Repo Runs

Antoine Martin David Skeie
Ernst-Ludwig von Thadden[1]

October 2011

Abstract

This paper develops a dynamic model of financial institutions that borrow short-term and invest into long-term marketable assets. Because such intermediaries perform maturity transformation, they are subject to potential runs. We derive distinct liquidity and collateral constraints that characterize the fragility of such institutions as a result of changing market expectations. The liquidity constraint depends on the intermediary's endogenous liquidity position that acts as a buffer against runs. The collateral constraint depends crucially on the microstructure of particular funding markets that we examine in detail. In particular, our model provides insights into the fragility and differences of the tri-party repo market and the bilateral repo market that were at the heart of the recent financial crisis.

Keywords: Investment banking, securities dealers, repurchase agreements, tri-party repo, runs, financial fragility.
JEL classification: E44, E58, G24

[1]Martin and Skeie are at the Federal Reserve Bank of New York. Von Thadden is at the University of Mannheim. Author e-mails are antoine.martin@ny.frb.org, david.skeie@ny.frb.org, and vthadden@uni-mannheim.de, respectively. We thank Viral Acharya, Sudipto Bhattacharya, Patrick Bolton, Fabio Castiglionesi, Douglas Gale, Gary Gorton, Todd Keister, Ed Nosal, Lasse Pedersen, Matt Pritzker, Jean-Charles Rochet, Jean Tirole, and Jos van Bommel for helpful comments. Part of this research was done while Antoine Martin was visiting the University of Bern, the University of Lausanne, and the Banque de France. Von Thadden thanks the Finance Department of the London School of Economics for its hospitality during the final phase of this project. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

This paper develops an equilibrium model of financial institutions that are funded by short-term borrowing and hold marketable assets. We show that such institutions are subject to the threat of runs similar to those faced by commercial banks and study the conditions under which runs can occur. The analysis yields distinct liquidity and collateral constraints for such institutions that must both be violated for runs to occur.

The liquidity constraint obtains because equilibrium profits of financial institutions in our model are positive. Profits therefore act as a liquidity buffer and are a key stabilizing element against runs. The collateral constraint arises because investors' incentives to run on a particular firm depend on the value they expect their collateral to have. Both constraints depend on the firms' size, their short-term funding, and other structural variables. The collateral constraint depends crucially on the microstructure of the short-term funding market. We model the differences between various repo and other related funding markets and examine the consequences of these differences.

Our framework is general and can be applied to various types of financial institutions that suffered from losses in short term funding during the financial crisis of 2007-09. Such institutions include money market mutual funds (MMMFs), hedge funds, off-balance sheet vehicles including asset-backed commercial paper (ABCP) conduits, and structured investment vehicles (SIVs). The primary application of our model is to large securities dealers who use the tri-party repo market as a main source of financing. In that market dealers borrow from institutional investors, such as MMMFs, against collateral that is held by a third-party clearing bank. Dealers' borrowing in the tri-party repo market reached over \$2.8 trillion outstanding in aggregate at its peak in 2008; individual dealer borrowing reached \$400 billion, most of which with overnight maturity.

Our model is motivated by the observation that the collapses of Bear Sterns and Lehman Brothers were triggered by a precipitous decrease in funding from the tri-party repo market. As noted by Bernanke (2009), these sudden stops were surprising because tri-party repo borrowing is collateralized by securities. The Task Force on Tri-Party Repo Infrastructure (2009), a private sector body sponsored by the Federal Reserve Bank of New York, noted that "tri-party repo arrangements were at the center of the liquidity pressures faced by securities firms at the height of the financial crisis."[2] As a response, the creation of the primary dealer credit facility (PDCF) was an attempt to provide a backstop for the tri-party repo market.

Given the importance of repo markets for some key events in the crisis, we compare the organization of the tri-party repo market, which is the primary repo market for borrowing by dealers, with the bilateral repo market, which is the primary repo market for lending by dealers. Comparing tri-party repos and bilateral repos is particularly interesting because the two markets behaved very differently in the crisis. As documented by Gorton and Metrick (2011), haircuts in bilateral repos increased dramatically during the crisis, consistent with the margin spirals described in Brunnermeier and Pedersen (2009). In contrast, Copeland, Martin, and Walker (2010) show that haircuts in the tri-party repo market barely moved and document large differences in haircuts between the two markets for comparable asset classes. Our model clarifies the distinction between increasing margins, which is a potentially equilibrating phenomenon, and runs, which can happen if margins do not increase sufficiently to provide protection to investors. Furthermore, our analysis shows that a particular institutional feature of the tri-party repo market, the early settlement of repos by clearing banks called the "unwind", can have a destabilizing effect on the market. This finding lends theoretical support to the recent reform proposals by the Tri-Party Repo Infrastructure Reform Task Force to eliminate the unwind procedure. A general lesson of our analysis, therefore, is that the market microstructure of the shadow banking system plays a critical role for the system's fragility.

Our work builds on the theory of commercial bank instability developed by Diamond and Dybvig (1983), Qi (1994), and others. As pointed out by Gorton and Metrick (2011), there are important similarities between the fragility of commercial banks that borrow unsecured deposits and hold non-marketable loan portfolios, and of "securitized" or "shadow" banks, which borrow in repo or other short-term funding markets against marketable secu-

[2] See http://www.newyorkfed.org/tripartyrepo/.

rities as collateral. In particular, repo markets perform maturity transformation by allowing investors with uncertain liquidity needs to lend short-term against longer term, less liquid securities. We provide a formal model of shadow banking to identify the determinants of equilibrium profits, liquidity, and collateral that support such maturity transformation.[3]

Krishnamurthy, Nagel and Orlov (2011) document that prior to the crisis, the two main providers of funds to the shadow banking system, MMMFs and securities lenders, invested heavily in ABCP and the corresponding conduits. As shown by Covitz, Liang, and Suarez (2009), ABCP has very short maturities that shortened even further during the crisis. ABCP conduits therefore are an important case in point for our theory. And indeed, Covitz, Liang, and Suarez (2009) argue that the precipitous drop in outstanding ABCP of roughly $190 billion in August 2007 had many characteristics of a traditional run.

Our theory of fragility differs from the classic literature on commercial bank runs in several ways. First, we model collateral and the different ways it can be handled explicitly. Second, we do not model bank contracts as insurance arrangements for risk-averse investors and place no constraints on investor preferences. And third, perhaps most importantly, we distinguish between collateral and liquidity concerns by endogenizing banks' liquidity. In our model, dealers have the choice between funding securities with their own cash or with short-term debt. We derive a dynamic participation constraint under which dealers will prefer to fund their operations with short-term debt and show that this condition implies that dealers make positive profits in equilibrium. These profits can be used to forestall a run and thus serve as a systemic buffer. If current profits are insufficient to forestall a run, dealers can cut investment at the expense of future profits in order to generate further cash, and if even this is not sufficient, dealers can sell their assets to generate liquidity, potentially at a discount (Shleifer and Vishny, 1992). We derive this

[3] Shleifer and Vishny's "Unstable Banking" (2010) formalizes some elements of securitized banking, but focusses mostly on the spillover of irrational investor sentiments into the securitized loan market. Rampini and Viswanathan (2010) examine a dynamic model of intermediary effects of bank capital and collateralizable assets on lending but do not examine the fragility of intermediaries' liabilities.

discount in equilibrium and show when such asset sales relax the liquidity constraint of distressed dealers.

Our theory uses a simple dynamic rational expectations model with multiple equilibria. However, unlike in conventional models of multiple equilibria, not "everything goes" in our model. The theory pins down under what conditions individual institutions are subject to potential self-fulfilling runs, and when they are immune to such expectations. The intermediaries in our model are heterogenous and the liquidity and collateral constraints are specific to each institution. The equilibrium is therefore consistent with observations of some institutions failing and others surviving in case of changing market expectations. In particular, our theory is consistent with the observation by Krishnamurthy, Nagel, and Orlov (2011) that "the effects of the run on repo seem most important for a select few dealer banks who were heavy funders of private collateral in the repo market" (p.6).

While our theory focuses on multiple equilibria, the history of the 2007-09 crisis clearly also has a fundamental component. Our choice of model is motivated, on the one hand, by the wish to simplify the exposition and, on the other hand, by the belief that illiquidity was an important issue at some key turning points during the crisis.

The remainder of the paper proceeds as follows. Section 2 describes our model. Section 3 characterizes its steady states. In particular, we derive the dealers' dynamic participation constraint in this section and show that equilibrium profits are positive. Section 4 studies the dealers' ability to withstand runs in terms of liquidity. Section 5 considers the fragility of different market microstructures and derives collateral constraints. Section 6 generalizes the liquidity constraint derived in Section 4 to the possibility of asset sales. Section 7 discusses extensions of the model in the form of market runs and liquidity provision. Section 8 concludes.

2 The Model

2.1 Framework

We consider an economy that lasts forever and does not have an initial date. At each date t, a continuum of mass N of "young" investors is born who live for three dates. Investors are born with an endowment of 1 unit of goods that they can invest at date t and have no endowment thereafter. Investors' preferences for the timing of consumption are unknown when born at date t. At date $t+1$, investors learn their type. "Impatient" investors need cash at date $t+1$, while "patient" investors do not need cash until date $t+2$. The information about the investors' type and age is private, i.e. cannot be observed by the market. Ex ante, the probability of being impatient is α. We assume that the fraction of impatient agents in each generation is also α (the Law of Large Numbers).

The timing of the investors' needs of cash is uncertain because of "liquidity" shocks. In practice, money market investors, such as MMMFs, may learn about longer term investment opportunities and wish to redeploy their cash or they may need to generate cash to satisfy sudden outflows from their own investors. We do not model explicitly what investors do with their cash in the event of a liquidity shock and, for the remainder of the paper, simply assume that they value it sufficiently highly to want to use it at the given point in time.[4] Their utility from getting payments (r_1, r_2) over the two-period horizon can therefore simply be described by

$$U(r_1, r_2) = \begin{cases} u_1(r_1) & \text{with prob. } \alpha \\ u_2(r_2) & \text{with prob. } 1-\alpha \end{cases} \tag{1}$$

with u_1 and u_2 strictly increasing.[5]

[4] This assumption is as in Diamond and Dybvig (1983). As we shall show in the next section, together with a no-arbitrage assumption it implies that dealers are funded short-term. This argument is different from that of Diamond and Rajan (2001) who argue that short-term liabilities are a way to commit for bankers to repay the proceeds of their investments to depositors. For a critical assessment of short-term borrowing see Admati et al (2010).

[5] We do not assume the traditional consumption-smoothing motive of the Diamond-Dybvig literature (concave u_t), which would make little sense in our context.

Everybody in the economy has access to a one-period storage technology, which can be thought of as cash and returns 1 for each unit invested.

The economy is also populated by M infinitely-lived risk-neutral agents called dealers and indexed by $i \in \{1, ..., M\}$. Dealers have no endowments of their own but access to an investment technology, which we think of as investment in, and possibly the creation of, securities. These investments are illiquid in the sense that they cannot be liquidated instantaneously, and they are subject to decreasing returns, which we model simply by assuming that there is a limit beyond which the investment provides no returns. Hence, investing I^t units at date t yields

$$\begin{cases} R_i I^t & \text{if } I^t \leq \overline{I_i} \\ R_i \overline{I_i} & \text{if } I^t \geq \overline{I_i} \end{cases} \tag{2}$$

with $R_i > 1$ at date $t+2$ for all t and yields nothing at date $t+1$.[6] To simplify things, we assume that the return on these investments is riskless. In order to have a role for collateral in our model, we assume that the return is not verifiable. This means that investors cannot be sure that a dealer has indeed realized $R_i I^t$ from his past investment. Although this is a probability zero event, a dealer who has received funds from investors can claim that he cannot repay the investors.

Investment returns can only be realized by the dealer who has invested in the asset, because dealers have a comparative advantage in managing their security portfolio. Other market participants only realize a smaller return. Investors could realize a return of $\gamma_i^t R_i$ from these assets, with $\gamma_i^t < 1$ for all t, and other dealers could realize $\hat{\gamma}_i^t \in [\gamma_i^t, 1]$. γ_i^t and $\hat{\gamma}_i^t$ reflect different skills in valuing or managing the assets, possible restrictions on the outsider's portfolio composition, transactions and timing costs, and similar asymmetries.[7] We allow γ_i^t to depend on the dealer, reflecting potential

[6] The need to assume such capacity constraints (or more generally, decreasing returns) in dynamic models of liquidity provision has been pointed out by van Bommel (2006).

[7] For T-bills, γ_i^t should be very close to 1 at all times. But dealers typically also finance large volumes of less liquid securities. Simplifying somewhat, the main categories of collateral in repo markets are (i) US treasuries and strips, (ii) Agency debentures, (iii) Agency ABS/MBS, (iv) Non-Agency ABS/MBS, (v) corporate bonds. We could have different γ_i^t for each class of collateral without changing the analysis.

differences in the portfolio of collateral that different dealers seek to finance, and allow it to depend on t since the outside value of such portfolios may change over time.

Dealers use the endowment of young investors to invest in securities. To make the model interesting, we must assume that the total investment capacity $\overline{I} = \sum \overline{I_i}$ strictly exceeds the investors' amount of cash available for investment, N.[8] Without this assumption, there would be no competition among dealers for short-term cash from investors. Dealers could extract all the surplus from investors by simply offering to pay the storage return of 1 each period, and there would be no instabilities or runs. Instead of the condition $\overline{I} > N$, we assume the slightly stronger condition

$$\sum_{j \neq i} \overline{I_j} > N \tag{3}$$

for all i. Hence no dealer is pivotal, and even if one dealer fails, there will still be competition for investor funds.

If dealer i in period t invests I_i^t, holds c_i^t in cash, receives b_i^t from young investors, repays r_{1i}^t after one period or r_{2i}^t after two periods, impatient investors do not roll over their funding when middle-aged, but patient investors do, then the dealer's expected profits are

$$\pi_i^\tau = R_i I_i^{\tau-2} + c_i^{\tau-1} + b_i^\tau - \alpha r_{1i}^{\tau-1} b_i^{\tau-1} - (1-\alpha) r_{2i}^{\tau-2} b_i^{\tau-2} - I_i^\tau - c_i^\tau. \tag{4}$$

At each date, dealers consume their profits. The dealer's objective at each date t then is to maximize the sum of discounted expected profits $\sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi_i^\tau$, where $\beta < 1$. In order to make the problem interesting, we assume that dealers are sufficiently patient and their long-term investment is sufficiently profitable:

$$\beta^2 R_i > 1 \tag{5}$$

for all i. Given the investors' preferences in (1), there is no scope for rescheduling the financing from investors. Hence, if $\pi_i^t < 0$ at any date t, the dealer is bankrupt, unless he is able to sell assets to other dealers, which we consider in Section 6.

[8] As usual, all quantities are expressed per unit mass of investors.

3 Steady-states

As a benchmark, this section characterizes steady-state allocations in which in each period young investors fund dealers and withdraw their funds precisely at the time of their liquidity shocks. We shall see that these are the only possible steady states. We assume that the Law of Large Numbers also holds at the dealer level: each period the realized fraction of impatient investors at each dealer is α. Hence, in every period, each dealer obtains funds from young investors, and repays a fraction α of middle-aged investors and all remaining old investors. Thus there is no uncertainty about dealers' profits, and each dealer's realized profit is equal to his expected profit (4).

Each period, dealers compete for investors' funds. Since dealers have a fixed investment capacity, they cannot make unconditional interest rate offers, but must condition their offers on the amount of funds they receive. The simplest market interaction with this feature is as follows.[9] At each date $t \in (-\infty, \infty)$:

1. Dealers offer contracts $(r^t_{1i}, r^t_{2i}, Q^t_i, k^t_i) \in \mathbb{R}^4_+$, $i = 1, ..., M$.
2. New and patient middle-aged investors decide whether to finance the dealer.
3. If the dealer is unable to repay all investors who demand repayment, he must declare bankruptcy. Otherwise, the dealer invests I^t_i and continues.

Here, $r^t_{\tau i}$ is the (gross) interest payment offered by dealer i on t-period funding, Q^t_i the maximum amount for which this offer is valid, and k^t_i is the amount of collateral posted per unit borrowed. Total new borrowing by the M dealers then is $(b^t_1, ..., b^t_M) \in \mathbb{R}^M_+$, with $b^t_i \leq Q^t_i$ for $i = 1, ..., M$ and $\sum b^t_i \leq N$. Since investment returns are non-verifiable, the collateral posted must be sufficient to incentivize dealers to repay, i.e. to honor the repurchase leg of the repo transaction. At the time of the contract offer to middle-aged

[9] Our analysis in this section would be unchanged if we assumed a competitive lending market, with competitive interest rates r_1 and r_2. Explicit interest rate competition only becomes relevant in the later analysis of runs.

investors, the dealer needs r_{1i}^{t-1} in cash and offers collateral maturing one period later; hence, at that time the dealer will be expected to repay instead of keeping his cash if

$$R_i k_i^t r_{1i}^{t-1} \geq r_{2i}^{t-1} \tag{6}$$

In order to obtain cash from young investors, the dealer offers to put up the assets he creates with these funds as collateral. One period later, he will want to repay instead of giving up the assets if

$$\beta R_i k_i^t \geq r_{1i}^t \tag{7}$$

We will abstract from more complicated considerations of default and ex post bargaining, and simply assume that collateral must satisfy the two repayment constraints (6) and (7).[10] A steady state equilibrium is a collection of $(r_{1i}, r_{2i}, k_i, b_i, I_i, c_i)$ for each dealer i, where b_i is new funding, k_i collateral, c_i cash holding, and $I_i \leq \overline{I_i}$ investment per dealer, such that no dealer and investor would prefer another funding and investment policy, given the behavior of all others.[11]

We now characterize the steady states in which dealers invest by a sequence of simple observations.

Lemma 1 *For each dealer i with $b_i > 0$, $r_{2i} = r_{1i}^2$.*

Proof. Clearly, $r_{2i} \geq r_{1i}^2$, because otherwise investors would strictly prefer to never roll over their funding, regardless of their type. Patient middle-aged investors would withdraw their funds and then invest again together with young investors. Suppose that this inequality is strict. In this case, an

[10] See, e.g., Hart and Moore (1998) or von Thadden, Berglöf and Roland (2010) for more complex models of default and renegotiation. We also abstract from reputational or other dynamic concerns, which would trade off the possible loss of future access to investor funds against current cash gains. Note that (6) and (7) are consistent with observed practice in the repo market in the sense that in the (rare) cases in which repos are not repaid investors usually choose to extend them for another night.

[11] For simplicity, we can ignore the bound Q_i in the description of the steady state, where it can be thought of as being set to $Q_i = b_i$. The bound plays no substantive role in steady state, but is important for runs in later sections.

impatient middle-aged investor can deviate by extending her funding and at the same time borrowing the amount q from a young investor at interest rate $\widehat{r}-1$. He can then claim back r_{2i} from the dealer one period later and repay $q\widehat{r}$. This deviation is feasible if $q\widehat{r} = r_{2i}$. It makes both parties weakly better off iff (i) $q \geq r_{1i}$, (ii) $\widehat{r} \geq r_{1i}$, and (iii) $\widehat{r}r_{1i} \geq r_{2i}$.

(i) and (iii) imply $\widehat{r}r_{1i} = r_{2i}$. Hence, (ii) holds with strict inequality, and there are gains from trade. ■

The proof is based on a simple no-arbitrage argument. It is different from the classical argument by Jacklin (1987) in the context of the Diamond-Dybvig (1983) model, because investors in our context do not have access to the long-term investment technology. It is also different from the argument by Qi (1994), who assumes and uses strict concavity of the investors' utility. In our market context, the no-arbitrage argument is natural and sufficient.[12] Note that although Lemma 1 forces the yield curve to be flat, dealers still provide maturity transformation if $r_{1i} > 1$.

Lemma 2 $r_{1i} = r_{1j}$ *for all dealers* i, j *with* $b_i, b_j > 0$.

Proof. Suppose that $r_{1i} < r_{1j}$ for some i, j with $b_i, b_j > 0$. Let $\mathcal{J}_i$ be the set of all dealers k with $r_{1k} > r_{1i}$ and $b_k > 0$. $\mathcal{J}_i$ is not empty because $j \in \mathcal{J}_i$. All $k \in \mathcal{J}_i$ must be saturated, i.e. have $b_k = Q_k$ (otherwise investors from i would deviate). Hence, any dealer $k \in \mathcal{J}_i$ can deviate to $r_{1k} - \varepsilon$ for $0 < \varepsilon < r_{1k} - r_{1i}$ and strictly increase his profit. ■

By Lemma 2 the Law of One Price holds, and we can denote the single one-period interest rate quoted by all active dealers by $r = r_1$. Then the steady-state budget identity of dealer i is

$$R_i I_i + b_i = I_i + \alpha r b_i + (1-\alpha) r^2 b_i + \pi_i \tag{8}$$

where the left-hand side are the total inflows per period and the right-hand side total outflows.

[12] "Early dyers" (as the Diamond-Dybvig literature calls them) do not die, and are perfectly able to transact after their liquidity shock.

Clearly, if $R_i > 1$, the higher is I_i the higher are profits.[13] We do not concern ourselves with showing how a steady state with $I_i > 0$ would emerge if there were a startup period. But under our assumption (5) that dealers are sufficiently patient, it is optimal for dealers to build investment up to maximum capacity.

Lemma 3 *In steady-state dealers do not hold cash: $c_i = 0$ for all i.*

Proof. Since $\beta < 1$, and $c_i > 0$ does not affect the dealer's budget constraint (8), each dealer does strictly better by consuming c_i. ■

Lemma 4 *If $r > 1$, total steady-state funding by investors is maximal: $\sum_{i=1}^{M} b_i = N$.*

Proof. The total supply of funds is inelastically equal to N in each period if $r > 1$. The scarcity constraint (3) implies that there is a dealer who invests less than full capacity, $I_i < \overline{I}_i$. Suppose that $\sum_{i=1}^{M} b_i < N$. If i makes strictly positive profits, he strictly increases his profits by setting $Q_i = \overline{I}_i$ and thus attracting more funds. If i makes zero profits, he can make strictly positive profits by reducing his interest rate marginally, setting $Q_i = \overline{I}_i$, and attracting the previously idle supply of funds. ■

Lemma 5 *If $\pi_i > 0$, steady-state investment of dealer i is maximal: $I_i = \overline{I}_i$.*

Proof. Suppose the lemma is wrong. The dealer can then increase investment slightly at any date t by using his own cash. By condition (5), this yields a strict increase in discounted profits. ■

[13] The literature on dynamic banking has not always been clear about the distinction between investment capacity ($\overline{I}_i$ in our model) and per capita borrowing (N/M). In particular, the implicit assumption that $\overline{I}_i = N/M$ in Qi (1994), Bhattacharya and Padilla (1996) and Fulghieri and Rovelli (1998) is not necessary, and may even ignore interesting dynamic features. See van Bommel (2006) for an excellent discussion.

Lemma 6 *If there exists a dealer i with $\pi_i > 0$ and $b_i > 0$ then steady-state interest rate satisfies*

$$(1-\alpha)\beta^2 r^2 + \alpha\beta r = 1 \tag{9}$$

Proof. For each unit of cash that dealer i receives and invests at date t, he pays back αr in $t+1$, generates returns R_i in $t+2$, and pays back $(1-\alpha)r^2$ in $t+2$. Hence, his expected discounted profits on this one unit are $\beta^2(R_i - (1-\alpha)r^2) - \beta\alpha r$. Alternatively he could invest his own cash. The discounted profits from not using the one unit of outside funds and rather investing his own money is $\beta^2 R_i - 1$. If the dealer receives funds from investors in steady state ($b_i > 0$) and has funds of his own ($\pi_i > 0$), this cannot be strictly better, which implies $(1-\alpha)\beta^2 r^2 + \alpha\beta r \leq 1$.

Suppose that this inequality is strict. For an arbitrary dealer j, this means that

$$\beta^2(R_j - (1-\alpha)r^2) - \beta\alpha r > \beta^2 R_j - 1 \tag{10}$$

which is strictly positive by (5). Hence, all dealers strictly prefer $b_i = \overline{I_i}$. This contradicts (3), because the demand for funds would exceed supply. ■

Lemma 6 is surprisingly strong: the existence of one active dealer with strictly positive profits pins down the equilibrium interest rate. We call condition (9) the dealers' "dynamic participation constraint." Basic algebra shows that its solution is

$$\overline{r} \equiv 1/\beta > 1.$$

This makes sense: at the margin, dealers discount profits with the market interest rate. But it is interesting to note that $\overline{r}$ does not depend on other supply and demand characteristics such as R_i and α. In steady-state, the cost of funds, $\overline{r}-1$, is determined exclusively by the dealers' discount factor. This makes them indifferent at the margin between attracting more cash from investors, which increases current dealer profits, or attracting less and using their own cash to finance investments, which increases future dealer profits. Overall therefore, since the marginal profit from using own funds is strictly positives by (5), (9) implies that dealers make positive profits in equilibrium, if one active dealer makes positive profits in equilibrium. The following proposition provides a simple condition for this to be the case.

Proposition 1 *If*

$$\frac{(1+\beta)\beta^3}{1-\alpha+\beta}\sum_i R_i\overline{I_i} \geq N \tag{11}$$

then steady states exist in which

- *investors roll over their loans according to their liquidity needs,*
- *all dealers make strictly positive profits,*
- $I_i = \overline{I_i}$, $c_i = 0$, *and* $r = \overline{r}$,
- *borrowing satisfies* $\sum_i b_i = N$,

$$b_i \leq \frac{(1+\beta)\beta^3 R_i}{1-\alpha+\beta}\overline{I_i}, \tag{12}$$

and is otherwise indeterminate,

- *collateral* k_i *satisfies*

$$\frac{1}{\beta^2 R_i} \leq k_i \leq \frac{(1+\beta)\beta\overline{I_i}}{(1-\alpha+\beta)b_i} \tag{13}$$

and is otherwise indeterminate.

Proof. Suppose there is an equilibrium in which at least one active dealer makes $\pi_i > 0$. Lemma 6 implies that $r = \overline{r}$.

Simple algebra shows that (12) and (5) imply

$$\begin{aligned} b_i &< \frac{\beta^2(R_i-1)}{1-\alpha+\alpha\beta-\beta^2}\overline{I_i} \\ &= \frac{(R_i-1)\overline{I_i}}{(1-\alpha)\overline{r}^2+\alpha\overline{r}-1} \end{aligned}$$

This implies that dealer equilibrium profits as defined in (8),

$$\pi_i = (R_i-1)\overline{I_i} - ((1-\alpha)\overline{r}^2+\alpha\overline{r}-1)b_i, \tag{14}$$

are strictly positive for all i. Hence, $I_i = \overline{I_i}$ for all i by Lemma 5. At the interest rate $\overline{r}$, every dealer i is indifferent at any date t between using outside funds and using his own cash π_i for investment, and thus finds it indeed optimal to borrow any positive amount b_i satisfying (12). Because $\overline{r} > 1$ and all dealers pay the same interest rate, patient middle-aged investors find it indeed optimal to roll over their funding and young investors find it optimal to invest all their endowment. By (11), there exist borrowing levels $(b_1, ..., b_M)$ satisfying (12) such that $\sum_i b_i = N$. This establishes the existence of equilibria with the first four features of Proposition 1.

The repayment condition (7) is equivalent to the first inequality in (13) and implies (6). For the second inequality in (13), note that in steady state the dealer has two types of securities to offer as collateral, those maturing at $t+1$ or maturing at $t+2$. Because $r = 1/\beta$, both dealers and investors value both types of securities identically. Hence, the maximum amount of collateral a dealer can pledge in steady state is $\bar{I}_i(1+\beta)$, in terms of securities maturing at $t+1$. The total amount of funds provided by investors per period is $b_i\,[1+(1-\alpha)\overline{r}] = b_i\,[1-\alpha+\beta]\,/\beta$. It follows that the maximum amount of collateral per unit that the dealer can offer is

$$\kappa_i \equiv \frac{\beta \bar{I}_i(1+\beta)}{b_i\,[1-\alpha+\beta]}. \tag{15}$$

The second inequality in (13) is the condition $k_i \leq \kappa_i$. Both inequalities in (13) are compatible by condition (12). ■

Condition (11) is similar in spirit to the scarcity condition (3), but may be slightly stronger. It can easily verified, using (5), that (3) implies (11) if $\beta^2 \geq 1-\alpha$.

The proof of Proposition 1 actually implies that all steady state equilibria are either of the form described in the proposition or have zero profits for all active dealers. Such zero-profit equilibria can exist in some parameter configurations, but "usually" don't. In the appendix we provide a full characterization.[14] Even if they do exist, such equilibria are implausible (given

[14]For example, such equilibria do not exist if dealers are sufficiently large (more precisely, if $\overline{I}_i \geq N/2$ for all $i = 1, ..., M$) or if dealers are sufficiently heterogeneous in their profitability R_i.

the large profits in practice) and knife-edge cases: they would disappear if dealers had a small exogenous endowment in at least one period. For all these reasons, we ignore these equilibria in the sequel and use the steady states identified in Proposition 1 as a benchmark for the rest of the analysis.

An important and novel feature of these equilibria is that condition (9) prevents competition from driving up interest rates to levels at which dealers make zero profits. The reason why dealer profits are positive is intuitive (but not trivial): dealers must have an incentive to use their investment opportunities on behalf of investors instead of using internal funds to reap those profits for themselves. This rationale of positive intermediation profits is different from the traditional banking argument of positive franchise values (e.g., Bhattacharya, Boot, and Thakor (1998), or Hellmann, Murdock and Stiglitz, (2000)), as it explicitly recognizes the difference between internal and external funds. Hence, the coexistence of internal and external funds and the internalization of all cash flows arising from them implies that financial intermediaries make positive profits.[15]

The steady states of Proposition 1 all feature maximum investment and the same interest rate $\overline{r}$, but dealers can differ in their reliance on borrowing b_i and the collateral k_i they post. In fact, in steady state the exact amount of collateral, subject to constraint (13), plays no role because investors never consume it. It is important nevertheless, because it makes sure that each period the cash changes hands as specified.

In steady state, the borrowing level b_i is only limited by the requirement that the dealer has sufficient collateral as defined by (13). The resulting restriction (12) by itself implies that the dealer makes strictly positive profits. It is important to realize that in steady state dealers have no incentive to change their exposure, but that they may prefer other steady states. Hence, Proposition 1 is consistent with the notion that dealers can be "trapped" in an equilibrium with high borrowing and low profits. In fact, as seen in (14), dealer profits are strictly decreasing in b_i. Therefore, to the extent that period profits act as a buffer against adverse shocks, as we show in the following sections, dealers with larger borrowing levels will be more fragile.

[15] This is different from Acharya, Myers, and Rajan (2010) where overlapping generations of bankers try to pass on the externality of debt.

4 Runs without asset sales

In this section, we study the stability of dealers in the face of possible runs. We analyze this problem under the assumption that behavior until date t is as in Proposition 1 and ask whether a given dealer can withstand the collective refusal of all middle-aged investors to extend their funding and of young investors to provide fresh funds.[16] In the next section we will describe the specific microstructure of the tri-party repo market and other institutions that can make such collective behavior of investors optimal and thus imply that the corresponding individual expectations are self-fulfilling.

The key question is how much cash the dealer can mobilize to meet the repayment demands by middle-aged investors in such a situation. In steady state, at the beginning of the period, a dealer, on the asset side of his balance sheet, holds $R_i\overline{I_i}$ units of cash from investments at date $t-2$, as well as securities that will yield $R_i\overline{I_i}$ units of cash at date $t+1$. The dealer holds investor claims for dates t and $t+1$ on the liability side of his balance sheet. In this section, we assume that the dealer cannot sell his assets.

The dealer's repayment obligations in case of a run are $(\overline{r}+(1-\alpha)\overline{r}^2)b_i$. If there is no fresh funding in the run and new investment is maintained at the steady-state level $\overline{I_i}$, the run demand can be satisfied by the individual dealer if

$$(R_i-1)\overline{I_i} \geq (\overline{r}+(1-\alpha)\overline{r}^2)b_i. \tag{16}$$

If (16) holds, a run would have no consequence whatsoever and all out-of-equilibrium investor demand would be buffered by the dealer's profits. Anticipating this, investors have no reason to run. But more is possible. In the event of a run at date t, the cash position of the individual dealer who satisfies the run demand is

$$I_0 = R_i\overline{I_i} - (\overline{r}+(1-\alpha)\overline{r}^2)b_i. \tag{17}$$

Clearly, if $I_0 < 0$ the dealer does not have the liquidity to stave off the run and is bankrupt. If $I_0 \geq 0$, but (16) does not hold, the dealer must adjust

[16] Note that in our infinite-horizon model, there are two sources of instability: middle-aged investors may not roll over their funding and new investors may not provide fresh funds. The former corresponds to the classical Diamond-Dybvig problem, the latter arises only in fully dynamic models.

his funding or investment in order to survive the run. Since after a run in $t+1$ the dealer will have $R_i\overline{I_i}$ in cash and nothing to repay, he can resume his operations by investing $\overline{I_i}$ at date $t+1$ and save and invest thereafter. Whether he can attract fresh funds after t depends on the market, but this is immaterial for his survival.

The liquidity constraint, (18) in the following proposition, is obtained by simply writing out the condition $I_0 \geq 0$ from (17).

Proposition 2 *In steady state, a run on dealer i who cannot sell her assets can be accommodated if and only if the dealer's liquidity constraint holds, i.e. if*

$$\beta^2 R_i\overline{I_i} \geq (1-\alpha+\beta)b_i. \tag{18}$$

Condition (18) is independent of the borrowing restriction (12) of Proposition 1, in the sense that (18) can hold or fail in steady state, depending on the parameters. Hence, a dealer who makes positive profits in steady state may still fail in a run. The comparative statics of the liquidity constraint are simple and we collect them in the following proposition.

Proposition 3 *The liquidity constraint (18) is the tighter,*

- *the higher is the dealer's leverage $b_i/\overline{I}_i$,*
- *the lower is the dealer's investment capacity $\overline{I}_i$,*
- *the lower is the dealer's profitability R_i.*

Proposition 3 shows that if dealers have sufficient access to profitable investment ($\overline{I_i}$ large), if these investments are sufficiently profitable (R_i large), or if they have sufficiently low exposure to short-term borrowing ($b_i/\overline{I}_i$ small), then dealers are more likely to be able to stave off runs individually, only by reducing their investment temporarily. In this case, unexpected runs cannot bring down dealers out of equilibrium. If condition (18) is violated, a run would bankrupt the individual dealer if he cannot sell his illiquid assets.

5 Fragility

This section examines different microstructures that are associated with repo markets or other money markets. We ask whether runs can occur in each of the institutional environments considered. The focus is on the tri-party repo market, but we also examine bilateral repos, MMMFs, ABS-backed conduits, and traditional bank deposits. We derive a collateral constraint for each market and show that if and only if the liquidity constraint and the collateral constraint are violated, then a run can occur for the particular market structure.

We study unanticipated runs that arise from pure coordination failures. As noted in the previous section, in a run at date t all investors believe that i) no middle-aged investors renew their funding to dealer i, so the dealer must pay $\left[\overline{r} + (1-\alpha)\overline{r}^2\right] b_i$ to middle-aged and old investors, and ii) no new young investors lend to the dealer. The question is whether such beliefs can be self-fulfilling in a collective deviation from the steady state.

Since the Law of One Price holds in steady state by Lemma 2, a trivial coordination failure may induce all investors of a given dealer to switch to another dealer out of indifference. This looks like a "run", but is completely arbitrary. We will therefore assume that investors if indifferent lend to the dealer they are financing in steady state. Hence, in order for a collective deviation from the steady state to occur we impose the stronger requirement that the individual incentives to do so must be strict.

The first insight, which applies to all institutional environments considered in this section, is simple but useful to state explicitly: a run cannot occur if a dealer is liquid in the sense of Proposition 2.

Lemma 7 *If a dealer satisfies the liquidity constraint (18), there are no strict incentives to run on this dealer.*

The proof is simple. In a run on this dealer, all middle-aged patient investors would be repaid in full regardless of what young investors do and without affecting the dealer's asset position. Hence, patient middle-aged and young investors are indifferent between lending to the dealer or to another one. By our assumption about the resolution of indifference, there is thus no reason to run in the first place. Intuitively, patient middle-aged investors would just "check on their money" before it is re-invested. Since the dealer has the money, such a check does not cause any real disruption, and the dealer may as well keep it until he invests in new securities.

5.1 The US tri-party repo market

This section briefly reviews the microstructure of the tri-party repo market and the key role played by the clearing bank.[17] In particular, we show that a practice called the "unwind" of repos increases fragility in this market.

The clearing banks play many roles in the tri-party repo market. They take custody of collateral, so that a cash investor can have access to the collateral in case of a dealer default, they value the securities that serve as collateral, they make sure the specified margin is applied, they settle the repos on their books, and importantly, they provide intraday credit to dealers.[18]

In the US tri-party repo market, new repos are organized each morning, between 8 and 10 AM. These repos are then settled in the afternoon, around 5 PM, on the books of the clearing banks. For operational simplicity, because dealers need access to their securities during the day to conduct their business, and because some cash investors want their funds early in the day, the clearing banks "unwind" all repos in the morning. Specifically, the clearing banks send the cash from the dealers' to the investors' account and the securities from the investors' to the dealers' account. They also finance the dealers' securities during the day, extending large amounts of intraday credit. At the time when repos are settled in the evening, the cash from the overnight investors extinguishes the clearing bank's intraday loan.

From the perspective of our theory, we can model the clearing bank as an agent endowed with a large amount of cash. By assumption, the clearing bank can finance the dealer only intraday. At each date, the clearing bank finances

[17] More details about the microstructure of the tri-party repo market can be found in Task Force (2010) and Copeland, Martin, and Walker (2010). The description of the market corresponds to the practice before the implementation of the 2010 reforms.

[18] The reform proposed by the Task Force would limit considerably the ability of the clearing banks to extend intraday credit (Task Force 2010).

dealers according to the following intra-period timing, which complements the timing considered in the previous section:

1. The clearing bank "unwinds" the previous evening's repos. For a specific dealer i this works as follows:
   (a) The clearing banks sends the cash amount $b_i\left[\overline{r}+(1-\alpha)\overline{r}^2\right]$ to all investors of dealer i, extinguishing the investors' exposure to the dealer they have invested in.
   (b) At the same time, the clearing bank takes possession of the assets the dealer has pledged as collateral.
   (c) In the process, the clearing bank finances the dealers temporarily, holding the assets as collateral for its loan.
2. $\bar{I}_i$ assets of a dealer mature (yielding $R_i\overline{I_i}$ in cash), allowing the dealer to repay some of its debt to the clearing bank.
3. Possibly a sunspot occurs.
4. The dealer offers a new repo contract $(\widehat{r}_i, \widehat{Q}_i, \widehat{k}_i)$.
5. New and patient middle-aged investors decide whether to engage in new repos with the dealer.
6. If the dealer is unable to repay its debt to the clearing bank, he must declare bankruptcy. Otherwise, the dealer continues.

This time line explicitly takes into account the sunspot that may cause a change of investor expectations. This is a zero-probability event that allows investors to coordinate on a run, if such out-of-equilibrium behavior is optimal for them.[19] For simplicity, we assume that the clearing bank extends the intraday loan to the dealer at a zero net interest rate. Also, since runs

[19] The sunspot also allows the dealer to react to the run. This adds realism to the model and makes runs more difficult (because the dealer's contract offer in stage 4 can now be different from the steady-state offer $(\overline{r}, Q_i, b_i, k_i)$).

are zero probability events the clearing bank has no reason not to unwind repos.[20]

In the tri-party repo market, traders choose only the interest rate applicable to the repo. The haircut for each collateral class is included in the custodial undertaking agreement between the investor, the dealer, and the clearing bank, and is not negotiated trade by trade. It is possible to change haircuts by amending the custodial agreement but this takes time. In practice, these changes appear to occur only rarely. We therefore assume that the contract offered in response to a sunspot must leave collateral unchanged from its steady state value, $\widehat{k}_i = k_i$, from Proposition 1.[21]

In response to the contract offer by the dealer, individual investors must compare their payoff from investing with the dealer in question to that from investing with another dealer. The latter decision yields the common market return $\overline{r}$,[22] the return from the former depends on what the other investors do. Table 1 shows the payoffs of the two decisions for the individual investor (rows) as a function of what the other investors do (columns), if the dealer is potentially illiquid (i.e. if the liquidity constraint (18) is violated). If the investor re-invests her funds with the dealer, the clearing bank will accept the cash, since it reduces its intraday exposure to the dealer, and give the investor assets that mature at date $t+1$. These are the only assets available in case of a run since the clearing bank will not let the dealer invest in new securities unless it obtains enough funding. Hence, in case of a run, an investor who agrees to provide financing receives securities that yield $\gamma_i^{t+1} R_i k_i$ at date $t+1$ if the dealer defaults.

[20] In the appendix, we consider the coordination problem between the clearing bank and the investors.

[21] Copeland, Martin, and Walker (2010) provide more details about haircuts in the tri-party repo market. In particular, they document that haircuts hardly moved, even at the peak of the crisis.

[22] This is obvious if the investor is the only one to deviate, because then he is negligible. If all investors of the dealer in question deviate, this follows from the slack in assumption (3).

		other investors	
		invest	don't
	invest	$\widehat{r}_i$	$\gamma_i^{t+1} R_i k_i$
	don't	$\overline{r}$	$\overline{r}$

Table 1: Payoffs in tri-party repo with unwind

Hence, investors will finance the dealer in case of a run iff[23]

$$\overline{r} \leq \gamma_i^{t+1} R_i k_i \tag{19}$$

Note that the investors' decision-making is completely dichotomous. If they anticipate a run, only collateral matters; if they anticipate no run, only interest matters. If condition (19) does not hold, the collective decision not to lend to the dealer in question is self-enforcing. In this case, the yield from the securities pledged as collateral is so low that an investor who believes that nobody will invest with dealer i would also choose not to invest. In our model, steady state collateral is not unique, but clearly, if constraint (19) is violated for the maximum possible amount of collateral κ_i in (15), then it cannot hold in any case.

Combining the above results with those of the previous section and writing out condition (19) for $k_i = \kappa_i$, the maximum amount of collateral per unit borrowed, yields the following prediction about the stability of the tri-party repo market.

Proposition 4 *In the tri-party repo market, a run on a dealer i can occur and bankrupt the dealer if and only if the dealer's liquidity constraint (18) and his collateral constraint*

$$\beta^2 R_i \overline{I}_i \geq \frac{1 - \alpha + \beta}{\gamma_i^t (1 + \beta)} b_i \tag{20}$$

are both violated for some time t.

[23]The weak inequality is due to the assumption that investors do not switch dealers if indifferent. If $\overline{r} = \gamma_i^{t+1} R_i \kappa_i$, there exists the trivial run equilibrium discussed at the beginning of this section.

Condition (20) is implied by the steady-state borrowing constraint (12) of Proposition 1 if γ_i^t is close to 1 and stronger than that constraint if γ_i^t is small. Hence, if investors can use the collateral almost as efficiently as dealers ("good" collateral in "normal" times), the collateral constraint is slack, and the dealer is run-proof. The collateral constraint becomes relevant only when there are larger differences in valuation between investors and dealers. Furthermore, condition (20) is independent of the liquidity constraint (18). The comparative statics of the collateral constraint for the tri-party model are again simple and we collect them in the following proposition.

Proposition 5 *The collateral constraint (20) is the tighter,*

- *the lower is the value γ_i^t of collateral to investors*
- *the higher is the dealer's short-term leverage $b_i/\overline{I}_i$,*
- *the lower is the dealer's investment capacity $\overline{I_i}$,*
- *the lower is the dealer's productivity R_i.*

Hence, the comparative statics with respect to b_i, $\overline{I_i}$, and R_i are identical for the two constraints (18) and (20). Both constraints are relaxed if dealers have sufficient access to profitable investment ($\overline{I_i}$ large), if these investments are sufficiently profitable (R_i large), or if they have sufficiently low leverage ($b_i/\overline{I}_i$ small). In this case, there is no reason for unexpected runs to occur on the investor side, and they cannot bring down dealers if they occur out of equilibrium. In the opposite case, a run can be a self-fulfilling prophecy and bankrupt the dealer.

5.2 Tri-party repo without unwind

To highlight the importance of the unwind mechanism for the fragility of the tri-party repo market, it is interesting to consider what would happen to the game described in the previous section if there were no unwind.[24] This case is similar to the tri-party repo markets in Europe. It is also similar to what the US tri-party repo market should become once the recommendation of the Task Force will be implemented.[25]

When there is no unwind, the timing of events intraday is as follows:

1. Possibly a sunspot occurs.
2. The dealer offers a new repo contract $(\widehat{r}_i, \widehat{Q}_i, k_i)$.
3. New and patient middle-aged investors decide whether to engage in new repos with a dealer.
4. If the dealer is unable to repay his debt to last period's repo investors, he must declare bankruptcy. Otherwise, the dealer continues.

From Lemma 7 it is again enough to consider the case in which the dealer is illiquid after a run. The situation without the unwind differs in two important respects from the one with unwind. First, without the unwind, an individual investor is repaid $\overline{r}$ if and only if the dealer can repay everybody - otherwise the dealer is bankrupt and repays everybody less than the contractual payment. Second, in contrast to the case with unwind, young and middle-aged investors are in a different situation when there is no unwind. Young investors hold cash while middle-aged investors hold a repo with the dealer, until the dealer is able to repay.

In case of a run, an illiquid dealer is bankrupt. All middle-aged investors then keep their collateral and may obtain additional cash as unsecured creditors depending on the bankruptcy rules. This payment is independent of whether an individual investor has demanded to be repaid or has agreed to roll over his loan. Hence, middle-aged investors are indifferent whether to

[24]In this paper, we do not model why the unwind may be beneficial. As described in Task Force (2010) and Copeland, Martin, and Walker (2010), the unwind makes it easier for dealers to trade their securities during the day. Automatic substitution of collateral, as is currently available in the European tri-party repo market and is being introduced in the US, allows dealers to have access to their securities even as investors remain collateralized.

[25]More information about the proposed change to settlement in the tri-party repo market can be found at http://www.newyorkfed.org/tripartyrepo/task_force_proposal.html.

run or not. Given the tie-braking rule assumed throughout this section, patient middle-aged investors therefore reinvest. This in turn induces young investors to invest with the dealer:

Lemma 8 *If middle-aged patient investors reinvest, investing is a (weakly) dominant strategy for new investors.*

Proof. If middle-aged patient investors do not withdraw their funds, the dealer is liquid, because

$$R_i\overline{I_i} - \left(\frac{\alpha}{\beta} + \frac{1-\alpha}{\beta^2}\right) b_i > 0$$

by (14) and (12). The dealer therefore has enough assets that will mature in the future to satisfy all future claims by young agents who invest today. ■

Hence, when there is no unwind, the incentives of investors are modified so that they never have a strict incentive to run. In essence, this is because the overnight repo market is an institution that creates simultaneity: if a sufficiently large number of investors do not re-invest, there is bankruptcy and all current creditors (the middle-aged investors) are treated equally, regardless of their intention to withdraw funding. This eliminates fragility due to pure coordination failures.

Proposition 6 *In the tri-party repo market without unwind, there are no strict incentives to run on dealers.*

5.3 Bilateral repos

In this section, we apply our model to bilateral repos. As noted in the introduction, MMMFs and other less sophisticated investors typically invest in tri-party repos rather than bilateral repos. Instead, the providers of funds in the bilateral repo market are mostly dealer banks and the borrowers are hedge funds and other dealers. In this section, we will therefore use the more general terms "borrowers" and "lenders".

Typically, bilateral repos have a longer term than tri-party repos. Hence, one period in our model should be thought of as representing a few days to a few weeks.[26] In terms of our assumptions this means that borrowers can adjust the whole contract offer in response to a sunspot.

To simplify the exposition of institutional details, we consider a borrower who funds "Fed-eligible" securities; securities that can be settled using the Fedwire Securities Service®. Fedwire Securities is a delivery versus payment settlement mechanism, meaning that the transfer of the securities and the funds happen simultaneously. The settlement is triggered by the sender of securities and reserves are automatically deducted from the Fed account of the institutions receiving the securities and credited to the Fed account of the institution sending the securities.

This procedure creates a "first come first serve" constraint. In the case of a run, lenders who send the securities they hold as collateral early are more likely to receive cash than lenders who send their securities late. With bilateral repos, the timing is as follows:

1. Possibly a sunspot occurs.
2. The borrower offers a new repo contract $(\widehat{r}_i, \widehat{Q}_i, \widehat{k}_i)$.
3. New and patient middle-aged lenders decide whether to engage in new repos with the borrower.
4. Patient middle-aged lenders are repaid in the order in which they send back their collateral, until the borrower runs out of cash. From that point on, lenders receive their collateral and any lender who chooses to invest receives his collateral.

The total amount of collateral available is as before. Yet, borrowers can now reduce their borrowing level by changing $\widehat{Q}_i$, which effectively allows

[26] Also, a borrower may choose to stagger the terms of its repos, so that only a small portion of these repos are due on any given day. Because of the distribution of lender liquidity needs, this cannot happen in our model. He and Xiong (2010) analyze the consequences of (exogenously determined) staggered short-term debt for the stability of financial institutions.

them to increase the collateral per unit borrowed. In order to withstand the run, the borrower must at least cover the missing amount

$$m_i \equiv (\overline{r} + (1-\alpha)\overline{r}^2)b_i - R_i\overline{I_i} \tag{21}$$

At the time when he must pledge the collateral the borrower has $\overline{I_i}$ units, which will mature in $t+1$. Hence, the maximum possible value of collateral per unit borrowed is

$$\overline{k}_i = \overline{I_i}/m_i^t. \tag{22}$$

Again, there are two different investor groups the borrower can borrow from, young lenders who hold cash and middle-aged lenders who hold a repo with the borrower that may be rolled over.

	other lenders invest	other lenders don't
invest	$\widehat{r}_i$	$\gamma_i^{t+1}R_i\widehat{k}_i$
don't	$\overline{r}$	$\overline{r}$

Table 2: Payoffs to young lenders in bilateral repos

Table 2 gives the payoff to an individual young lender as a function of the collective behavior of all other lenders. The payoffs are as in Table 1, with the exception that the promised collateral can differ from the steady-state value. Hence, the run outcome (don't, don't) is not a strict equilibrium if and only if

$$\gamma_i^{t+1}R_i\widehat{k}_i \geq \overline{r} \tag{23}$$

Now, if the funding shortfall m_i is small, the borrower can increase his collateralization beyond κ_i, and this condition is weaker than (19) in the tri-party context.

Note that the borrower can attract as many young lenders as necessary to fund the shortfall m_i if he has the collateral, because he can compete away lenders from other borrowers if his offer is sufficiently attractive. Inserting m_i from (21) into (22) yields the collateral constraint of the following proposition.

Proposition 7 *In bilateral repo markets, a run on borrower i can occur and bankrupt the borrower if and only if the borrower's collateral constraint*

$$\beta^2 R_i \overline{I_i} \geq \frac{1-\alpha+\beta}{1+\gamma_i^t \beta} b_i \tag{24}$$

is violated.

Proof. Condition (24) is (23) evaluated at $\overline{k}_i = \overline{I_i}/m_i$. We already have shown that this condition is sufficient to prevent a run, because young lenders will fund the shortfall if it holds. In order to prove necessity, we must examine the incentives of middle-aged patient lenders to roll over their existing repos.

Suppose therefore that condition (24) is violated. From (21), only a fraction

$$\varphi \equiv \frac{R_i \bar{I}_i}{b_i \left[\overline{r} + (1-\alpha)\overline{r}^2\right]} \in (0,1) \tag{25}$$

of middle-aged lenders can stop renewing their repos before the dealer becomes illiquid. With probability $1-\varphi$, patient middle-aged lenders who run are forced to keep their collateral. Lenders who are able to obtain their cash back can invest it with another borrower. The payoffs of patient middle-aged lenders (per unit of funds) are therefore as in the following table.

	other lenders	
	invest	don't
invest	$\widehat{r}_i$	$\gamma_i^{t+1} R_i \widehat{k}_i$
don't	$\overline{r}$	$\varphi\overline{r} + (1-\varphi)\gamma^{t+1} R_i k_i$

Table 3: Payoffs to middle-aged patient lenders in bilateral repos

Table 3 differs from Table 2 in the lower right cell, which reflects the different positions of young and patient middle-aged lenders. The outcome (don't, don't) is strictly optimal for the individual patient middle-aged investor if and only if

$$\gamma_i^{t+1} R_i \widehat{k}_i < \varphi\overline{r} + (1-\varphi)\gamma_i^{t+1} R_i k_i \tag{26}$$

This condition holds for all k_i and $\widehat{k}_i$ iff it holds for $\widehat{k}_i = \overline{k}_i$ from (22) and $k_i = 1/\beta^2 R_i$ from (13). Re-writing (26) for these two extreme values and setting $d_i = \overline{I}_i/b_i$ yields

$$\gamma_i^{t+1} R_i \frac{d_i}{\overline{r} + (1-\alpha)\overline{r}^2 - R_i d_i} < \varphi\overline{r} + (1-\varphi)\gamma_i^{t+1} R_i \frac{1}{\beta^2 R_i} \tag{27}$$

$$\Leftrightarrow \quad \frac{\gamma_i^{t+1}\beta^4 d_i R_i}{1-\alpha+\beta-\beta^2 d_i R_i} < \frac{\beta^3 d_i R_i}{1-\alpha+\beta} + \gamma_i^{t+1}\frac{1-\alpha+\beta-\beta^2 d_i R_i}{1-\alpha+\beta} \tag{28}$$

Since (24) is violated, we have

$$1-\alpha+\beta-\beta^2 d_i R_i > \gamma_i^{t+1}\beta^3 d_i R_i \tag{29}$$

Hence, (28) is equivalent to

$$(1-\alpha+\beta)\gamma_i^{t+1}\beta^4 d_i R_i < \left[\beta^3 d_i R_i + \gamma_i^{t+1}(1-\alpha+\beta-\beta^2 d_i R_i)\right](1-\alpha+\beta-\beta^2 d_i R_i) \tag{30}$$

Suppose first that $\gamma_i^{t+1} > \beta$. By (29), it is enough to show that

$$\begin{aligned} & \beta(1-\alpha+\beta) \leq \beta^3 d_i R_i + \gamma_i^{t+1}(1-\alpha+\beta-\beta^2 d_i R_i) \\ \Leftrightarrow \quad & (\beta-\gamma_i^{t+1})(1-\alpha+\beta) \leq (\beta-\gamma)\beta^2 d_i R_i \end{aligned}$$

which is implied by (29).

Now suppose that $\gamma_i^{t+1} \leq \beta$. (30) is linear in γ_i^{t+1} and holds for $\gamma_i^{t+1} = 0$ and for $\gamma_i^{t+1} = \beta$. Hence, it holds for all $\gamma_i^{t+1} \leq \beta$.

Finally, note that condition (24) is strictly weaker than the liquidity constraint (18). Hence, if it is violated, (18) is violated as well, and (24) is necessary and sufficient for the stability of bilateral repos. ■

As condition (20) in the tri-party case, condition (24) is implied by the steady-state borrowing restriction (12) if γ_i^t is close to 1 and stronger if γ_i^t is small. Hence, for "good" collateral in "normal" times, the collateral constraint is slack, and it becomes relevant only in "stress" times.

Furthermore, and differently from the tri-party case, condition (24) is strictly weaker than the liquidity constraint (18). Hence, if it is violated, (18) is violated as well. This means that (24) is necessary and sufficient for the stability of bilateral repos.

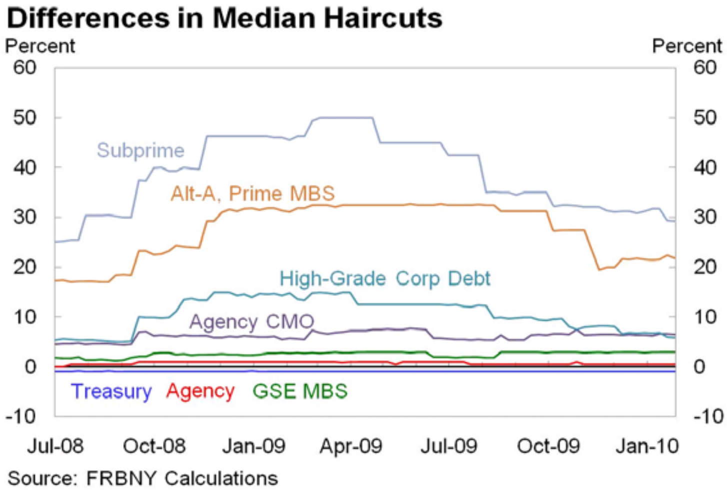

Figure 1: Differences in median haircut between bilateral and tri-party repos per asset class

Finally, the bilateral collateral constraint is strictly weaker than the tri-party constraint (20). This implies that there are borrowers who are run-proof in the bilateral repo market but can fail in the tri-party market. In this sense, the tri-party market is more fragile than the bilateral market. This problem is exacerbated by the fact that cash investors in the tri-party market are generally considered to be less sophisticated and more restricted in processing collateral than the lenders in the bilateral market, hence have a lower γ_i^t.[27]

Our analysis of the bilateral market has assumed that collateral can adjust in response to a run and has shown that this can be achieved by reducing borrowing and is indeed optimal. This is consistent with the evidence in Gorton and Metrick (2011) of sharply rising haircuts during the crisis of 2008.[28] However, the behavior of haircuts was very different in the tri-party and bilateral repo markets. Figure 1 provides some graphical evidence of

[27] See, e.g., Krishnamurthy, Nagel and Orlov (2011, pp. 9-10).

[28] If the price of the collateral (the loan size) is p and the market value of collateral is v, then the haircut is $(v - p)/v$.

this striking difference, taken from Copeland, Martin, and Walker (2010). In the tri-party repo market, haircuts barely moved (this information is not in the figure) while there were large increases in haircuts in some bilateral repo markets. Lehman experienced a sudden reduction of funding in the tri-party repo market that led to its downfall with hardly any adjustment in haircuts. We are not aware of similar sudden losses of funding in the bilateral repo market. Instead, all institutions in this market saw a gradual increase in haircuts that reduced the amount of funding they could obtain (Gorton and Metrick, 2011). Our results in Sections 5.1 and 5.3 are consistent with these two different developments in the bilateral and tri-party repo markets.

5.4 Money market mutual funds

In this section, we adapt our model to the case of money market mutual funds (MMMFs) that can offer shares at a fixed net asset value (NAV). These funds are also known as 2a-7 funds, named after SEC rule 2a-7. MMMFs offer their investors shares that can be redeemed at a fixed price, typically \$1. Positive returns by the fund increase the number of shares, without affecting the shares' price. If the fund loses value, however, the number of shares cannot decrease. In such a case, the fund is said to have "broken the buck" and is liquidated. Investors' shares give them a pro-rata claim on the proceeds from the liquidation of the assets.

The fixed NAV makes MMMFs similar to banks since, under most circumstances, investors can obtain their funds on demand at a fixed price. However, MMMFs do not hold a capital buffer and do not have access to the discount window. MMMFs invest mainly in marketable safe assets, such as ABCP-backed special investment vehicles, in ABCP directly, and other short-term notes. As a percentage of their balance sheet, MMMFs have invested relatively little in tri-party repos backed by non-Agency MBS/ABS (and hardly anything in bilateral repos), although overall they were an important source of funds to the tri-party repo market.[29] In contrast to repo investors, MMMF investors do not have a claim on a specific piece of collateral.

[29] See Krishnamurthy, Nagel and Orlov (2011).

In our framework, a MMMF can be thought of as an agent who invests $I_i = b_i \leq \overline{I_i}$ and offers to pay investors a short-term "interest rate" r obtained by increasing their shareholdings by $100(r-1)$ percent. Since MMMFs do not invest capital of their own, the argument used to establish the dynamic participation constraint (9) cannot be applied in this context. However, this characterization ignores the important role played by MMMFs' parent institutions. A MMMF is typically part of a larger financial institution that provides start-up funding, is the claimant to returns on the form of fees, and even provides discretionary financial support if the MMMF experiences difficulties. Support by parent institutions has been an important source of stability for MMMFs during the recent financial crisis and earlier episodes, as documented by Shilling, Serrao, Ernst, and Kerle (2010).

When applied to the parent institution, the same argument as in Lemma 6 shows that the MMMF's implied interest rate in steady state equilibrium must be $\overline{r} = 1/\beta$. Hence, Proposition 1 applies, with the exception that investment $I_i = b_i \leq \overline{I_i}$ is required to equal borrowing.

Abusing our terminology slightly and recognizing the important role of the parent institution, we can describe the run scenario for a MMMF by the following extensive form.

1. Possibly a sunspot occurs.

2. The MMMF offers a new contract $(\hat{r}_i, \hat{Q}_i)$.

3. New and patient middle-aged investors decide whether to withdraw from the MMMF.

4. The parent institution decides whether to inject liquidity into the MMMF.

5. Investors who redeem their shares get cash until the MMMF runs out. At that time, the MMMF has broken the buck and the remaining investors get a pro-rata claim on the fund's illiquid assets.

In our simple framework, the parent company will always inject liquidity in stage 4 if the fund is illiquid, because the fund is in principle profitable. The only reason why the parent may not do so in our model is that the parent, too, does not have sufficient liquidity. This was indeed the case in 2008 and threatened to bring down the whole money market fund industry in September.[30]

Compared to our lead example of Section 2, the liquidity of MMMFs therefore differs in two respects. First, MMMFs do not invest beyond the level of their short-term funding b_i. As (14) shows, this reduces their liquidity and thus tightens their liquidity constraint (18). Second, however, MMMFs can obtain liquidity support from their parent, which loosens their liquidity constraint. If the parent is expected to inject sufficient liquidity in stage 4 of the game, the fund is expected to be liquid, and there is no run in stage 3. In order to analyze the run scenario, we therefore assume that the liquidity constraint is violated and that the parent does not inject liquidity.[31] Since the liquidity constraint is violated, the withdrawals $b_i\left[\bar{r} + (1-\alpha)\bar{r}^2\right]$ exceed the fund's cash $R_i b_i$, which implies

$$\beta^2 R_i < 1 - \alpha + \beta.$$

As in (25), the probability that a withdrawing investor is able to obtain cash therefore is

$$\varphi = \frac{R_i}{\bar{r} + (1-\alpha)\bar{r}^2} \in (0,1).$$

With probability $1-\varphi$, the investor is unable to withdraw quickly enough to obtain cash. The investor thus gets a claim on the fund's assets. The amount of these assets divided by the total claims outstanding is

$$\mu_i \equiv \frac{I_i}{I_i\left[\bar{r} + (1-\alpha)\bar{r}^2\right] - R_i I_i}.$$

[30] Perhaps the most prominent case was that of the Reserve Primary Fund, which broke the buck on September 16. "Despite efforts to calm share holders in the Primary Fund, Bruce Bent II reported to the board that morning that redemption requests as of 9 A.M. stood at $24.6 billion. He also told the board that Reserve Management had not arranged any credit facility or injected any capital to maintain the one-dollar net asset value. And State Street had refused to extend additional overdraft privileges to the fund. The parent company, Reserve, did not have adequate capital to buy the Lehman assets at par. The Bents were unable to inject any of their own personal funds, contrary to representations they had made the previous day" (James Stewart, *New Yorker*, 9/21/2009).

[31] More generally, the parent may be able to inject some cash, but not enough to plug the liquidity hole m_i. In this case, the parent will optimally not inject any cash at all, because the fund will not survive anyhow and the cash will go to the investors.

Note that the denominator is again $m_i \equiv (\overline{r}+(1-\alpha)\overline{r}^2-R_i)I_i$. The payoffs to middle-aged patient investors as a function of how the other middle-aged patient investors behave are therefore given by the following matrix.

	other investors	
	invest	don't
invest	$\widehat{r}_i$	$\gamma_i^{t+1}R_i\mu_i$
don't	$\overline{r}$	$\varphi\overline{r}+(1-\varphi)\gamma_i^{t+1}R_i\mu_i$

Table 4: Payoffs to middle-aged patient investors in MMMFs

If $\mu_i \geq \bar{r}$, then investors do not have a strict incentive to run on an MMMF. Rewriting this condition we get

$$\beta^2 R_i \geq \frac{1+\beta-\alpha}{1+\gamma_i^t\beta}. \tag{31}$$

Interestingly, this condition is the same as (24), evaluated at $b_i = \overline{I}_i$. Note, however, that condition (31) is independent of fund size (which is equal to outside funding). This is consistent with the observation that the crisis of MMMFs in the wake of the Lehman bankruptcy hit funds across the board, regardless of their size. Again, if (31) is violated, so is the liquidity constraint. Hence, if (31) is violated, the survival of the fund depends on whether the parent company has the cash m_i necessary to stabilize the fund. As (21) shows, this cash shortfall depends on the size of the fund.

5.5 Asset-backed commercial paper conduits

In this subsection we briefly describe the structure of ABCP conduits that Acharya, Schnabl, and Suarez (2009) and Krishnamurthy, Nagel, and Orlov (2011) have identified as an important destination of funds in the shadow banking system and as a main mechanism of contraction during the crisis. While Krishnamurthy, Nagel, and Orlov (2011) consider the evolution of funding from 2007 to 2009 more broadly, Covitz, Liang, and Suarez (2009) focus on the turmoil of the ABCP market in the second half of 2007, which marked the onset of the Great Financial Crisis.

ABCP conduits are institutions that are "sponsored" (i.e., set up, managed, and guaranteed) by banks mainly for the purpose of regulatory arbitrage (or to "optimize yield"). They mostly invest in relatively short-term assets such as receivables or notes and are funded by commercial paper that is of very short maturity. Covitz, Liang, and Suarez (2009) report that "more than half of ABCP daily issuance has maturities of 1 to 4 days, and the average maturity of outstanding paper is about 30 days" (p. 7). ABCP can be liquidated daily, and ABCP conduits are usually opaque. However, unlike traditional banks they are not insured by the government and rather rely on the liquidity support by their sponsoring bank, very much like MMMFs.

We do not provide a formal model of ABCP conduits, which would be similar to that of MMMFs sketched previously, and only report the findings of Covitz, Liang, and Suarez (2009) about the precipitous fall in ABCP finance in August 2007. They find a decrease of outstanding ABCP of $187 billion, almost 20 percent, in August alone, which moreover was mostly concentrated in the two weeks following August 9. More importantly, they analyze the incidence of runs, defined as weeks in which a conduit has more than 10 percent of its outstanding paper maturing but does not issue new paper. Their most important econometric finding, corroborated by various robustness checks, is that "runs are related importantly to program fundamentals, but there is strong evidence that programs that would be sound in more stable market conditions were also subject to runs in the early weeks of the financial crisis" (p. 19).

5.6 Traditional banks

The investors in traditional banks, depositors, are different from the money market participants whom we have considered up to now. Although much of the analysis in Sections 2 - 4 does not change in substance for the case of banks, the no-arbitrage argument underlying Lemma 1 does not apply to depositors. However, in this case it is appropriate to assume that the utility functions u_1 and u_2 are strictly concave, which again implies a flat yield curve, as shown by Qi (1994). Apart from that the analysis for traditional banks is similar to the analysis for MMMFs. With $b_i < \bar{I}_i$, the assets $\left(\bar{I}_i - b_i\right)(1+\beta)$ can be thought of as the equity of the bank. Like MMMF investors, bank depositors do not get a claim on a specific piece of collateral, but rather a claim on the bank's assets in case of bankruptcy. The major difference between a MMMF and a bank is that banks hold largely nonmarketable assets. Hence, the outside value of assets γ_i^t is low in the case of a bank.

The timing of bank funding in our model structure is as follows.

1. Possibly a sunspot occurs.

2. The bank offers a new deposit contract (r_i, Q_i).

3. New and patient middle-aged investors decide whether to deposit (again) with the bank.

4. Investors can withdraw cash until the bank runs out. At that time, the bank is bankrupt and the remaining investors get a claim on the remaining assets.

The analysis and the payoff table is as in the case of a MMMF, with the exception that the bank (hopefully) has equity, i.e. that $b_i < \bar{I}_i$. The collateral constraint therefore becomes

$$\beta^2 R_i \overline{I_i} \geq \frac{1-\alpha+\beta}{1+\gamma_i^t \beta} b_i$$

which is identical to the bilateral constraint (24). The main difference here is that the collateral value γ_i^t of assets of a failing bank is likely to be very low. Hence, the collateral constraint is unlikely to be satisfied and the liquidity constraint (18) thus crucial for bank stability.

Our work therefore nests the classic literature on bank stability which emphasizes the importance of liquidity. It adds to this literature by endogenizing the profits that can serve as liquidity buffers and therefore can make predictions which banks are likely to be subject to runs if investor sentiment changes.

6 Runs and Asset Sales

In this section, as in Section 4, we ask whether, if behavior until date t is steady state as in Proposition 1, the collective refusal to lend to a dealer can bankrupt the dealer. However, as pointed out by, e.g., Shleifer and Vishny (1992) or Diamond and Rajan (2011), "fire sales", i.e. asset sales under distress, can mitigate the dealer's illiquidity problem. We therefore introduce the possibility of asset sales as a reaction to a run and thus generalize the analysis of Section 4.

To investigate this possibility, consider a dealer, say i, at date t who holds assets that will yield $R_i\overline{I_i}$ at date $t+1$. We assume that in response to a run, the dealer can sell these assets to other dealers at some market price p. If the dealer under distress sells an amount A of assets, this improves his current liquidity by pA and reduces his cash at date $t+1$ by R_iA. Generalizing (17), his cash position after the run at date t therefore is

$$I_0 = R_i\overline{I_i} + pA - (\overline{r} + (1-\alpha)\overline{r}^2)b_i \tag{32}$$

Since the maximum amount of assets the dealer can sell is $A = \overline{I_i}$, (32) implies that the dealer can survive if and only if

$$(R_i + p)\overline{I_i} - (\overline{r} + (1-\alpha)\overline{r}^2)b_i \geq 0 \tag{33}$$

If p satisfies (33) the dealer will survive by selling a sufficient amount of assets, if not he will be bankrupt. Whether the dealer can raise enough cash through the asset sale depends on the cash in the market (Allen and Gale 1994), i.e. on the total amount of cash held by all other dealers. At the moment of the run, i.e. when the dealers have repaid their steady-state borrowing, received their new loans including the funds $b_i + (1-\alpha)\overline{r}b_i$ that have not gone to dealer i, but before they have invested their funds, this cash is

$$\begin{aligned} C_i &= b_i + (1-\alpha)\overline{r}b_i + \sum_{j\neq i}\left[R_j\overline{I_j} - (\alpha\overline{r} + (1-\alpha)\overline{r}^2 - 1)b_j\right] \\ &= N + \sum_{j\neq i} R_j\overline{I_j} - (\alpha\overline{r} + (1-\alpha)\overline{r}^2)(N - b_i) + (1-\alpha)\overline{r}b_i \end{aligned}$$

By (3) and (14), this cash is clearly sufficient to cover dealer i's missing amount m_i as defined in (21); intuitively, the run on the dealer simply means a redistribution of his liquidity to the other dealers.

The question is whether this cash can be mobilized to save the dealer. The benefits from mobilizing this cash are the asset returns in $t+1$, the cost is the foregone investment that yields benefits in $t+2$ and thereafter. The demand for cash is easily described. The dealer must raise m_i. From (33), the proceeds from the asset sale will be sufficient to cover m_i if and only if

$$p \geq \frac{m_i}{\overline{I}_i} \equiv \underline{p_i}$$

Since the assets sold by the dealer yield only $\hat{\gamma}_i R_i$ to outsiders next period, the demand for these assets, hence the supply of cash, will be 0 if $p > \beta\hat{\gamma}_i R_i$. This implies the following characterization of when asset sales can save a distressed dealer.

Proposition 8 *Asset sales give a distressed dealer sufficient liquidity if and only if* $\underline{p_i} \leq \beta\hat{\gamma}_i^t R_i$, *which means*

$$\beta^2 R_i \overline{I_i} \geq \frac{1-\alpha+\beta}{1+\hat{\gamma}_i^t \beta} b_i \tag{34}$$

Proof. Suppose $\underline{p_i} > \beta\hat{\gamma}_i R_i$. Then either $p \geq \underline{p_i}$, in which case other dealers do not purchase the dealer's assets, or $p < \underline{p_i}$, in which case the price is too low to save the dealer.

Now suppose that $\underline{p_i} \leq \beta\hat{\gamma}_i R$. Consider any $p < \beta\hat{\gamma}_i R$. If all dealers $j \neq i$ invest $\overline{I_j}$ into their assets as in steady state they have a total of $\sum_{j \neq i} \pi_j + (1+(1-\alpha)\overline{r})b_i$ in cash. Investing this cash into the distressed dealer's assets yields a return of $\hat{\gamma}_i R_i/p$ next period, which is strictly preferred to consuming the cash now. This cash is sufficient to cover m_i, because of positive profits, (14), and because $b_i \leq \overline{I_i}$ implies

$$(1+(1-\alpha)\overline{r})b_i \geq (\overline{r}+(1-\alpha)\overline{r}^2)b_i - R_i\overline{I_i}$$

Figure 2 provides a graphical illustration.

■

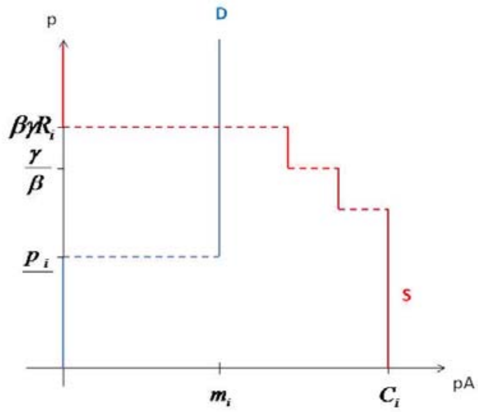

Figure 2: Demand and supply of cash in the asset market

Condition (34) is exactly the same condition as the collateral constraint for the bilateral repo market, (24), if $\hat{\gamma}_i^t = \gamma_i^t$ and is weaker if $\hat{\gamma}_i^t > \gamma_i^t$. If $\hat{\gamma}_i^t = \gamma_i^t$, investors and other dealers realize the same return from the assets. Hence, asset sales cannot loosen a dealer's liquidity constraint beyond the limitations of the collateral constraint. In contrast, if $\hat{\gamma}_i^t > \gamma_i^t$ dealers realize a higher return from the assets than investors would. Other dealers compete to purchase the assets, raising their price up to the point where the returns are greater than the return investors would get from the assets.

Condition (34) is strictly weaker than the liquidity constraint (18), confirming the intuition that sales of assets relax the dealer's liquidity constraint. In other words, some dealers who would go bankrupt if they could not sell their assets can survive if asset sales are possible. However, if in a distressed asset sale $\hat{\gamma}_i^t$ is sufficiently small, condition (34) does not provide much relief and the dealer is illiquid despite the asset sale.

6.1 Interpretation

A simple calculation shows that (34) holds in steady state if $\widehat{\gamma}_i^t = 1$ and is in general violated if $\widehat{\gamma}_i^t = 0$. Most assets serving as collateral in the tri-party repo market are liquid and of little if any risk, so we should expect $\hat{\gamma}_i^t$ to be close to 1 at least in normal times. Hence, we can interpret the result of this section as suggesting that when markets are not stressed, dealers in the tri-party repo market are expected to accommodate the demand that would arise from an idiosyncratic run. This is broadly consistent with the conventional wisdom before the financial crisis.

In other wholesale funding markets such as ABCP and bilateral repo, collateral is known to have been of lower quality on average (see, e.g., Pozsar, Adrian, Ashcraft, and Boesky, 2010). In these markets, the liquidity constraint (34) is more likely to be violated, in particular for some types of borrowers. In stress times, the main determinant of fragility in these markets is therefore likely to be the relevant collateral constraint. As argued above, in the ABCP market in August 2007, the market seems to have viewed this constraint as violated in a number of SIVs.

There are two cases where we might expect $\hat{\gamma}_i^t$ to be low also in the tri-party repo market. A low $\hat{\gamma}_i^t$ should be expected for the less liquid or riskier collateral used by dealer to back its repos. In such a case, it will be more difficult for whoever tries to liquidate the collateral to obtain a high value. Anecdotal evidence suggests that the share of less liquid collateral in the tri-party repo market had been increasing before the crisis, probably reaching 30 percent of the collateral in that market before Bear Stearns on average.[32] This would have made dealers who borrow against such collateral more susceptible to runs.

A low $\hat{\gamma}_i^t$ may also apply if the quantity of a relatively liquid asset used as collateral in tri-party repos is so large that the market may not be able to absorb all the collateral in case of a dealer default. For example, Agency

[32] This is not inconsistent with the estimate of Krishnamurthy, Nagel, and Orlov (2011) that repos only accounted for around 8 percent of all holdings of non-Agency MBS/ABS by MMMFs and securities lenders (p. 21). Their numbers are about the asset side of these cash lenders, whereas we are interested in the liability side of dealers on the other side of the tri-party repo market.

MBSs are considered liquid securities, but the amount of such securities financed in tri-party is so large that the market may not have been able to absorb them without some price effect. This effect is likely to be particularly strong in times of aggregate market stress, which we discuss in the next section. Indeed, Paul Friedman, Senior Managing Director at Bear Stearns, testified before the Financial Crisis Inquiry Commission on May 5, 2010, that "repo market lenders declined to roll over or renew repo loans, even when the loans were supported by high-quality collateral such as agency securities."[33]

It is also worth pointing out that our model probably overstates dealer's ability to accommodate the demand for cash in a run and the ability of other dealers to purchase assets. In our model, the share of repos held by old and impatient middle aged investors is close to half of all the repos made by a dealer.[34] Hence, the demand for funds in the case of a run is about twice as large as the steady state demand. Anecdotal evidence suggests that the share of repos being rolled over in the tri-party repo market is much larger, probably over 80 percent. This would mean that the run demand is five times as large as the steady-state demand, which would be more difficult for a dealer to accommodate.

Our model could be adapted to increase the share of repos rolled over every period. For example, we could consider an economy in which agents lived longer lives and assets matured after more periods. In such an economy, the share of cash and maturing assets would be a smaller share of all assets. Similarly, the share of new and withdrawing investors, which must be equal in steady state, would represent a smaller fraction of the population of all investors. Hence, the demand for funds in case of a run would be much larger than the steady state demand, compared to the economy we consider. The share of unmatured assets that can be sold, compared to the available cash, would also be greater, increasing the fire sale effect.

[33] See "http://fcic-static.law.stanford.edu/cdn_media/fcic-testimony/2010-0505-Friedman.pdf"

[34] The exact share will vary depending on the parameters α and β

7 Extensions: Market Runs and liquidity provision

As noted above, the more dealers are in trouble, the more assets troubled dealers are trying to sell and the fewer dealers are available to buy these assets. This puts pressure on the price of assets and makes it less likely that a run can be avoided. In the extreme case of a market run, all dealers are facing a run demand and only a small number may have enough liquidity to satisfy their demand and at the same time buy the assets that are put up for sale by a large number of distressed dealers. This is an extreme version of the externality of short-term debt identified by Stein (2011) and may justify liquidity provision by a lender of last resort.

7.1 Market runs

First, we consider the conditions for the case where no dealer survives. Proposition 2 continues to apply, so a dealer will be illiquid if condition (18) is violated. The collateral constraint is slightly different in the case where no dealer survives. Indeed, in that case, investors who do not survive get a payoff of 1, rather than $\bar{r}$. Hence, the collateral constraint is $1 \leq \gamma_i^t R_i k_i$. After replacing k_i with κ_i we can write

$$\beta R_i \overline{I_i} \geq \frac{1-\alpha+\beta}{\gamma_i^t(1+\beta)} b_i. \tag{35}$$

This condition is less restrictive than (20) because investors do not have as good an outside option. If all dealers' collateral and liquidity constraints are violated, no dealer survives the market run.

If at least one dealer survives the market run, then this dealer can attract depositors and purchase assets from other dealers. If only one dealer survives, this dealer may be able to act monopolistically. To simplify the argument, we assume here that at least two dealers survive and behave competitively.

If there are no restrictions on investors switching from troubled to healthy dealers, the healthy dealers can attract all investors from troubled dealers and use these fund to purchase assets.[35] In this case, the same argument as in section 6 shows that there is enough cash in the market to cover the troubled dealers' need. This is because the total liquidity in the market is simply redistributed among all dealers. In this case, Proposition 8 applies. Skeie (2004) obtains a similar result in the context of traditional banks.

However, it may be difficult for healthy dealers to take on all the troubled dealers' investors. On-boarding new clients can be costly and take time and there may simply be economies of scale similar to (2). When the redistribution of liquidity among dealer is not frictionless, there may not be enough cash in the market to cover the troubled dealers' need. If the supply of cash is sufficient, then proposition 8 applies. If the supply of cash in insufficient, then troubled dealers will bid down the price of assets until $p_i = \beta\widehat{\gamma}_i^t R_i$. If the price of the assets drops any lower, then the troubled dealer's investors generate a higher yield from the assets. In this case, the collateral constraint is

$$\beta^2 R_i \overline{I_i} \geq \frac{1-\alpha+\beta}{1+\widehat{\gamma}_i^t\beta} b_i. \tag{36}$$

Note that condition (36) is tighter than condition (34). This suggests that multiple equilibria may be possible. If investors expect low asset prices, more dealers will be in trouble, which increases the supply of assets in the market. Because the redistribution of liquidity among dealers is not perfect, the supply of cash in the market does not increase as much, which justifies the low price of assets. In contrast, if the price of assets is expected to be high, then fewer dealers are in trouble. This means that the supply of assets is low and the available cash is high, justifying high prices.

7.2 Liquidity provision

Access to a lender of last resort is a standard tool used to strengthen the banking sector in the face of financial fragility. Theoretical work has shown how access to a lender of last resort can prevent bank runs (see, for example, Martin 2006, Skeie 2004). In the U.S., the broker dealers that rely on the triparty repo market as a source of short-term funding did not have direct access

[35] This occurs, for example, if the capacity limit $\bar{I}_i$ applies to investment, but not the stock of assets held.

to the discount window. This lack of access to emergency liquidity proved destabilizing during the crisis and motivated the Federal Reserve to introduce the Primary Dealer Credit Facility (PDCF). Similar concerns about MMMFs, who represent an important share of investors in the tri-party repo market, motivated the creation of the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF), and the Money Market Investor Funding Facility (MMIFF). These facilities were temporarily created under section 13.3 of the Federal Reserve Act, which allows the Federal Reserve to lend to a variety of institutions under unusual and exigent circumstances.

The Task Force on Tri-Party Repo Infrastructure (2009) noted the need to "consider establishing an industry-sponsored utility with the ability to finance the securities portfolio of a faltering or defaulted dealer and limit the associated stress on the market while their portfolio is liquidated." The model in our paper suggests that there would be benefits to the creation of a lender-of-last-resort facility for wholesale funding markets in general and the tri-party repo market in particular.[36] The argument is similar to the case of banking. In case of a run, investors do not refuse to roll over their loans because they need cash, but because they are concerned about the default of the dealer and having to hold collateral that they might have to liquidate. As in classic banking theory, a lender of last resort could lend cash to the dealer taking securities as collateral. The cash could be used to pay all investors who do not roll over their loans. This would prevent the default of the dealer and allow it to manage the collateral until it matures. Knowing that the dealer will not default, investors no longer have to worry about having to hold or liquidate assets, so their incentive to run is reduced.

8 Conclusion

We have studied a model of short-term collateralized borrowing and the conditions under which runs can occur. The framework resembles the dynamic bank model studied in Qi (1994), but extends that model beyond the

[36] A full discussion of the policy implications of our model is beyond the scope of this paper. Stein (2011) provides an illuminating discussion of monetary policy when short-term debt imposes externalities on the econmy.

pure theory of commercial banking. We derive a dynamic participation constraint that must hold for dealers to agree to purchase securities on behalf of investors. Under this constraint, dealers will make profits that can be mobilized to forestall runs.

Our model sheds light on the panic in the ABCP market in August 2007 that triggered the Great Financial Crisis and on the puzzling behavior of margins in different repo markets. We can account for the difference between the bilateral repo market, where haircuts increased dramatically during the crisis, and the tri-party repo market, where the haircuts barely moved. The model also clarifies the distinction between increasing margins, which is a potentially equilibrating phenomenon, and runs, which can happen if margins do not increase sufficiently to reassure investors. The model also shows that the practice of early settlement of tri-party repos, called the "unwind", can increase fragility in the market; this result lends support to reforms currently underway to eliminate the unwind. Our results on the particular fragility of the tri-party repo market show how a lack of increase in haircuts and the practice of "unwind", each of which may appear to provide additional liquidity for dealers in normal times, actually can explain the sudden collapse of securitized lending that contributed to the runs on Bear Stearns and Lehman Brothers.

A key difference between traditional banks and modern financial intermediaries is that the former mainly hold opaque assets while the latter's assets are much more liquid and marketable. The value of collateral and the liquidity of asset markets therefore become crucial for the liquidity of intermediaries. Runs can be forestalled by mobilizing sufficient liquidity and having sufficiently valuable collateral. This gives rise to two constraints that can be interpreted as a liquidity and a collateral constraint. As analyzed in Section 6, part of a dealer's liquidity comes directly from his balance sheet and part from the ability to liquidate illiquid assets. It is therefore tempting to augment Brunnermeier and Pedersen's (2009) distinction between market liquidity and funding liquidity by the notion of "balance sheet liquidity". This balance sheet liquidity has been isolated in Section 4 by ruling out asset sales. Even in the general model, however, Section 4 becomes relevant again in the case of a market run, in which no dealer may be able to purchase other dealers' assets.

It should be noted that we consider runs only as out-of-equilibrium phenomena. They are triggered by sunspots that occur with probability 0. If in a more general model sunspots occur with some probability q, then our model corresponds to the limiting case $q \to 0$. Since the more general model would be continuous in q, our results carry over to equilibrium sunspots that occur with sufficiently small probability. This is the standard practice in other dynamic models such as Kiyotaki and Moore (1997), Brunnermeier and Pedersen (2009) or Uhlig (2010).

Our framework can be used to consider a number of policy questions related to the fragility of short-term funding markets. For the tri-party repo mechanism, for example, Lehman's demise highlighted the problem that there is no process to unwind the positions of any large bank that deals in repo should it fail. Lehman required large loans from the Federal Reserve Bank of New York to settle its repo transactions. Our framework can be used to study a liquidation agent, as suggested in the Task Force on Tri-Party Repo Infrastructure (2009), with the objective to unwind the positions of a defaulting dealer. Similarly, our analysis sheds light on the role of institutional features such as the unwind mechanism in the tri-party market or the difference between bilateral and tri-party repo lending and thus should contribute to a better understanding of the fragility of wholesale banking markets.

9 Appendix: Coordination problem between the clearing bank and investors

The tri-party repo market is also vulnerable to another coordination problem, this time between the clearing bank and the investors. Suppose that, in the timing described in section 5.1, just before step 1 the clearing bank comes to believe that at step 5 all investors will refuse to engage in repos with dealer i. In this case, the clearing bank will refuse to unwind if the loan it makes to the dealer, $b_i\left[\bar{r} + (1-\alpha)\bar{r}^2\right]$, exceeds the proceeds it could obtain from the assets, $R_i\bar{I}_i(1+\beta\gamma_i^t)$.[37] This condition can be written as

$$\beta^2 R_i \bar{I}_i \geq \frac{1+\beta-\alpha}{1+\gamma_i^t\beta} b_i. \tag{37}$$

This condition is the same as the collateral condition for bilateral repos, (24).

The flip side of this coordination problem is that investors may choose not to invest with dealer i if they believe that the clearing bank will refuse to unwind that dealer's repos the next morning.[38] In this case, the condition for investors to have a strict incentive to run is the same as in the case where investors believe other investors may not engage in repos.

[37] Here we assume that the clearing bank faces the same γ_i^t as the investors.

[38] Clearing banks have the contractual right not to unwind a dealer's repos. Failure to unwind the repos would almost certainly force the dealer into bankruptcy.

References

[1] Acharya, V., D. Gale, and T. Yorulmazer (2011), Rollover Risk and Market Freezes, forthcoming, *Review of Financial Studies*.

[2] Acharya, V., S. Myers, and R. Rajan (2010), "The Internal Governance of Firms." forthcoming *Journal of Finance*.

[3] Acharya, V., Ph. Schnabl, and G. Suarez (2009), "Securitization without Risk Transfer", manuscript, New York University.

[4] Admati, A., P. DeMarzo, M. Hellwig, and P. Pfleiderer (2010), "Fallacies, Irrelevant Facts, and Myths in Capital Regulation: Why Bank Equity is Not Expensive." Manuscript, Stanford University.

[5] Allen, F. and D. Gale (1994). "Limited Market Participation and Volatility of Asset Prices." *American Economic Review* 84, 933-955.

[6] Bernanke, B. S. (2009). "Financial Reform to Address Systemic Risk." Council on Foreign Relations, Washington, D.C. March 10, 2009. (http://www.federalreserve.gov/newsevents/speech/bernanke20090310a.htm)

[7] Bhattacharya, S. and J. Padilla (1996). "Dynamic Banking: A Reconsideration." *Review of Financial Studies* 9, 1003-1032

[8] Bhattacharya, S., A. Boot, and A. Thakor (1998). "The Economics of Bank Regulation." *Journal of Money, Credit and Banking* 30, 745-770.

[9] Brunnermeier, M. K. and L. H. Pedersen (2009). "Market Liquidity and Funding Liquidity." *Review of Financial Studies* 22, 2201-2238.

[10] Copeland A., A. Martin, and M. W. Walker (2010). "The tri-party repo market before the 2010 reforms." Federal Reserve Bank of New York Staff Report.

[11] Covitz, D., N. Liang, and G. Suarez (2009). "The Evolution of a Financial Crisis: Panic in the Asset-Backed Commercial Paper Market." Federal Reserve Bank of New York, Staff Report.

[12] Diamond, D. and P. Dybvig (1983). "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91, 401-419.

[13] Diamond, D. W. and R. G. Rajan (2001). "Liquidity risk, liquidity creation, and financial fragility: A theory of banking." *Journal of Political Economy* 109, 287-327.

[14] Diamond, D. and R. Rajan (2011), "Fear of Fire Sales, Illiquidity Seeking, and Credit Freezes." *Quarterly Journal of Economics*, forthcoming.

[15] Duffie, D. (2010). "How Big Banks Fail and What To Do About It." *Princeton University Press.*

[16] Fulghieri, P. and R. Rovelli (1998). "Capital Markets, Financial Intermediaries, and Liquidity Supply." *Journal of Banking and Finance* 22, 1157-79.

[17] Garbade, K. D. (2006). "The Evolution of Repo Contracting Conventions in the 1980s." Federal Reserve Bank of New York *Economic Policy Review.*

[18] Gorton, G. B. and A. Metrick (2011). "Securitized Banking and the Run on Repo." *Journal of Financial Economics*, forthcoming.

[19] Hart, O. and J. Moore (1998). "Default and Renegotiation: A Dynamic Model of Debt." *Quarterly Journal of Economics* 113, 1-41

[20] He, Z. and W. Xiong (2010), "Dynamic Debt Runs." Manuscript Princeton University.

[21] Hellmann, T., K. Murdock, and J. Stiglitz (2000). "Liberalization, Moral Hazard in Banking and Prudential Regulation: Are Capital Requirements enough?" *American Economic Review* 90, 147-165.

[22] Kiyotaki, N. and J. H. Moore (1997). "Credit Cycles." *Journal of Political Economy* 105, 211-248.

[23] Krishnamurthy, A., S. Nagel and D. Orlov (2011). "Sizing up Repo." Working paper, February 2011.

[24] Martin A. (2006). "Liquidity provision vs. deposit insurance: preventing bank panics without moral hazard." *Economic Theory* 28, 197–211.

[25] Pozsar, Z., T. Adrian, A. Ashcraft, and H. Boesky (2010). "Shadow Banking." Working Paper, Federal Reserve Bank of New York.

[26] Qi. J. (1994). "Bank Liquidity and Stability in an Overlapping Generations Model." *Review of Financial Studies* 7, 389-417.

[27] Rampini, A. and S. Viswanathan (2010). "Financial Intermediary Capital." Working paper.

[28] Shleifer, A. and R. W. Vishny (2010). "Unstable Banking." *Journal of Financial Economics* 97, 306-318..

[29] Shilling, Serrao, Ernst, Kerle (2010). "Sponsor Support Key to Money Market Funds." Moody's Investors Service *Special Comment*.

[30] Skeie, D. R. (2004). "Money and Modern Bank Runs." Dissertation, Princeton University.

[31] Stein, J. (2011). "Monetary Policy as Financial-Stability Regulation", *Quarterly Journal of Economics*, forthcoming.

[32] Task Force on Tri-Party Repo Infrastructure (2009). "Progress report." (http://www.ny.frb.org/prc/report_091222.pdf)

[33] Uhlig, H. (2010). "Crises in Repo Markets with Adverse Selection." Manuscript, University of Chicago.

[34] van Bommel, J. (2006). "Intergenerational Risk Sharing and Bank Raids." Manuscript, Oxford University.

[35] von Thadden, E., E. Berglöf and G. Roland (2010). "The Design of Corporate Debt Structure and Bankruptcy." *Review of Financial Studies* 23, 2648-2679.