of demand-deposit contracts, it denotes the withdrawals, which are endogenous: An “early consumer” withdraws in $t = 1$, while a “late consumer” withdraws in $t = 2$.

**Intermediaries**

There is a mass $m$ of intermediaries. While consumers cannot invest in the technologies directly, intermediaries face no investment restrictions. Intermediaries have no market power, they compete for the consumers’ funds which they collect in exchange for a demand-deposit contract, and they invest the funds in the technologies. Moreover, they may choose to invest some of their own funds in the intermediation business.

Intermediaries only care about $t = 2$ consumption. Each intermediary is endowed with $E$ units of the good. We assume that $E$ is large, implying that no result will be driven by the aggregate intermediaries’ endowment $mE$ becoming a binding resource constraint. Importantly, intermediaries are assumed to have an outside option, resulting in a required return of $\rho > R$ in $t = 2$ for each unit invested in $t = 0$. Because the required return is larger than the technologies’ returns, it is costly for the consumers if the intermediaries invest their own endowment for investment. As we will see later, this assumption makes it costly to use a skin-in-the-game mechanism in order to provide

---

7It is assumed that $m$ is small compared to the mass of depositors such that each bank has a very large number of depositors, and thus does not face aggregate liquidity risk by a law of large numbers argument.

8As the model has no aggregate uncertainty, the shape of intermediaries’ utility is not important. They may be risk-neutral or risk-averse. Only for the case of sunspot runs with positive probability we will assume that intermediaries are risk-neutral in order to keep the analysis tractable.
intermediaries with incentives to invest in the productive technology instead of in the shirking technology in the presence of a deposit insurance.

On the liability side, intermediaries initially offer the deposit contract \((c_1, c_2)\) to consumers in exchange for one unit of initial deposits. Moreover, intermediaries choose to invest \(e_0\) units of their endowment in the intermediation business in \(t = 0\), in exchange for receiving \(e_2\) units in \(t = 2\). While we do not initially impose restrictions on how intermediaries finance intermediation, equity financing will turn out to be optimal.

On the asset side, intermediaries make the following investment decision: We denote by \(I\) the investment in the productive technology, by \(I_{shirk}\) the investment in the shirking technology, and \(1 + e_0 - I - I_{shirk}\) denotes the investment in storage. We assume that an intermediary’s investment decision is unobservable in \(t = 0\), but becomes public information in \(t = 1\).

**Investors**

There is a continuum of investors of mass \(n\). Investors only become active in the interim period and can provide liquidity to intermediaries: Investors can transfer some of their endowment to intermediaries in exchange for a claim on some of the future returns of the intermediaries’ technologies. We refer to this activity as “secured wholesale funding”. Later, we will show that the presence of these investors makes investment in storage inefficient, i.e., it is optimal for intermediaries to rely on wholesale funding from outsiders instead of storing real goods. However, this will also give rise to the main contagion channel between regulated and unregulated banking: When a run on shadow banks induces a fire sale, a cash-in-the-market constraint can become binding, and wholesale funding conditions for regulated banks deteriorate as well.

Investors are born in \(t = 1\) and receive an endowment of \(A/n\), so the investors’ aggregate endowment is given by \(A\). The endowment \(A\) will be one of the crucial parameters of the model: while it may be a sufficient source of liquidity in normal times, it may lead to a binding cash-in-the-market constraint in case of systemic runs. Given that investors are born in \(t = 1\), it is not possible to contract with them in \(t = 0\). Investors care about consumption\(^9\) in period 2, and they are assumed to have an outside option, which induces a required return of \(\gamma\), where \(\gamma \in [1, R]\). That is, for each unit they transfer to intermediaries in \(t = 1\), they need to receive at least \(\gamma\) units in \(t = 2\).

Investors have no market power and thus are price takers as long as their liquidity \(A\) is not scarce, i.e., they take the conditions of wholesale funding as given. The required return \(\gamma\) implies that they are willing to provide liquidity as long as the return \(r\) they

---

\(^9\)Again, the shape of their utility function is not important as long as it is compatible with the specified outside option.
receive satisfies \( r \geq \gamma \).

There are two different contractual specifications of secured wholesale funding (i.e., of how investors provide liquidity to intermediaries) which are economically equivalent: asset sales and collateralized lending. If investors purchase assets with a face value of \( R \) in period 2 at a price \( p \) in period 1, they get a return of \( r = R/p \). On the other hand, they can also lend one unit to banks at the interest rate \( r = R/p \) while receiving \( r/R = 1/p \) units of asset as collateral.

As long as liquidity is not scarce, competition among investors will induce \( r = \gamma \). However, if liquidity becomes scarce, we assume that the asset price and the interest charged in collateralized lending is determined by a cash-in-the-market constraint.

3. Optimal Intermediation and Runs

3.1. First-Best

We will now derive the allocation that maximizes the expected utility of consumers, subject to the participation constraints of intermediaries and investors, and subject to the resource constraints. Since our objective is to maximize consumers’ welfare, we treat these participation constraints as resource constraints. We refer to the resulting allocation as the first-best allocation and denote it by \( (c^*_1, c^*_2) \).

In the first-best, the shirking technology is not made use of as the productive technology strictly dominates the shirking technology, i.e., \( I_{\text{shirk}} = 0 \). We denote by \( L \) the units of the productive technology that get transferred from intermediaries to investors, in exchange for \( Lp \) units of the good (“liquidity”) from investors to intermediaries in the interim period. This is referred to as “secured wholesale funding” in the following.

The first-best maximization program is given by
\[
\max_{(c_1, c_2, e_0, e_2, I, L, p) \in \mathbb{R}_+^7} \pi u(c_1) + (1 - \pi) u(c_2),
\]
subject to
\[
\pi c_1 \leq (1 + e_0 - I) + Lp, \quad (3) \quad \{\text{eq:period1}\}
\]
\[
(1 - \pi)c_2 \leq (I - L)R - e_2, \quad (4) \quad \{\text{eq:period2}\}
\]
\[
e_2 \geq \rho e_0 \geq 0, \quad (5) \quad \{\text{eq:PC_intermediary}\}
\]
\[
R \geq \gamma p, \quad (6) \quad \{\text{eq:PC_investor}\}
\]
\[
pL \leq A, \quad (7) \quad \{\text{eq:arbitrage_budget}\}
\]
\[
I \leq 1 + e_0, \quad (8) \quad \{\text{eq:investment}\}
\]
\[
L \leq I. \quad (9) \quad \{\text{eq:liquidation}\}
\]

The budget constraints for periods one and two are given by (3) and (4). Investors may transfer \(Lp\) to consumers in \(t = 1\) in exchange for \(L\) units in \(t = 2\). As indicated above, we refer to this as wholesale funding. (5) represents the participation constraint of the intermediary and non-negativity constraint, and (6) represents the participation constraint of investors. The resource constraint on investors’ capital \(A\) ("interim liquidity") in the interim period is given by (7). Finally, (8) and (9) denote the constraint on initial investment as well as the constraint on the units of assets that can be sold in the interim period.

As discussed in the Appendix, depending on the model parameters \(A, R, \gamma, \) and \(\pi\), as well as on the shape of the utility function, the first-best program now has three solution candidates. As discussed in detail in the appendix, investment in storage is only optimal if \(A\) is small, and becomes unnecessary when \(A\) is sufficiently large. For the remaining part of the paper, we will assume that we are in the case in which the endowment of the investors \(A\) is large enough such that the investors’ budget constraint (26) is not binding. In this case, storage is not used, and there is only investment in the productive technology, i.e., \(I^* = 1\). This translates into the following assumption:

**Assumption 1.** \(A \geq \xi \equiv \pi \gamma \frac{R}{(1 - \pi) + \pi \gamma} \frac{1}{\frac{1}{\gamma} + \frac{1}{\beta}}\) \{assu:lowerA\}

For a detailed discussion of the implications of Assumption 1, see Appendix A, in which we also characterize the the first-best for the case that investors’ capital is scarce. Assumption 1 allows us to focus on a setup where intermediation optimally relies exclusively on investors providing interim liquidity through wholesale funding and refrains from the use of storage.

**Lemma 1** (First-Best Allocation). The first-best allocation is characterized by

\[
I^* = 1, \quad L^* = \xi \gamma / R, \quad \text{and} \quad e_0 = e_2 = 0,
\]
and the optimal consumption profile is given by

\[ c_1^* = \gamma^{-\frac{1}{\eta}} \frac{R}{(1 - \pi) + \pi \gamma^{1 - \frac{1}{\eta}}} \quad \text{and} \quad c_2^* = \frac{R}{(1 - \pi) + \pi \gamma^{1 - \frac{1}{\eta}}}. \] (10)

The risk-sharing between early and late consumers is described by the FOC \( u'(c_1) = \gamma u'(c_2) \) because under wholesale funding, the technological rate of substitution between period 1 and 2 is given by the investors’ required return \( \gamma \). Diamond and Dybvig (1983) restrict attention to utility functions with a relative risk aversion larger one. In their setup, risk-sharing between patient to impatient consumers is optimal, implying that \( c_{1DD} > 1 \), where 1 is the technological rate of return between periods 0 and 1 (storage). However, this condition also enables self-fulfilling runs. In contrast, we focus on the special case of constant relative risk aversion. The parameter of relative risk aversion is constant and given by \( \eta > 1 \). We get a similar result with respect to risk-sharing: It holds that \( c_1^* > R/\gamma \), where \( R/\gamma \) is the rate of return between periods 0 and 1 under wholesale funding. But as we shall see in next subsection, this condition also has similar and important implications for fragility and self-fulfilling runs.

Lemma \[ \text{in the Appendix describes the first-best if Assumption 1 does not hold. For} \ A < \xi, \ the \ investors’ \ endowment \ constraint (26) \ becomes \ binding. \ Furthermore, \ there \ exists \ some \ threshold \ \xi_0 < \xi \ below \ which \ partial \ investment \ in \ storage \ becomes \ optimal. \ In \ the \ extreme \ case \ of \ \gamma = R \ or \ A = 0, \ the \ optimal \ consumption \ profile \ is \ identical \ to \ that \ in \ the \ Diamond \ and \ Dybvig \ model \ with \ CRRA \ utility, \ which \ is \ thus \ nested \ in \ our \ model. \]

### 3.2. Intermediary Implementation

In the following, we will show that the first-best allocation can be implemented by demand-deposit contracts offered by the intermediaries. We first show that consumers are willing to lend to intermediaries, although intermediaries have the option of investing in the shirking technology. We will show that the demand-deposit contracts allow depositors to discipline the intermediary. In a second step, we show that the disciplining element of the demand-deposit contract is associated with financial fragility in the sense that panic-based runs may take place in the interim period.

**Disciplining Demand-Deposit Contracts**

We assume that consumers cannot invest in the technologies directly, but only via intermediaries. Let us first consider the agency problems on the part of the intermediary.
To this end, let us first devote more attention to the timing and the action space of consumers and intermediaries\footnote{Notice that in case of unsecured wholesale funding, one would have to worry about the behavior of investors as well; compare Luck and Schempp (2014b). However, for the case of asset sales or collateralized lending, we do not have to worry about the investors’ behavior as long as they cannot collude in order to extract rents from consumers.}

A consumer can choose whether and where to deposit her endowment in period 0, and an intermediary can then choose how to invest this endowment on her behalf. In period 1, consumers learn their type and observe the intermediary’s investment choice from the initial period, and they can decide whether to withdraw based on this information.

Let us assume that competition among intermediaries forces them offer to the first-best demand-deposit contract \((c_1^*, c_2^*)\) in exchange for the consumers’ endowment. In period 1, consumers have the possibility to withdraw the promised amount of \(c_1^*\), or to wait until period 2. We have assumed that an intermediaries investment decision \(I_{\text{shirk}}\) is not observable in \(t = 0\), but becomes publicly observable before consumers make their withdrawal decision in \(t = 1\).

**Proposition 1** (Implementation of the First-Best: Demand-Deposit Contracts). There exists a subgame-perfect Nash equilibrium in which the first-best consumption profile \((c_1^*, c_2^*)\) is implemented by the intermediaries offering demand-deposit contracts.

Consider the following strategy of a consumer for the period-1 subgame: She withdraws if she turns out to be impatient or if the intermediary has chosen \(I_{\text{shirk}} > 0\), and she does not withdraw if she turns out to be patient and the intermediary has chosen \(I_{\text{shirk}} = 0\). We will now show that if all consumers use this strategy, this strategy profile constitutes a Nash equilibrium in the period-1 subgame for any investment decision of the intermediary, and the optimal strategy of the intermediary is to choose \(I_{\text{shirk}} = 0\).

Assume that the intermediary has chosen \(I_{\text{shirk}} > 0\). Because all other consumers withdraw, it is a best response to do so as well because the intermediary is illiquid and insolvent already in \(t = 1\). Notice further that if \(I_{\text{shirk}}\) is large enough, withdrawing actually becomes a dominant strategy because the intermediary will be illiquid and insolvent in \(t = 2\) even without a run.

Now assume the intermediary has only invested in the productive technology, i.e., \(I = 1\) and \(I_{\text{shirk}} = 0\). Given that only impatient consumers withdraw, the intermediary will be able to serve all early consumers by selling \(L^*\) units of her investment to investors. Because \(A \geq \xi = \pi c_1^*\) by assumption, the investors’ funds are sufficient to serve all early depositors. As \(c_2^* > c_1^*\), it is a best response for patient consumers to wait.
This withdrawal strategy is a credible punishment strategy, and it uses the threat of a bank run as a disciplining device: Because the intermediary anticipates that all consumers will withdraw in \( t = 1 \) whenever she invests in the shirking technology, she knows that she will not be able to enjoy the private benefit \( B \). Therefore, she does not invest in the shirking technology in the first place. This disciplining effect of short-term debt is reminiscent of the findings of Calomiris and Kahn (1991), and Diamond and Rajan (2001), and allows intermediaries to implement the first-best allocation via demand-deposit contracts.

Note that there also exists a continuum of subgame-perfect Nash equilibria in which the bank chooses to invest a positive fraction in the shirking technology, but is not disciplined by the depositors up to this fraction. We discuss such equilibria in Appendix B. In the following, we restrict attention to the equilibrium proposed above. This is equivalent to assuming that an intermediary can only exclusively invest in either the productive or the shirking technology.

### 3.3. Fragility

While short-term debt is disciplining in our model, it is also a source of fragility. In fact, the model exhibits multiple equilibria in the period-1 subgame. Depending on the amount of investors’ funds \( A \), qualitatively different run equilibria emerge. As long as the amount of funds \( A \) is sufficiently large, potential runs on some intermediaries do not affect other intermediaries. However, if the endowment of investors \( A \) is relatively small, liquidity can become scarce in case of a run on many intermediaries. This puts the market for liquidity under stress and deteriorates the funding conditions of other intermediaries.

The price \( p \) of assets sold in period 1 depends on the aggregate amount \( L \) of assets sold if and only if the investors’ resource constraint becomes binding. As long as the resource constraint is not binding, competition among investors ensures that the price is equal to the investors’ willingness to pay. Thus, if \( A \) is so large such that \( L \) units of the asset can purchased by investors at price \( p = R/\gamma \), this is the market-clearing price, i.e., the price is equal to the assets’ rate of return \( R \) divided by the rate of the investors’ outside option \( \gamma \). If, however, \( A \) is scarce relative to the amount \( L \) of assets sold (i.e., if \( A \) is not sufficient to purchase \( L \) units at price \( R/\gamma \)), the market clears via cash-in-the-market pricing, i.e., it must hold that \( pL = A \).
Given $L$, the amount of assets sold, the price of the assets in period 1 is given by

$$p(L) = \begin{cases} \frac{R}{\gamma} & \text{if } \frac{AR}{\gamma} \geq L \\ \frac{A}{L} & \text{if } \frac{AR}{\gamma} < L. \end{cases}$$ (11)

Recall that by Assumption 1 we restrict our attention to the case in which intermediaries exclusively invest in the productive technology, i.e., $I^* = 1$, so the amount of assets sold is at most one. The price for assets in period 1 is depicted in Figure 1 for the case that $A < \frac{R}{\gamma}$.

![Figure 1: This graph depicts the potential fire-sale price for the case $R/\gamma > A$. In this case, a run may lead to depressed fire-sale prices via the binding cash-in-the-market constraint.](fig:firesale)

As long as $A \geq \frac{R}{\gamma}$, liquidity cannot become scarce, and the price is always given by $\frac{R}{\gamma}$. However, if $A < \frac{R}{\gamma}$, runs on some intermediaries can have negative external effects on others. If sufficiently many intermediaries experience a run, the price $p$ gets depressed, thus deteriorating the refinancing condition of other intermediaries.

**Micro-Fragility: Runs on Single Institutions**

Let us start by considering the stability of a single intermediary. Notice that, on the one hand, the price on the secondary market is limited by the investors’ willingness to pay, i.e., $p \leq \frac{R}{\gamma}$, but, on the other hand, the optimal demand-deposit contract promises an early consumption level that is strictly larger than this amount, $c^*_1 > \frac{R}{\gamma}$. Because $p < c^*_1$, it holds true that if all depositors of one specific intermediary $i$ run, this intermediary has to sell all assets, but still cannot fulfill all her obligations to her depositors. This particular intermediary becomes illiquid and insolvent already in period
1, and in particular could not serve any late consumer. Thus, a run on intermediary $i$ constitutes an equilibrium.

**Lemma 2 (Single-Institution Runs).** Assume that intermediaries choose the first-best investment level and demand-deposit contract. There exists a Nash equilibrium in the period-1 subgame in which there is a run on some intermediary $i$, inducing a complete asset sale and immediate illiquidity and insolvency of this intermediary. In particular, there exists an equilibrium in which there is a run on all intermediaries.

Notice that the run on a mass $j$ of intermediaries does not necessarily affect the remaining mass $1 - j$ of other intermediaries. If there is sufficient investor capital $A$, the price on the market remains high enough to make it possible that there exists an equilibrium where some mass $j$ of intermediaries face a run, but the rest does not face a run. The reason is that if $A$ is large enough and if (conditional on $A$) the mass $j$ of intermediaries who face a run is sufficiently low, the price in the secondary market is high enough to make “prudent” behavior at the intermediaries $1 - j$ compatible in equilibrium with runs elsewhere. Nonetheless, it may be true that all intermediaries are experiencing a run at the same time.

**Macro-Fragility: Systemic Runs and Cash-in-the-market-pricing**

Notice first that, if $A > R/\gamma$, it holds that $p(L) = R/\gamma$ for all $L$. This means that even in case of an economy-wide run, the price on the secondary market is unaffected and there is no binding cash-in-the-market constraint. This also implies, that if all intermediaries except for $i$ had a run, this run would not affect $i$ at all, because it can sell the designated amount $L^*$ at the expected price $p = R/\gamma$, so it could refinance at the ex-ante expected conditions.

Now consider the case of $A < R/\gamma$. This implies that there is cash-in-the-market pricing in case of an economy-wide run, implying that $p(1) = A$. Thus, if all intermediaries but one are experiencing a run, the intermediary who is not experiencing a run will yet face deteriorated funding conditions. We refer to runs as “systemic runs” if they induce cash-in-the-market pricing and thus affect the overall funding conditions.

**Proposition 2 (Systemic runs).** Assume that $A < R/\gamma$, and assume that intermediaries choose the first-best investment level and demand-deposit contract. Then there exist “systemic runs”, i.e., an economy-wide run in the period-1 subgame leads to cash-in-the-market pricing and thus a deterioration of overall funding conditions.

Proposition [1] shows that the ability to withdraw early induces diligent behavior of the intermediary, so short-term debt has a disciplining effect in our model. However, there
always exist multiple equilibria. In one class of equilibria, only some single institutions experience runs while others do not, and the latter ones remain completely unaffected. From Lemma 2 we learn that runs are always possible (on single institutions, but also economy-wide runs). However, the runs on single institutions occur independently. From Proposition 2 we learn that runs are contagious via deteriorated funding conditions only if $A < R/\gamma$, i.e., if investor capital is scarce. Whenever $A < R/\gamma$, there exists a second class of equilibria, in which runs also become contagious in the sense that they affect funding conditions of other institutions. The second type of run will be particularly important when we analyze later how runs in the shadow banking sector may affect funding conditions for the regulated banking sector.

Finally, it is important to notice that there is an implicit assumption underlying the existence of systemic runs: Our model does not allow that funds withdrawn at one shadow bank immediately re-enter the system as deposits at another intermediary or via the secondary market. Without this restriction, further frictions would be needed to explain systemic runs, such as frictions in interbank trade as pointed out by Skeie (2008).

4. Deposit Insurance and Optimal Bank Regulation

As we have seen in the previous section, the first-best is implementable through non-regulated intermediaries, but this implementation is fragile in the sense that there always exist run equilibria in the period-1 subgame. To eliminate such panic-based bank runs, assume that the regulator provides a credible deposit insurance: The regulator guarantees each bank depositor that she will receive exactly the amount in $t = 2$ that she was promised. In a setup without aggregate uncertainty and with multiple equilibria, introducing a deposit insurance that is credible may eliminate the adverse run equilibrium at no cost, e.g., as discussed by Diamond and Dybvig (1983). By guaranteeing patient consumers that they will receive their promised payment in the final period, the strategic complementarity is eliminated. Thus, the deposit insurance is never tested in equilibrium and is costless.

11 Alternatively, one could consider a lender of last resort, who grants access to the discount window of the central bank in order to prevent panics. The role of the lender of last resort is discussed in a later section.
12 It would be sufficient to guarantees each bank depositor that she will receive at least the amount in $t = 2$ that she was promised for $t = 1$. However, the specification of a complete deposit insurance will turn out more convenient when calculating the cost of the deposit insurance below.
13 An alternative measure often discussed in the literature is to allow intermediaries to suspend con-
In our setup, however, a deposit insurance – if implemented without further regulatory policy measures – can give rise to opportunistic behavior on the part of intermediaries, which imposes costs on the provider of such deposit insurance. The reason is that in the presence of deposit insurance, consumers do not care about the investment behavior of the intermediary, thus eliminating the disciplining effect of short-term debt. Even if they know that the intermediary will be insolvent in the second period, they do not run because they know that the deposit insurance will pay them at least the amount that the demand deposit contract entitles them to withdraw in the interim period. Therefore, an intermediary may have incentives to invest in the shirking technology.

Given the moral hazard problem arising from the deposit insurance, there exists an optimal regulatory response. In the first-best, the intermediary does not invest any of his funds in the intermediation business. The intermediary has no skin in the game, and the participation constraint $e_2 \geq \rho e_0$ is trivially satisfied by $e_0 = e_2 = 0$. This is efficient because there is no need to provide the intermediary with incentives, and given $\rho > R$, it would be costly for consumers to use the intermediary’s funds.

If a regulator wants to rule out moral hazard, she can do so by requiring the intermediaries to hold junior claims on their intermediation business. Optimal regulation thus calls for a minimal equity requirement via a classic skin-in-the-game argument. To insure diligence, the incentive compatibility constraint of the intermediary has to be satisfied. It is given by

$$e_2 \geq (1 + e_0)B. \quad \text{(12)}$$

At the same time, the intermediary’s participation constraint, $e_2 \geq \rho e_0$, still needs to be fulfilled.

In the second-best, both constraints are binding, i.e., $e_2 = (1 + e_0)B$ and $e_2 = \rho e_0$, yielding the second-best equity stakes

$$e_0^{**} = \frac{B}{\rho - B}, \quad \text{and}$$
$$e_2^{**} = \frac{\rho B}{\rho - B}. \quad \text{(13)}$$

Because it is costly to use intermediary’s funds, it holds that in an optimal regulatory regime that tries to prevent the intermediary from investing in the shirking technology, as little as possible intermediary capital is used, but enough to ensure diligent behavior.

vertibility. One can easily see that the discussion below would be equivalent under suspension of convertibility: suspension of convertibility may successfully prevent panic-based runs, but also undermines the disciplining effect of demand deposit contracts. If banks are able to suspend convertibility, regulation will also be necessary to ensure diligent behavior of intermediaries.
Given this necessary equity level, we can derive the second-best demand-deposit contract. The FOC is again given by \( u'(c_1) = \gamma u'(c_2) \).

**Lemma 3 (Second-best Contract).** Assume that demand deposits are protected by a credible deposit insurance. Optimal bank regulation requires intermediaries to satisfy an equity-to-debt ratio of \( B/(\rho - B) \), and intermediaries will hold exactly \( e_0^* = B/(\rho - B) \).

There exists no run equilibrium in the period-1 subgame. Given that \( \xi \leq A \), investment and sales are given by

\[
I^{**} = 1 + \frac{B}{\rho - B} \quad \text{and} \quad L^{**} = \frac{\pi \gamma^{1 - \frac{1}{\eta}} R - \frac{B}{\rho - B}(\rho - R)}{(1 - \pi) + \pi \gamma^{1 - \frac{1}{\eta}}},
\]

and the optimal consumption is given by

\[
c_1^{**} = \gamma - \frac{1}{\eta} R - \frac{B}{\rho - B}(\rho - R) \quad \text{and} \quad c_2^{**} = \frac{R - \frac{B}{\rho - B}(\rho - R)}{(1 - \pi) + \pi \gamma^{1 - \frac{1}{\eta}}}.
\]

In the regime with a deposit insurance, the consumption levels are decreasing in the private benefit \( B \) as well as in the required return of intermediaries \( \rho \). Obviously, first-best (Lemma 1) and second-best coincide if \( B = 0 \) or \( \rho = R \). For any other \( B > 0 \) and \( \rho > R \), the second-best consumption levels are strictly lower. In fact, the case of \( B > 0 \), but \( \rho = R \), is very interesting. In this case, using intermediary capital is not costly, and the first-best can always be implemented by using intermediary capital and investing it in the production technology until incentives are provided.

Importantly, there are no run equilibria in the interim period. The allocative inefficiency comes with the benefit of financial stability. However, as we will emphasize in the next section, this overall stability can only be attained if we exclude the possibility of regulatory arbitrage.

### 5. Regulatory Arbitrage and Fragility

In the previous section, we have abstracted from the possibility of regulatory arbitrage. In the following, we assume that the regulator provides a safety net and imposes a capital requirement on those intermediaries that are covered by the deposit insurance, hereafter referred to “commercial banks” or “regulated banks”. However, we assume that it is also possible for intermediaries to place themselves outside of the regulatory perimeter of banking. Intermediaries who engage in this kind of regulatory arbitrage are referred to as “shadow banks” in the following. In this case, they will neither be
regulated nor covered by the deposit insurance. However, shadow banks are disciplined in their investment behavior by short-term debt contracts.\textsuperscript{14}

First, we analyze a situation in which both regulated commercial banks and shadow banks coexist, taking the size of the sectors as given, and analyze how systemic risk that emerges in the shadow banking sector can spread to the commercial banking sector. Second, we will pin down the equilibrium size of the shadow banking sector when we assign positive probability to systemic runs. We then show that the equilibrium size of the shadow banking sector is larger than what would be socially optimal.

5.1. Coexistence of Banks and Shadow Banks

Assume that in $t=0$, intermediaries can decide whether they want to become a commercial bank or a shadow bank\textsuperscript{15}. An intermediary that sets up a commercial bank will be subject to the capital requirement as described above, and thus be required to inject some of his endowment in her banking business. In exchange, her business will be covered by the safety net. An intermediary that operates a shadow bank in turn will not need to invest any own funds.

A regulated bank can offer consumption levels of $(c^b_1, c^b_2) = (c^{b*}_1, c^{b*}_2)$ in exchange for a consumer’s endowment, where the superscript $b$ stands for bank. The expected utility of a bank customer is thus decreasing in $B$ and $\rho$. A shadow bank can offer a consumption profile given by $(c^{sb}_1, c^{sb}_2) = (c^*_1, c^*_2)$, where the superscript $sb$ stands for shadow bank. The drawback of being a customer at a shadow bank is that the shadow bank sector is not covered by the safety net and thus panic-based runs are possible.

We assume that in $t=0$, consumers only deposit either in the banking sector or the shadow banking sector. Consumers will compare the expected utility that they will receive from holding a bank account to holding a shadow bank account. In order to pin down the actual expected utility we first need to analyze what happens in case of a run before determining the equilibrium composition of the economy.

\textsuperscript{14}While by legal standards shadow banks have historically not offered demand deposits in reality, they do issue claims that are essentially equivalent to demand deposits, such as equity shares with a stable net assets value (stable NAV), or other instruments such as asset-backed commercial papers or repurchase agreements. For tractability, we will assume that shadow banks are literally taking demand deposits.

\textsuperscript{15}This decision is assumed to be binary. We relax this assumption in Section 7 when discussing the role of liquidity guarantees.
5.2. Fire Sales and the Deterioration of Funding Conditions

Let \( \sigma \in [0, 1] \) denote the size of the shadow banking sector. Taking the size of the shadow banking sector as given, we now analyze under which conditions a run on shadow banks leads to cash-in-the-market pricing, i.e., to systemic runs in the spirit of Proposition 2.

There exists some threshold \( \bar{\sigma} \) such that in case of a run on shadow banks, the fire sale price of assets is given by

\[
p(\sigma) = \begin{cases} 
  \frac{R}{\gamma} & \text{if } \sigma \leq \bar{\sigma} \\
  A - (1 - \sigma)\pi_c b_1 & \text{if } \sigma > \bar{\sigma}.
\end{cases}
\]  

(17) \{eq:fire_sale\}

In case of a run on the shadow banking sector, all shadow banks try to serve all withdrawing depositors. In order to fulfill their obligations, they sell all their assets, i.e., the shadow banking sector sells a total amount of \( \sigma \) units. As long as there is no cash-in-the-market pricing, shadow banks thus absorb an amount of liquidity \( \sigma R/\gamma \). However, this does not suffice to serve all customers because \( c^b_1 > p \geq R/\gamma \). In addition, commercial banks also need an amount \((1 - \sigma)\pi_c b_1\) of liquidity to satisfy their withdrawing impatient consumers. A run on shadow banks is thus not compatible with a price \( p = R/\gamma \) (i.e., it leads to cash-in-the-market pricing) and thus has a negative effect on regulated banks if the sum of these two terms exceeds the available funds \( A \).

**Proposition 3.** A run on shadow banks is systemic and affects commercial banks if \( \sigma > \bar{\sigma}(A) \), i.e., if the shadow banking sector is larger than \( \bar{\sigma}(A) \), given by

\[
\bar{\sigma}(A) = \frac{A - \pi_c b_1^{**}}{R/\gamma - \pi_c b_1^{**}}.
\]

\{prop:contagion\}

If \( A < R/\gamma \), it holds that \( \bar{\sigma}(A) < 1 \), i.e., the price in case of a run decreases whenever the shadow banking sector is large enough. Although regulated banks cannot be subject to runs because they are covered by the deposit insurance, they are affected via a deterioration of their funding conditions as the cash-in-the-market constraint becomes binding due to the runs on shadow banks. That is, in case of a run on all shadow banks, the asset price in the secondary market will be lower than \( R/\gamma \). Thus, commercial banks will have to refinance themselves at conditions that may make it impossible to fulfill their promise \((c^1_1, c^2_1)\) to each consumer. Unless the banks’ equity cushion may be able to absorb these losses, banks will not be able to payout \((c^1_1, c^2_1)\). In particular, if the shadow banking sector is large, the fire-sale price will be so low such that banks become insolvent in \( t = 2 \), or already illiquid and insolvent in \( t = 1 \). This implies in turn that the deposit insurance is required to pay out actual funds and thus becomes costly.
The underlying mechanism is that banks and shadow banks essentially hold the same assets and rely on the same sources of funding. This gives rise to contagion in case of a crisis via a simple binding cash-in-the-market constraint. The fire sales that are induced by a run in the shadow banking sector lead to a deterioration of funding conditions for regulated banks. Commercial banks may become illiquid and insolvent because of runs outside of the banking sector even if there are no actual runs in the banking sector. In canonical bank run models, a deposit insurance eliminates run equilibria at no cost. Here, runs in the shadow banking sector ultimately may make a funding of the deposit insurance scheme for regulated banks inevitable.

Many post-crisis reforms have targeted the contractual linkages between the two sectors. In particular, any type of ex-ante credit and liquidity enhancement as well as any type of ex-post support has been prohibited. However, in our setup, contagion from the shadow banking sector to the banking sector may take place even in the absence of any contractual linkages between the two sectors. Thus, financial stability of the banking sector may not be achieved by simply prohibiting contractual linkages of the regulated banking sector with allegedly other risky activities.

5.3. Equilibrium Size of the Shadow Banking Sector

If runs are zero probability events, the contract offered by shadow banks strictly dominates the contract offered by regulated banks, as \((c_{sb}^1, c_{sb}^2) \succ (c_b^1, c_b^2)\). In expectation, fire-sale pricing does not occur, and no expected benefit is retrieved from the deposit insurance.
Therefore, let us now assume that agents assign a positive probability to runs taking place in the shadow banking sector. We do so by assuming that agents coordinate their behavior depending on the realization of a sunspot as in Cooper and Ross (1998) and Gertler and Kiyotaki (2015). Generally, it would be desirable to pin down the probability of a run is by relaxing the assumption of common knowledge and using global games techniques as pioneered by Morris and Shin (1998, 2003) and applied to bank runs by Rochet and Vives (2004) and Goldstein and Pauzner (2005). Under this approach, the run probabilities are tied to the fundamentals. However, introducing aggregate uncertainty in our model would introduce many other issues and thus shift the focus away from liquidity crises. Therefore, we make the simple assumption that consumers coordinate with an exogenous probability.

**Assumption 2.** With probability $q$ agents coordinate on withdrawing from the shadow banking sector in $t = 1$.

Moreover, we assume that the contracts offered by banks and shadow banks are not altered by the fact that agents can anticipate runs. It is important to notice the contracts may not actually be the optimal contracts once we assign positive probabilities to a run in the shadow banking sector. Thus, the analysis is as if consumers anticipate runs, but intermediaries do not. We discuss the importance of this assumption in Section 2.

Given the deposit insurance scheme, the payoff of a bank deposit remains at $(c_{b1}^*, c_{b2}^*)$ in any contingency. The expected utility is thus given by $EU_b = U(c_{b1}^*, c_{b2}^*)$. In contrast, the payoff of shadow bank customers changes in case of a run, and is given by $(\tilde{c}_{sb1}(\sigma), \tilde{c}_{sb2}(\sigma))$ where

$$(\tilde{c}_{sb1}(\sigma), \tilde{c}_{sb2}(\sigma)) = \begin{cases} (R/\gamma, R/\gamma) & \text{if } \sigma < \bar{\sigma} \\ (p(\sigma), p(\sigma)) & \text{otherwise.} \end{cases}$$

Whenever the shadow banking sector is small ($\sigma < \bar{\sigma}$), shadow banks will be able to sell all their entire assets at $p = R/\gamma$ and distribute the proceeds among their depositors. In turn, for $\sigma > \bar{\sigma}$, runs will be systemic and cash-in-the-market pricing prevails, and thus only $p(\sigma)$ can be distributed among the claimholders. Therefore, $EU_{sb}(\sigma)$ is decreasing in the size of the shadow banking sector $\sigma$, as a larger size implies a lower fire sale price in case of a run.

The expected utility of depositing in the shadow banking sector is thus given by:

$$EU_{sb}(\sigma) = (1 - q)U(c_{b1}^*, c_{b2}^*) + qU(\tilde{c}_{sb1}(\sigma), \tilde{c}_{sb2}(\sigma)).$$

In equilibrium, consumers have to be indifferent in $t = 0$ between holding a regular bank deposit or a shadow bank deposit and thus the equilibrium size of the shadow
banking sector is given by $\sigma^*$, such that

$$EU_b = EU_{sb}(\sigma) \iff \text{(18)}$$

We can derive the following proposition:

**Proposition 4.** It holds that $\sigma^* > \bar{\sigma}$ if and only if $q \leq \bar{q}$, where

$$\bar{q} = \frac{U(c^*_1, c^*_2) - U(c^b_1, c^b_2)}{U(c^*_1, c^*_2) - U(R/\gamma)}.$$  

Moreover, it holds that

$$\frac{\partial \sigma^*}{\partial \rho} > 0, \quad \text{and} \quad \frac{\partial \sigma^*}{\partial q} \leq 0.$$  

The proof can be found in the appendix.

In equilibrium, consumers need to be indifferent between holding deposits at banks and shadow banks. The equilibrium is attained via the following mechanism: The fire sale price is decreasing in the amount of assets that are sold in case of a fire sale and thus in the mass of people holding claims on the shadow banking sector. The equilibrium is attained when sufficiently many agents are in the shadow banking sector such that the expected fire sale is sufficiently low and thus the expected utility of holding a shadow bank deposit equals the expected utility in the regulated banking sector.

Moreover, we have that the equilibrium size is weakly decreasing in the sunspot probability (see Figure 3) and decreasing in the required return of intermediaries. The intuition for both results is straightforward: The more costly intermediary capital is, the less attractive regulated banks become. Moreover, in equilibrium this also implies that more consumers move in the shadow banking sector, leading to lower expected fire sale prices. Thus, shadow banks become less attractive as well. Moreover, the more likely runs are, the less attractive is the shadow banking sector and thus the shadow banking sector shrinks. At the extreme ends, when runs are either very likely or very unlikely, either all financing is intermediated via shadow banks or there is no shadow banking sector altogether.

### 5.4. Social Optimum

We can contrast the equilibrium size of the shadow banking sector of Proposition 4 with the size the social planner would choose. The social planner in contrast to individuals can internalize the effect of the size of the shadow banking sector on the fire sale price. In turn, the fire sale price will affect the expected utility in the shadow banking sector via
Figure 3: This graph depicts the size of the shadow banking $\sigma$ sector as a function of the sunspot probability $q$.

The fire sale price directly, and the expected cost of the deposit insurance that arise from the deterioration of funding conditions. We calculate the expected cost for the provider of the the deposit insurance, $DI(\sigma)$ in Appendix G. Importantly, the cost imposed on the deposit insurance scheme are increasing in $\sigma$. The logic is as in Proposition 3: the larger the shadow banking sector is, the lower fire sales will be in case of a run and thus the stronger the deterioration of funding conditions for commercial banks and thus the larger the shortfall.

We assume that the social planner solves the following problem:

$$\max_{\sigma} (1 - \sigma)[EU_b] + \sigma[EU_{sb}(\sigma)] - qDI(\sigma)$$

I.e., we assume that the social planner maximized the weighted sum of utilities and values the cost that accrue in case of a run for providing the deposit insurance scheme additively separate and linearly.

The corresponding first-order conditions imply that the social planner chooses $\sigma^{SP}$.

---

A more complex model would determine the optimal tax scheme to finance the deposit insurance. While the results would be quantitatively different, there would be no qualitative differences. For simplicity, we abstract from such considerations.
such that
\[
\frac{EU_b}{EU_{sb}(\sigma)} = 1 + \frac{1}{EU_{sb}(\sigma)} \left[ \sigma q U_\sigma(c_{1}^{sb}(\sigma), c_{2}^{sb}(\sigma)) - q DI'(q) \right]
\]
\[
< 0
\]
(19) \{eq:FOC_SP\}

We can compare the FOC to the condition for the equilibrium size of the shadow banking sector. Rewriting Equation (18), we have that \( \sigma^* \) satisfies
\[
\frac{EU_b}{EU_{sb}(\sigma)} = 1
\]
(20) \{eq:equilibrium\}

We thus can see that the LHS of the social planner’s FOC condition is lower than the one determining the equilibrium. Given that \( EU_{sb}(\sigma) \) is decreasing in \( \sigma \), the social planner thus prefers to choose a smaller \( \sigma \):

**Proposition 5.** It holds that
\[
\sigma^* > \sigma^{SP}.
\]
{prop:socialplanner}

The logic underlying the result is simple: unlike private agents that are atomistic, the social planner is able to internalize the effect of \( \sigma \) on the fire sale price as well as the effect of the fire sale price on the expected cost for funding the deposit insurance scheme. The social planner would thus like to choose a lower \( \sigma \) which would give those agents that hold deposits in the shadow banking sector a higher utility than those in the banking sector.

In equilibrium, the shadow banking sector is thus too large for two reasons: First, the fire sales in case of a run are too low as too many agents are in the shadow banking sector. This is a standard pecuniary externality as in Lorenzoni (2008). Moreover, this increases the expected costs of the deposit insurance scheme as lower fire sale prices also imply a deterioration of funding conditions for regulated banks, that may thus not be able to fulfill their obligations.

The social planners’s choice Pareto dominates the equilibrium allocation. Agents that would remain in the shadow banking sector would receive a higher utility while those in the banking sector would still receive the same payoff.

**6. Wholesale Funding Restrictions**

Under the premise that the social planner cannot control the extent of regulatory arbitrage \( \sigma \), a natural question that arises is whether limiting wholesale funding of commercial banks can improve financial stability. This question is particularly important as the
proposed liquidity regulation acts like a restriction on the wholesale funding capacity of banks, as discussed in Appendix I.

In the following, we analyze the most simple and the most extreme case: a complete prohibition of wholesale funding for regulated banks, i.e., a restriction that \( L = 0 \) for regulated banks. We show that this shuts down the contagion channel, but the associated allocation is less efficient and implies a further growth of the shadow banking sector.

The optimal consumption profile a bank can offer when wholesale funding is prohibited is identical to the second best for the case of \( \gamma = R \). In this latter case, the bank could sell assets in the interim period, however, the return of doing so is identical to using storage. The FOC is given by \( u'(c_1) = Ru'(c_2) \).

**Lemma 4.** The constrained optimal allocation under the prohibition of wholesale funding requires an investment of

\[
I^r = 1 + \frac{B}{\rho - B} - \pi c_1^r,
\]

and the consumption profile is given by

\[
c_1^r = R^{-\frac{1}{\eta}} R - \frac{B}{\rho - B} (\rho - R), \quad \text{and} \quad c_2^r = \frac{R - \frac{B}{\rho - B} (\rho - R)}{\pi R^{1 - \frac{1}{\eta}} + (1 - \pi)}.
\]

It holds that \( c_t^b > c_t^r \), and thus

\[
U(c_1^b, c_2^b) > U(c_1^r, c_2^r).
\]

Note that this optimal contract under the exclusion of wholesale funding \((c_1^r, c_2^r)\) coincides with the first-best allocation of the Diamond and Dybvig model for \( B = 0 \) or \( \rho = R \). However, this allocation is not efficient in the presence of investors that can provide liquid funds in the interim period.

Shadow banks still offer the same contract as above, i.e., \((c_1^{sb}, c_2^{sb}) = (c_1^*, c_2^*)\). The size of the shadow banking sector \( \sigma^r \) will now be determined by

\[
U(c_1^r, c_2^r) = (1 - q)U(c_1^*, c_2^*) + qU(\tilde{c}_1^{sb}(\sigma), \tilde{c}_2^{sb}(\sigma))
\]

**Proposition 6.** Wholesale funding restrictions successfully shield regulated banks from the adverse consequences of runs in the shadow banking sector. However, the shadow banking sector is larger than without these restrictions,

\[
\sigma^r > \sigma^* > \sigma^{SP}.
\]
Wholesale funding restrictions, together with a deposit insurance, eliminate the fragility in the banking sector altogether. They do, however, induce a further allocative inefficiency by deteriorating the consumption profile of bank customers and by pushing more depositors into the shadow banking sector.

**Welfare Implications**

The answer to the question whether a wholesale funding restriction ultimately leads to higher or lower welfare level is not straightforward. Clearly, a wholesale funding restriction reduces the direct utility agents receive in banking as well as in the shadow banking sector. However, the social planner must trade-off these losses against the fact that he does not need to finance the deposit insurance scheme. Therefore, the effect of wholesale funding restrictions on welfare crucially depend on how the cost of the deposit insurance scheme are valued. In 5.4 we assumed that they enter the social planners objective additively separate and linearly. However, in order to consider more complex trade-offs, one needs a more general model in which the deposit insurance scheme is financed via a tax and thus imposed on agents. The actual tax scheme would then determine whether restrictions are desirable or not.

**7. Liquidity Guarantees**

Until this point, we have restricted attention to intermediaries becoming either regulated banks or unregulated shadow banks. An interesting question is how our results change when banks and shadow banks are interdependent not only via effects on secondary markets, but if they are operated by the same intermediary. As indicated in the introduction, this had been practice prior to the recent financial crisis, as documented by Acharya et al. (2013), and has been targeted by post-crisis reforms.

To this end, we analyze a version of our model in which intermediaries operate a bank and a shadow bank at the same time. We investigate under which conditions intermediaries may have incentives to use funds from their regulated banking branch to support their shadow-banking activities in case of distress. We will thus analyze the effect of explicit contractual linkages.

**Private Optimality of Liquidity Guarantees**

In the previous section, we assumed that a systemic run in the shadow banking sector is a sunspot phenomenon that occurs with a probability $q$. Moreover, we assume that shadow banks experience idiosyncratic sunspot runs: With a probability $q_i$, each individual shadow bank experiences a run. Again we keep the contracts offered fixed. Again, a commercial bank optimally offers $(c_{1b}^k, c_{2b}^k)$ as before, and a shadow bank $(c_{1sb}^{sh}, c_{2sb}^{sh})$. 
Given that there is a run on the shadow bank with probability \( q_i \), the intermediary may now have an incentive to guarantee the liquidity of her shadow bank to protect her from idiosyncratic runs. Observe that the possibility of an idiosyncratic sunspot run can be eliminated if a regulated branch of an intermediary provides a credible liquidity guarantee for its unregulated operations. Moreover, observe that it is optimal to provide this support guarantee for each institution as it makes the offered contract more attractive and will thus attract more consumers.

The liquidity guarantee is credible if the bank can serve all its impatient bank customers as well as all those who own a shadow bank contract by selling all its assets. The regulated banking sector has conducted an initial investment of \( (1 - \sigma)(1 + e_0^*) \) and the shadow banking an investment of \( \sigma \). In total, the bank can thus raise \( R/\gamma(1 + (1 - \sigma)e_0^*) \).

The funds raised by selling all assets need to be sufficient to serve all impatient consumers, \( (1 - \sigma)c_{b1}^1 \), of the regulated bank and all consumers of the shadow bank, \( \sigma c_{sb1}^1 \). A liquidity guarantee is thus credible whenever

\[
\sigma \leq \frac{R/\gamma(1 + (1 - \sigma)e_0^*) - \pi c_{b1}^1}{c_{b1}^1 - \pi c_{b1}^1}
\]

A guarantee can thus only be credible if the shadow banking operations are not too large compared to the regulated banking activities.

**Systemic Runs**

While it is optimal from the perspective of a single institution to provide a liquidity guarantee, it leads to an increased parameter space for runs on the aggregate level. In case of a run, commercial banks have to provide an amount of \( \sigma c_{sb1}^1 \) to shadow banks. In addition, they require an amount of \( (1 - \sigma)c_{b1}^1 \) to satisfy their own impatient customers. Therefore, a run is systemic whenever \( \sigma c_{sb1}^1 + (1 - \sigma)c_{b1}^1 > A \).

**Proposition 7.** Assume that intermediaries can operate a regulated bank and a shadow bank at the same time and that liquidity guarantees are credible. It is privately optimal for each intermediary to guarantee the liquidity for her shadow bank branch by using funds from her regulated bank. In turn, this decreases the threshold size above which the shadow banking becomes systemic: A systemic run can now already occur and affect regulated banks if

\[
\sigma \geq \frac{A - \pi c_{b1}^1}{c_{b1}^b - \pi c_{b1}^1} < g(A).
\]

Without liquidity guarantees, systemic runs are only possible if the shadow banking sectors size exceeds \( \bar{\sigma} \). With liquidity guarantees, this is already true for a sector size $28$.}{prop:guarantee}
of $\sigma \geq \frac{A - \pi c_1}{\pi c_1 - \pi c_2}$. The underlying mechanism is as follows: while liquidity guarantees are optimal from the individual intermediary’s perspective, they increase the number of assets sold in case of a systemic run in the shadow banking sector. This shows that there is a clear benefit of preventing direct contractual linkages via regulation, as it reduces the parameter space in which systemic runs may take place. However, as shown in the earlier section, it may not be sufficient to rule out adverse effect for regulated banks entirely.

8. Lender of Last Resort and Market Maker of Last Resort

So far, we have only considered a deposit insurance and capital regulation as policy measures, and we have ignored other government interventions. In practice, regulated banks also have access to the discount window of a central bank (Lender of Last Resort, LoLR). Alternatively, the central bank may decide to intervene in the secondary market in order to prevent cash-in-the-market pricing (Market Maker of Last Resort, MMLR). We now want to ask whether the possibility of such interventions may help to shield commercial banks from turmoil in the shadow banking sector.

Assume that there is an institution called central bank that has unlimited funds at its disposal in the interim period and can commit to being a LoLR or a MMLR. In a richer setup that distinguishes between nominal and real values, one could also consider potential costs of central bank interventions. We abstract from such trade-offs in our analysis and briefly discuss them at the end of this section.

LOLR

Let us consider a central bank that acts as a LoLR or a MMLR, and analyze the policy it would choose in order to protect regulated banks from systemic crises. In our setup, a LoLR would intervene in case of a systemic crisis (i.e., if $p < R/\gamma$). Regulated banks would be allowed to use the discount window to borrow funds, so they do not have to sell assets at fire-sale prices. If we assume that the discount window lending has the same terms as the wholesale funding in normal times, then LoLR effectively shields regulated banks from turmoil in the shadow banking sector. Thus, even there may be runs in the shadow banking sector, the deposit insurance scheme will never be used. However, shadow banks also benefit from such intervention, although they do not have access to the discount window. Because regulated banks do not have to sell their assets, fewer assets are on the market compared to a situation without central bank intervention, implying that fire-sale prices are higher. Thus, the equilibrium size of the shadow banking sector will be larger.
MMLR

A MMLR in our setup could ensure that the price of assets is always equal to \( R/\gamma \) by buying assets at this price in case of a systemic crisis. In contrast to the LoLR’s discount window lending, the MMLR conducts an open market operation, and it appears plausible to assume that the MMLR cannot selectively buy assets from regulated banks only. As in the case of the LoLR, regulated banks are again unaffected by runs in the shadow banking sector and can offer the second-best. Shadow bank customers again benefit from the intervention, and in this case even more than in case of a LoLR intervention because the MMLR eliminates cash-in-the-market pricing altogether. Thus, the shadow banking sector will also increase under a MMLR policy. Moreover, the central bank will end up purchasing actual assets whenever there is a run.

It is worth noticing that both the LoLR and the MMLR as discussed, are effective in shielding regulated banks, but they do not prevent runs in the shadow banking sector as the strategic complementary among consumers prevails independent of the wholesale funding conditions.

Therefore, we conclude that, for both types of interventions, it holds true that they make the shadow banking sector relatively more attractive in the presence of positive probabilities for runs and thus lead to an increased shadow banking sector; and this effect is more pronounced for the MMLR than for the LoLR intervention. Still, a LoLR or a MMLR that shields regulated banks from the adverse consequences of runs on shadow banks could be considered as an alternative to restrictions on wholesale funding. However, as mentioned above, our model is not suited to draw such conclusions in general, as we do not consider the costs and potential distortions resulting from interventions. While interventions have to be financed via taxation or inflation, there would be a trade-off in a richer setup that allows to distinguish between nominal and real values. Moreover, an ex-post intervention by a LoLR or a MMLR may not always be desirable in richer setups because of other potential frictions, e.g., if the government cannot distinguish between insolvent and illiquid banks or if interventions give rise to moral hazard.

9. Literature

Our paper connects to the recent literature on theoretical aspects of shadow banking. The paper by Bolton et al. (2011) presents a model in which intermediaries originate assets, and can sell them on a secondary market to raise “outside liquidity”. Similar to their paper, we discuss how the presence of such outside liquidity may affect the business model of banks. However, we abstract from adverse selection and emphasize the role of
coordination.

Our modeling approach is also related to the papers by Martin et al. (2014a), Martin et al. (2014b), and Luck and Schempp (2014a). All three rely on the setup of Qi (1994). Martin et al. (2014a) investigate the differences between bilateral and tri-party repo in determining the stability of single financial institutions. Martin et al. (2014b) focus on the difference between runs on single institutions and systemic runs in secured funding markets. While our model has a similar notion of cash-in-the-market pricing, our focus lies on the effects of regulation and in particular on the adverse effects of shadow banking on regulated banks. In this respect, the paper is close to Luck and Schempp (2014a), who analyze the effects of shadow banking on financial stability in the presence of contractual linkages between the two sectors. The crucial difference to Luck and Schempp (2014a) is that this paper focuses on contagion in the absence of contractual linkages.

Segura (2014) explicitly discusses the role of liquidity guarantees. In particular, he gives a theoretical answer to the question posed by Acharya et al. (2013) on why banks supported their SIVs despite the absence of direct contractual linkages. While Segura elegantly shows how regulated banks may have incentives to support their shadow bank operation in order to preserve their reputation, we argue that there may be reasons to believe that regulatory arbitrage may affect stability of regulated banks beyond such contractual linkages.

Other important contributions that deal with shadow banking are Ordoñez (2013), Gennaioli et al. (2013), Plantin (2014), and Hanson et al. (2015). Gennaioli et al. (2013) provide a model in which the demand for safe debt drives securitization, and fragility in the shadow banking sector arises when tail-risk is neglected. Ordoñez focuses on potential moral hazard on the part of banks. In his model, shadow banking is potentially welfare-enhancing as it allows one to circumvent imperfect regulation. However, it is only stable if shadow banks value their reputation and thus behave diligently; it becomes fragile otherwise. Plantin studies the optimal prudential capital regulation when regulatory arbitrage is possible. In his paper, relaxed capital requirements lead to a decline of the shadow banking sector, potentially improving welfare. In contrast to all three, we focus on the destabilizing effects of shadow banking in the sense that it gives rise to run equilibria. Finally, Hanson et al. argue that the fact that regulated banks are subject to capital requirements and covered by the deposit insurance allows them to hold relatively illiquid assets whereas shadow banks are subject to runs and thus hold relatively liquid assets.
10. Discussion

This paper provides a banking model in which intermediation optimally relies on secured
wholesale funding next to demand deposits as a source of funding. We show that in an
unregulated banking environment, short-term debt is disciplining, but panic-based runs
are also possible. If a safety net is in place to eliminate the possibility of panic-based runs,
regulation becomes necessary. This in turn leads to the emergence of an unregulated and
fragile shadow banking sector. We then show that because regulated banks optimally
rely on wholesale funding, runs on shadow banks are contagious via a deterioration of
the wholesale funding conditions. Therefore, turmoil in the shadow banking sector may
affect the commercial banking sector even in the absence of any contractual linkages
between the two sectors.

We emphasize two main findings of our analysis: First, we argue that prohibiting
contractual linkages is not sufficient to shield the regulated banking sector from financial
fragility. Naturally, commercial banks and shadow banks hold similar assets and rely
on similar sources of funding. Thus, a turmoil in the shadow banking sector is likely to
have an impact on the regulated banking sector even in the absence of any contractual
linkages. Second, we argue that restricting wholesale funding activities of commercial
banks may (as proposed in the Third Basel Accords liquidity regulation) may be an
effective way of shielding commercial banks from turmoil that originates outside the
regulated banking sector. However, it also leads to allocative efficiency in the regulated
banking sector and stronger fire sale effects in the shadow banking sector. The welfare
effects are thus ambiguous and depend crucially on the cost of funding the deposit
insurance scheme in case of a crisis.
Appendix A  First-Best

We start out with the following first-best maximization program:

\[
\begin{align*}
\max_{(c_1, c_2, e_0, e_2, I, L, p) \in \mathbb{R}_+^7} & \quad \pi u(c_1) + (1 - \pi) u(c_2), \\
\text{subject to} & \quad \pi c_1 \leq (1 + e_0 - I) + Lp, \\
& \quad (1 - \pi)c_2 \leq (I - L)R - e_2, \\
& \quad e_2 \geq \rho e_0 \geq 0, \\
& \quad R \geq \gamma p, \\
& \quad pL \leq A, \\
& \quad I \leq 1 + e_0, \\
& \quad L \leq I.
\end{align*}
\]  

The budget constraints for periods one and two are given by (22) and (23). Investors may transfer \(Lp\) to consumers in \(t = 1\) in exchange for \(L\) units in \(t = 2\). As indicated above, we refer to this as wholesale funding. (24) represents the participation constraint of the intermediary and non-negativity constraint, and (25) represents the participation constraint of investors. The resource constraint on investors’ capital \(A\) (“interim liquidity”) in the interim period is given by (26). Finally, (27) and (28) denote the constraint on initial investment as well as the constraint on the units of assets that can be sold in the interim period.

The constraint (28) cannot be binding because the Inada conditions require that \(c_2 > 0\) and thus \(L < I\). The two budget constraints are always binding. Furthermore, the participation constraint of intermediaries must also be binding. Moreover, \(e_0 > 0\) cannot be optimal. Because \(\rho > R\), we can reduce \(e_0\) and thus relax the second-period constraint. It therefore follows that \(e_0^* = e_2^* = 0\). Because the intermediaries’ required return is higher than the asset return, intermediaries’ funds are not used for intermediation in the first-best. As we will see below, moral hazard may make it necessary to force the intermediary to invest some of his endowment.

Let us now turn towards the use of interim liquidity, i.e., wholesale funding in period 1. In the first-best, it also has to hold that the participation constraint of investors is binding. Whenever \(p < R/\gamma\) and \(pL < A\), we can increase \(p\) and thereby relax the period 1 constraint. Whenever \(p < R/\gamma\) and \(pL = A\), we can increase \(p\) and decrease \(L\) as much as necessary, thereby relaxing the period 2 constraint. Therefore, it holds that \(p = R/\gamma\).

We are now left with a maximization problem with two weak inequalities.
max_{(c_1,c_2,I,L)\in \mathbb{R}_+^4} \pi u(c_1) + (1 - \pi)u(c_2), \quad (29) \{eq:max_b\}

subject to \pi c_1 = (1 - I) + LR/\gamma, \quad (30) \{eq:period1_b\}
(1 - \pi)c_2 = (I - L)R, \quad (31) \{eq:period2_b\}
LR/\gamma \leq A, \quad (32) \{eq:arbitrage_budget_b\}
I \leq 1. \quad (33) \{eq:investment_limit_b\}

Let us define the following thresholds:

$$\xi \equiv \gamma^{-\frac{1}{\eta}} \frac{\pi R}{1 - \pi + \pi \gamma^{1 - \frac{1}{\eta}}} < R/\gamma,$$ \quad (34)

and

$$\xi_0 \equiv \frac{\pi R^{-\frac{1}{\eta}}}{(1 - \pi + \pi R^{-\frac{1}{\eta}})} < \xi, \quad (35)$$

Depending on the model parameters $A, R, \gamma,$ and $\pi,$ as well as on the shape of the utility function, the first-best program now has three solution candidates.

**Lemma 5** (First-best). If $A \geq \xi,$ then the first-best allocation is characterized by

$$I^* = 1 \text{ and } L^* = \xi \gamma / R,$$

and optimal consumption is given by

$$c_1^* = \gamma^{-\frac{1}{\eta}} \frac{R}{1 - \pi + \pi \gamma^{1 - \frac{1}{\eta}}} \text{ and } c_2^* = \frac{R}{1 - \pi + \pi \gamma^{1 - \frac{1}{\eta}}}.$$ \quad (36)

For $A \in (\xi_0, \xi),$ we have that

$$I^* = 1 \text{ and } L^* = A \gamma / R,$$ \quad (37)

and optimal consumption is given by

$$c_1^* = \frac{A}{\pi} \text{ and } c_2^* = \frac{R - A \gamma}{1 - \pi}. \quad (38)$$

Finally, if $A \leq \xi_0,$ then the first-best allocation is characterized by

$$I^* = \frac{(1 - \pi)(1 + A) + \pi R^{-\frac{1}{\eta}} A \gamma}{(1 - \pi) + \pi R^{-\frac{1}{\eta}}} \text{ and } L^* = A \gamma / R,$$ \quad (39)

and optimal consumption is given by

$$c_1^* = R^{-\frac{1}{\eta}} \frac{R + (R - \gamma)A}{1 - \pi + \pi R^{-\frac{1}{\eta}}} \text{ and } c_2^* = \frac{R + (R - \gamma)A}{1 - \pi + \pi R^{-\frac{1}{\eta}}}.$$ \quad (40)
In the first case \((A \geq \xi)\), it holds that \(I^* = 1\) and \(L^* < 1\), and the optimal allocation is characterized by

\[
u'(c_1) = \gamma u'(c_2) \quad \text{(41)}
\]

\[
\pi c_1 \gamma + (1 - \pi) c_2 = R. \quad \text{(42)}
\]

In the third case \((A < \xi_0)\), we have \(L^* = A\gamma/R\), and \(I^* < 1\), and the optimal allocation is characterized by

\[
u'(c_1) = Ru'(c_2) \quad \text{(43)}
\]

\[
\pi c_1 R + (1 - \pi) c_2 = R + (R - \gamma)A. \quad \text{(44)}
\]

**Appendix B Shirking Equilibria**

We have stated above that in a subgame where the intermediary has chosen \(I_{shirk} = 1\), it is the dominant strategy of all consumers to withdraw their funds early, so this constitutes the unique Nash equilibrium of this subgame. This illustrates the potential for short-term debt to be disciplining. Because the intermediary enjoys the benefit only when he does not experience a bank run in the interim period, there exists a run strategies in which consumers credibly threaten to punish any shirking. Thus, there exists a subgame-perfect Nash equilibrium in which the intermediary does not shirk, but implements the first-best.

However, as we pointed out in the main text, there also exist subgame-perfect Nash equilibria in which that short-term debt is only partially disciplining. In fact, some investment in the shirking technology may be tolerated by depositors before they run. The reason is that if the intermediary has chosen some \(I_{shirk} \geq 0\) that is small enough, “running” is not the dominant strategy.

In order to illustrate this, assume that \(R_{shirk} = 0\). Assume the contract \((c_1^*, c_2^*)\) has been promised. We can ask what level of \(I_{shirk}\) the intermediary can choose before consumers will run, irrespectively of the behavior of other in the interim period. We have that the following two budget constraints hold:

\[
\pi c_1^* = LR/\gamma
\]

\[
(1 - \pi)c_2 = (1 - I_{shirk})R - LR
\]

Consumers will run after observing \(I_{shirk}\) whenever \(c_2 < c_1\). That is, whenever

\[
c_2 = \frac{1}{1 - \pi} \left( (1 - I_{shirk})R - \frac{\pi c_1^* \gamma}{R} \right) < c_1^*
\]
\[ \Leftrightarrow I_{\text{shirk}} > 1 - \frac{c_1^* (1 - \pi) + \pi c_1^* \gamma}{R} \equiv I_x \]

Therefore, there exist multiple subgame-perfect Nash equilibria in pure strategies which differ in the extent of the discipline that they ensure. In particular, for any \( \Psi \in [0, I_x] \), there exist an equilibrium in which each consumer runs if and only if \( I_{\text{shirk}} > \Psi \), and intermediaries choose \( I_{\text{shirk}} = \Psi \). In the main part of our analysis, we ignore such equilibria, which is equivalent to assuming that the intermediaries can only invest in either the productive or the shirking technology, i.e., \( I_{\text{Shirk}} \in \{0, 1\} \).

**Appendix C  Financial Markets Implementation**

We have by assumption ignored the possibility of implementing an allocation via a financial market instead via intermediaries. The allocation that can be attained via a financial market in which consumers invest in the technologies directly and trade with investors in \( t = 1 \), is \((c_{1m}^{fm}, c_{2m}^{fm}) = (R/\gamma, R)\). This allocation, however, only coincides with the first-best if \( \eta \to 1 \), i.e., if \( u(c) = \ln(c) \). This is reminiscent of the result of Diamond and Dybvig (1983).

Nonetheless, we need to make the investment restriction: If we allowed for the co-existence of financial markets and intermediaries, the incentive to conduct side trading would destroy the ability to implement the first-best via intermediaries, due to the same reasoning as in Jacklin (1987) and Farhi et al. (2009). If intermediaries offered the first-best demand-deposit contract, a consumer has an incentive to invest his endowment in the productive technology and consume the returns \( R > c_2^* \) if he turns out to be patient, and to trade with a patient depositor otherwise, thereby consuming \( c_1^* \) units.

**Appendix D  Second-Best**

In the presence of equity regulation, the second-best is the solution of the maximization problem of the first-best subject to the equity requirement. If the investors’ capital \( A \) is not scarce (i.e., if \( pL < A \)), it also holds that \( L < I \), and the following constraints are binding:
\[
\pi c_1 = (1 + e_0 - I) + Lp,
\]
\[
(1 - \pi)c_2 = (I - L)R - e_2,
\]
\[
e_0 = \frac{B}{\rho - B},
\]
\[
e_2 = \rho e_0,
\]
\[
p = \frac{R}{\gamma},
\]
\[
I = 1 + e_0.
\]

The maximization problem is reduced to

\[
\max_{(c_1, c_2, L)} \pi u(c_1) + (1 - \pi)u(c_2),
\]
subject to \( \pi c_1 = LR/\gamma, \)
\[
(1 - \pi)c_2 = (1 + e_0 - L)R - \rho e_0,
\]
\[
e_0 = \frac{B}{\rho - B}.
\]

As the utility function is characterized by CRRA with the RA-parameter \( \eta \), the FOC can be rewritten as
\[
c_1 = \gamma^{-\frac{1}{\eta}} c_2,
\]
and the period budget equations yield the following consumption levels:

\[
c_1 = \frac{RL}{\pi \gamma}, \quad \text{and}
\]
\[
c_2 = \frac{1}{1 - \pi} [(1 + e_0 - L)R - \rho e_0].
\]

**Appendix E  Optimal Contracts with Sunspot Runs**

**Appendix F  Proof of Proposition 4**

*Proof.* We want to show that it holds that \( \sigma^* > \bar{\sigma} > 0 \) if and only if \( q \leq \bar{q} \), where

\[
\bar{q} = \frac{U(c_1^*, c_2^*) - U(c_1^b, c_2^b)}{U(c_1^*, c_2^*) - U(R/\gamma)}
\]

In equilibrium: \( \sigma^* \)
\[
EU_b = EU_{sb}(\sigma) \iff U(c_1^b, c_2^b) = (1 - q)U(c_1^*, c_2^*) + qU(c_1^{sb}(\sigma), c_2^{sb}(\sigma))
\]

37
EU_b is independent of \( \sigma^* \), while \( EU_{sb}(\sigma) \) is weakly decreasing in \( \sigma \). Recall the fire sale price:

\[
p(\sigma) = \begin{cases} 
R/\gamma & \text{if } \sigma \leq \bar{\sigma} \\
\frac{A-(1-\sigma)\pi c_1^b}{\sigma} & \text{if } \sigma > \bar{\sigma}.
\end{cases}
\]

For any \( \sigma \) less than \( \bar{\sigma} \), the price is not changing in \( \sigma \), but decreasing afterward. Therefore, if \( \sigma = 0 \), but

\[
U(c_1^b, c_2^b) < (1-q)U(c_1^*, c_2^*) + qU(c_{sb 1}^b(\sigma), c_{sb 2}(0))
\]

it must hold in equilibrium that \( \sigma > \bar{\sigma} \) to satisfy the above equilibrium conditions. Thus, \( \sigma^* > \bar{\sigma} \), whenever \( q \leq \bar{q} \).

\[\square\]

**Appendix G  Cost of Deposit Insurance**

From Proposition 3 we know that a run in the shadow banking sector will be contagious thus affecting the regulated banking sector. This implies that a run in the shadow banking sector may potentially lead to illiquidity or insolvency of the regulated banks and thus impose costs on the provider of the deposit insurance. In the following we quantify these costs.

Recall the specification of the deposit insurance as described in Section 4: the deposit insurance guarantees a regular bank depositor to receive \( (c_1^b, c_2^b) \). If the fire sale price falls below \( R/\gamma \), a bank will have to sell more assets on the secondary market in order to pay out \( c_1^b \) to all \( \pi \) early consumers. Each commercial bank must sell amount \( L > L^* \) such that

\[
p(\sigma)L = \pi c_1^b
\]

This implies that there may be a potential shortfall of funds in \( t = 2 \) that need to be injected by the provider of the deposit insurance of:

\[
(1-\pi)c_1^b - (1+e^* - \frac{\pi c_1^b}{p(\sigma)})R
\]

However, if the bank turns already illiquid in \( t = 1 \), i.e. selling all assets is not sufficient to pay out early consumer, i.e. \( I = 1+e^* \) and \( p(\sigma) < \pi c_1^b \), the deposit insurance requires a funding of

\[
(1-\pi)c_2^b + \pi c_1^b - p(\sigma)
\]

In total, the funds required as a function of the fire sale price are given by:

\[
DI(\sigma) = (1-\sigma) \max \left[ (1-\pi)c_2^b + \pi c_1^b - p(\sigma), (1-\pi)c_1^{**} - (1+e^* - \frac{\pi c_1^b}{p(\sigma)})R, 0 \right]
\]

38
Importantly, the cost of the deposit insurance is decreasing in the fire sale price and thus increasing in the size of the shadow banking sector $\sigma$.

### Appendix H  Wholesale Funding Restrictions

In the case that wholesale funding is forbidden, commercial banks can only offer an allocation where liquidity is provided by investing in storage. The optimal allocation is the solution of the following problem:

\[
\max_{(c_1,c_2,I)} \pi u(c_1) + (1 - \pi)u(c_2), \quad \text{(45)}
\]

subject to

\[
\pi c_1 = (1 + e_0 - I), \quad \text{(46)}
\]
\[
(1 - \pi)c_2 = IR - \rho e_0, \quad \text{(47)}
\]
\[
e_0 = \frac{B}{\rho - B}. \quad \text{(48)}
\]

We see that if we set $\rho = R$ in the second-best problems, the two optimization problems are identical in terms of the consumption allocation they imply, and it holds that

\[
L^{**} = (1 + e_0 - I^r), \quad \text{(49)}
\]

i.e., the liquidation $L^{**}$ in the second best is equal to the investment in storage under wholesale funding restrictions.

### Appendix I  Liquidity Regulation in Basel III

In the following, we argue that the liquidity regulation in Third Basel Accord (Basel III) is isomorphic to a wholesale funding restriction in our model. Therefore, our model can shed light on the economic consequences of liquidity regulation as proposed.

Basel III introduces a new assessment and regulation of liquidity risk by defining two minimum standards of funding liquidity, first described in Basel Committee (2010). The two central measures are the Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR). The LCR requirement aims to ensure that a bank can withstand a “significantly severe liquidity stress scenario” with a horizon of 30 days and is described in detail in Basel Committee (2013). It aims to ensure that a bank has a sufficient stock of liquid assets in order to cover its liquidity needs during the next month. The objective of the NSFR requirement, elaborated in Basel Committee (2014), is to ensure stable
funding over a one-year horizon. It requires a bank to have an amount of equity, long-term debt, and other “stable” funding that is sufficient to finance its stock of illiquid assets during the next year.

In the following, we briefly summarize the Basel committee’s proposal for liquidity regulation and analyze its implications in the context of our model.

**Liquidity Coverage Ratio (LCR)**

The LCR is defined as the ratio of High Quality Liquid Assets (HQLA) and the (hypothetical) total net cash outflows over the next 30 calendar days, and Basel III requires this measure to be above 1.\(^{17}\) Thus, the LCR sets a lower bound for the stock of liquid assets, conditional on a bank’s (expected) cash flows. “Total net cash outflow” is defined as the maximum of “total expected net cash outflow” and “25% of total expected cash outflow”. In this context, “expected” denotes a scenario of a “combined idiosyncratic and market-wide shock” that entails (among others) a partial run-off of retail deposits and a partial reduction in unsecured wholesale funding and secured short-term financing.

HQLA consist of two categories: Level 1 assets are cash, central bank reserves, and government bonds with 0% risk weight. Level 2 assets can again be divided in two subcategories, Level 2A and Level 2B assets. A minimum haircut of 15% has to be applied to all Level 2 assets, which is supposed to capture their devaluation in a crisis scenario. After applying this haircut, Level 2 assets must not make up more than 40% of the whole stock of HQLA. Level 2A assets include government bonds with risk weights below 20%, as well as corporate debt securities (including commercial papers) and covered bonds with a rating of at least AA-. Level 2B assets also include Residential Mortgage-Backed Securities (RMBS) with ratings of at least AA, corporate debt securities with ratings of at least BBB-, and common equity shares which are constituent of a major stock index. These assets are subject to a haircut between 25% and 50% and must not make up more than 15% of the stock of HQLA.\(^{18}\)

In addition to these requirements, the Basel Committee specifies liquidity requirements of eligible assets in the following way: Level 2 assets must be

“traded in large, deep and active repo or cash markets characterised by a low level of concentration [and] have a proven record as a reliable source of liquidity in the markets (repo or sale) even during stressed market conditions (i.e. maximum decline of price not exceeding 10% or increase in haircut not exceeding 10 percentage points over a 30-day period during a relevant period of significant liquidity stress).” (Basel Committee, 2013).

\(^{17}\)For details on the LCR, see Basel Committee (2013).

\(^{18}\)Note that they must also be included in the 40% cap of all Level 2 assets.
The requirements for Level 2 assets are thus defined in terms of their past and present liquidity. The underlying notion seems to be that an asset’s past and present liquidity predicts its future liquidity. This is particularly evident in the condition that an asset’s value must have been stable in a “period of significant liquidity stress”.

**Net Stable Funding Ratio (NSFR)**

As a second measure of liquidity regulation, Basel III requires the NSFR to be above 1. The NSFR is defined as the ratio of available stable funding (AFS) and required stable funding (RSF), both with a horizon of one year.  

Thus, the NSFR sets a lower bound for the amount of stable funding, conditional on a bank’s portfolio of illiquid assets and off-balance sheet exposures. The ASF is defined in order to capture the “capital and liabilities expected to be reliable over the time horizon [of] one year”. ASF comprises regulatory capital, preferred stock, and liabilities with maturities of at least one year, but also

Liabilities of the latter categories have to be multiplied by an ASF factor of less than one. ASF aims to exclude unstable short-term funding, i.e., funding that might quickly be withdrawn or not rolled over. It excludes short-term wholesale funding, such as interbank lending, but includes retail deposits, because deposit insurance is supposed to make deposits a source of stable funding.

RSF is a measure of a bank’s illiquid asset portfolio. It is defined as the sum of the value of a bank’s assets, multiplied by a specific RSF factor that should reflect an asset’s liquidity risk, plus a similarly weighted sum of the bank’s off-balance sheet activities or potential liquidity exposures. An asset’s RSF factor is lower the more liquid this asset is. Cash and securities with a maturity of less than one year have an RSF factor of 0%; other securities and corporate bonds with good ratings have low, but positive RSF factors; other bonds, mortgages and loans have higher RSF factors, and other assets (particularly encumbered assets) have RSF factors of 100%.

The notion behind the NSFR requirement is that the ASF serves a bank to finance its illiquid asset contained in the RSF in times of a liquidity crisis. It is assume that those assets not contained in the RSF are liquid and can thus be sold even in times of a crisis in order to compensate the “unstable” funding that might disappear in a crisis.

**Comparison and Discussion**

Although the definitions of the LCR and NSFR appear quite different at first sight, a closer look reveals that their time horizon and the rhetoric are the only distinct differences. To illustrate this point, let us consider a stylized bank balance sheet; see Figure 4.

---

19 For details on the NSFR, see Basel Committee (2014).
The bank’s assets can be divided into a portfolio of liquid and a portfolio of illiquid assets, and the liability side consists of short-term debt, long-term debt, and equity.

If we abstract from the different time horizons of the two liquidity measures, we see that the illiquid assets count as RSF, and the liquid assets count as HQLA. Which asset is considered to be liquid or illiquid is determined by the scenario of stress that is specified by the regulator. On the liability side, the stress scenario specifies which kind of funding is expected to disappear and which is expected to stay during a crisis. The expected net cash outflow measures the expected change in the bank’s short-term liabilities. The part of short-term funding which (in the relevant scenario) is assumed not to disappear, together with long-term debt and equity, forms the ASF. Because total liabilities equal total assets, the LCR and the NSFR requirements are equivalent: HQLA exceed expected net cash outflow if and only if ASF exceeds RSF. It follows that the two measures only vary in their time horizon.

*Why equivalent to wholesale funding restriction in our model*
References


