GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 448

Repo Runs

Antoine Martin *
David Skeie **
Ernst-Ludwig von Thadden ***

* / ** Federal Reserve Bank of New York
*** Universität Mannheim

November, 2013

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Repo Runs

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

final version
November 2013

[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, Amar Bhide, Patrick Bolton, Fabio Castiglionesi, Douglas Gale, Gary Gorton, Todd Keister, Ed Nosal, Lasse Pedersen, Matt Pritzker, Adriano Rampini, Jean-Charles Rochet, Bogdan Stacescu, Jean Tirole, and Jos van Bommel for helpful comments. We are particularly grateful to the editor, Pietro Veronesi, and an anonymous referee. Part of this research was done while Antoine Martin was visiting the University of Bern, the University of Lausanne, and the Banque de France. Support by the Deutsche Foschungsgemeinschaft throug SFB-TR 15 is gratefully acknowledged. 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.

Abstract

The recent financial crisis has shown that short-term collateralized borrowing may be a highly unstable source of funds in times of stress. The present paper develops a dynamic equilibrium model and analyzes under what conditions such instability can be a consequence of market-wide changes in expectations. We derive a liquidity constraint and a collateral constraint that determine whether such expectations-driven runs are possible and show that they depend crucially on the microstructure of particular funding markets that we examine in detail. In particular, our model provides insights into the differences between the tri-party repo market and the bilateral repo market, which were both at the heart of the recent financial crisis.

Keywords: Investment banking, repurchase agreements, tri-party repo, bilateral repo, money market mutual funds, asset-backed commercial paper, bank runs.
JEL classification: E44, E58, G24

The sudden collapse of highly levered financial institutions was a key factor in the financial crisis of 2007-09. In this paper, we develop a model of short-term collateralized finance and study the conditions under which runs can occur in various short-term funding markets. Leveraged financial institutions may not be able to roll over their short-term borrowing, despite it being collateralized. We derive a liquidity constraint and a collateral constraint that depend crucially on the microstructure of the short-term funding market. We model the differences between various segments of the U.S. repo market and similar short-term 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. The primary application of our model is to large securities dealers who use the repo market as a main source of financing. In that market dealers borrow from institutional investors, such as money markets fund (MMFs), against collateral. Dealers' borrowing in the tri-party repo market, a particularly important segment of the repo market, reached over $2.8 trillion outstanding in aggregate at its peak in 2008; individual dealer borrowing reached $400 billion, most of it with overnight maturity.

Data about the repo market is limited. Copeland, Davis, LeSueur, and Martin (2012) estimate the size of the various segments of the U.S. repo market as of May 2012 and find that the value of all repos was about $3 trillion, with approximately two-thirds from the tri-party market, both interdealer and investor-to-dealer, and one-third from the bilateral market. The value of reverse repos was a little less than $2.5 trillion, with almost $2 trillion coming from the bilateral repo market. The same methodology can be used to estimate the size of the repo market as of July 2008. Total repos represented $6.1 trillion at the time, with bilateral repos representing almost 60 percent of the value. The decrease in the size of bilateral repos is consistent with Gorton and Metrick (2013). The value of reverse repos was about $4 trillion, with the bilateral market accounting for close to $3.7 trillion.[2]

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 collateral-

[2] See Garbade (2006) for an excellent overview of the history of repos and Duffie and Skeel (2012) for a recent debate about the social value of this type of debt instrument.

ized 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."[3] The Federal Reserve took exceptional policy measures to reduce the stress in financial markets and created the Primary Dealer Credit Facility (PDCF) and the Term Securities Lending Facility (TSLF) in an attempt to provide a backstop for the tri-party repo market.

Empirical studies have shown that different segments of the U.S. repo market behaved very differently during the crisis. Gorton and Metrick (2012) and Copeland, Martin, and Walker (2010) show that haircuts increased dramatically in some segments of the bilateral market. Copeland, Martin, and Walker (2010) and Krishnamurthy, Nagel and Orlov (2011) show that in the tri-party repo market, haircuts barely moved. Our model allows us to study the relative fragility of the different segments of these markets and 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. While increases in margins reduce available funding, as argued by Gorton and Metrick (2012), some funding is better than none. Hence, we argue that the bilateral repo market kept functioning, albeit at a reduced level. Lehman Brothers, on the other hand, which depended heavily on funding from the tri-party repo market, was almost completely cut off from such funding within a few days, while margins did not move. The available evidence suggests that Bear Stearns experienced similar stress.

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. Eliminating the unwind has been a major goal of tri-party repo reform, and our model lends theoretical support to this measure. 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 theory uses a simple infinite-horizon rational expectations model in which equilibrium may break down because of coordination failure. An important feature of this model is that borrowers make positive profits from intermediation in equilibrium.

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

This distinguishes our model from Qi (1994) who also considers an infinite-horizon model of liquidity provision along the lines of Diamond and Dybvig (1983). Qi's (1994) banks do not maximize profits, but are assumed to make zero profits. We generalize this framework to profit-maximzing banks and in fact show that competition does not necessarily drive up interest rates to zero-profit levels in equilibrium, because borrowers with liquidity of their own must have an incentive to borrow rather than use their own funds for investment. This equilibrium argument for positive profits relies on a trade-off between the use of external and internal funds, which are endogenous in our infinite-horizon model, but would need to be exogenously specified in a static model.

Hence, borrowers in our model have "inside liquidity" (Bolton, Santos, and Scheinkman (2011)), which arises endogenously. This inside liquidity can be used as a buffer to stave off runs. For that reason, unlike in conventional models of multiple equilibria, not "everything goes" in our model. The theory predicts 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 there exists a liquidity constraint as well as a collateral constraint that is 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). In fact, our model predicts that banks can be subject to such runs if they rely heavily on short-term borrowing, are relatively unprofitable or relatively small. For such firms, illiquidity was indeed an important issue at some key turning points during the crisis.

The remainder of the paper proceeds as follows. Section 1 introduces our model of collateralized lending. Section 2 characterizes the steady states of the model that serve as a benchmark for the rest of the analysis. Section 3 studies the borrowers' ability to withstand runs in terms of liquidity. Section 4 considers the fragility of different market microstructures and derives collateral constraints. The Appendix contains longer proofs, extends the model to related funding markets, such as money market funds and asset-backed commercial paper, and shows how our model nests the traditional model of commercial banking.

1 The Model

The economy is populated by two classes of agents, "investors" and "borrowers". The time horizon is infinite and there is no initial date. Investors live for three dates, and at each date t a continuum of mass N of them is born. On the first date of their lives, "young" investors receive an endowment of 1 unit of goods that they can invest at date t. Investors have no endowment in the following two periods when they are "middle-aged" and "old".

Investors' want to consume either in middle or in old age, but do not know when at the time they are born. The timing of the investors' needs of cash is uncertain because of "liquidity" shocks. Investors learn their type in middle age: "Impatient" investors need cash immediately, while "patient" investors do not need cash until their old age. The information about the investors' type and age is private. In particular, an outsider cannot distinguish between a young and a patient middle-aged investor.

In practice, investors such as MMFs 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. 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). As in the seminal paper by Diamond and Dybvig (1983), their utility from consumption (c_1, c_2) over the two-period horizon then is given as

$$U(c_1, c_2) = \begin{cases} u_1(c_1) & \text{with prob. } \alpha \\ u_2(c_2) & \text{with prob. } 1 - \alpha \end{cases} \tag{1}$$

with u_1 and u_2 strictly increasing.[4] Investors can trade with each other. But since they can generate no value from their endowments, they must invest with intermediaries, to which we turn now.

[4]Note that we do not assume concavity of the utility function. In practice, short-term investors such as MMFs or large corporates may well be risk-averse. But in our framework, risk aversion is not required for the results to hold. Differently from the classical literature on banking following Diamond and Dybvig (1983), investors in our model part with their money because (investment) banks have better investment opportunities, and not for risk-sharing motives.

The economy is also populated by M infinitely-lived risk-neutral agents, called "borrowers", who have profitable investment opportunities, but no endowments. Investors and borrowers have access to a one-period storage technology, which can be thought of as cash and returns 1 for each unit invested. In addition, borrowers have 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. They yield constant returns to scale up to a capacity constraint, which we interpret as the firm's size and which is exogenous. Specifically, 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 $t+2$ and yields nothing at date $t+1$. To simplify things, we assume that the return on these investments is riskless. Yet, the returns are not verifiable, which creates a role for collateral in our model. Indeed, investors cannot be sure that a borrower has realized $R_i I^t$ from his past investment. Although this is a probability zero event, a borrower who has received funds from investors could claim that he is unable repay. Our lead examples are repo transactions, where the collateral actually changes ownership. Sufficient collateral is then needed to give the borrower the incentive to repurchase the collateral.

Investment returns can only be realized by the borrower who has invested in the asset, because borrowers have a comparative advantage in managing their securities portfolio. Other market participants only realize a smaller return. Investors could realize a return of $\gamma_i^t R_i$ from these assets, with $\gamma_i^t < 1$, where the discount reflects different skills in valuing or managing the assets, possible restrictions on the outsider's portfolio composition, transactions and timing costs, and similar asymmetries.[5] We allow γ_i^t to depend on the borrower, reflecting potential differences in the portfolio of collateral that different borrowers seek to finance, and on time, reflecting changing market environments.[6]

[5] For T-bills, γ_i should be very close to 1. But many market participants 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 for each class of collateral without changing the analysis.

[6] In this paper, we take γ_i^t as a reduced-form description of asset market frictions. In

The total investment capacity of borrowers, $\overline{I} = \sum \overline{I_i}$, strictly exceeds the investors' amount of cash available for investment, N, so borrowers must compete for investors' funds. The alternative assumption that investors compete for the right to lend to borrowers is neither realistic nor analytically interesting, because then borrowers 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 borrower is pivotal, and even if one borrower fails, there will still be competition for investor funds.[7]

Since investors cannot generate investment returns on their own, they must invest with borrowers, and borrowers optimally offer them the possibility to request repayment after one period or after two periods, depending on their realized ex-post preferences. Hence, borrowers offer contracts that resemble demand deposits as in the classical banking literature following Diamond and Dybvig (1983) and Qi (1994). This corresponds to Gorton and Metrick's (2012) analogy of shadow banking.[8]

If borrower 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 and r_{2i}^t after two periods, impatient investors do not roll over their funding when middle-aged, but patient investors do, then the borrower's expected cash flow, which we also refer to as profits, is

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

At each date, borrowers consume their profits (or pay them out as dividends). The borrower's objective at each date t, then, is to maximize the

Martin, Skeie, and von Thadden (2012), we model asset market trading explicitly under cash-in-the-market constraints.

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

[8] Alternatively, borrowers could offer only fixed-term two-period contracts and rely on trading between investors with different liquidity needs. Prima facie, this would eliminate the fragility from short-term funding in the borrowing contracts. But one can show that under the alternative scenario the interest rate structure on the resulting short-term borrowing market would be identical to the structure that we derive in the following section. Therefore, the fragility that we analyze in this paper would arise identically in the short-term borrowing market. For simplicity, we therefore restrict borrowing contracts directly to the contingent form described above.

sum of discounted expected cash flows $\sum_{\tau=t}^{\infty}\beta^{\tau-t}\pi_i^{\tau}$, where $\beta<1$.[9] In order to make the problem interesting, we assume that borrowers are sufficiently patient and their long-term investment is sufficiently profitable:

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

By (5), borrowers will invest any funds they own up to their capacity rather than consume them. Since we only consider steady-states and do not model the build-up to the steady states, we therefore simply assume that all borrowers are active in the sense that $I_i^t > 0$ for all t.[10]

Given the corner preferences of investors in (1), there is no scope for rescheduling the financing from impatient middle-aged or old investors. Hence, if $\pi_i^t < 0$ at any date t, the borrower is bankrupt. Of course, his assets may be taken over by other borrowers to realize the borrower's going concern value, as in the case of Bear Sterns, but from the point of view of the individual borrower the relevant bankruptcy condition is $\pi_i^t < 0$.

We assume that the Law of Large Numbers also holds at the borrower level: each period the realized fraction of impatient investors at each borrower is α. Hence, in every period, a borrower who has obtained funds from young investors repays a fraction α of middle-aged investors and all remaining old investors. Thus there is no uncertainty about borrowers' profits, and each borrower's realized profit is equal to his expected profit (4).

Each period, borrowers compete for investors' funds. Since borrowers have a limited 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. At each date $t \in (-\infty, \infty)$:

1. borrowers offer contracts $(r_{1i}^t, r_{2i}^t, Q_i^t, k_i^t) \in \mathbb{R}_+^4$, $i = 1, ..., M$.

2. New and patient middle-aged investors decide whether to finance the borrower.

[9] The discount factor of borrowers, β, can be interpreted in a number of ways. The standard interpretation is that it represents a "psychological" preference for present consumption on the part of the owners of the borrower. Another is that it corresponds to the (inverse of) the return of some unmodeled alternative investment.

[10] This will be true, for example, if borrowers have a small initial endowment in one period or build up their investment through an appropriate dynamic borrowing strategy (which we do not model here).

3. If the borrower is unable to repay all investors who demand repayment, he must declare bankruptcy. Otherwise, the borrower invests I_i^t and continues.

Here, $r_{\tau i}^t$ is the (gross) interest payment offered by borrower i on τ-period funding, Q_i^t the maximum amount for which this offer is valid, and k_i^t is the amount of collateral posted per unit borrowed. Total new borrowing by the M borrowers then is $(b_1^t, ..., b_M^t) \in \mathbb{R}_+^M$, with $b_i^t \leq Q_i^t$ for $i = 1, ..., M$ and $\sum b_i^t \leq N$.

Since investment returns are non-verifiable, the transaction must be collateralized such that the borrower honors the repurchase leg of the repo transaction. At the time t of the contract offers, the borrower wants to roll over $(1-\alpha)r_{1i}^{t-1}b_i^{t-1}$ in loans from patient middle-aged investors and attract b_i^t from young investors. The borrower can offer collateral of two kinds: I_i^{t-1} units of assets created last period and maturing in $t+1$ ("seasoned collateral"), and I_i^t units of assets created by this period's investment ("fresh collateral"). To simplify the presentation, we assume that borrowers collateralize all their loans by an equally weighted mix of the two types of collateral: if a loan of size 1 is collateralized at the rate k_i, then the collateral is composed of $\frac{1}{2}k_i$ units of seasoned collateral and fresh collateral each.[11]

Hence, if the loan has the gross interest rate r, then the borrower will repay the loan (i.e. redeem all his collateral) next period instead of keeping his cash, if the value next period of all the collateral to the borrower is greater than the total repayment promise:

$$\frac{1}{2}k_i R_i + \frac{1}{2}k_i \beta R_i \geq r \tag{6}$$

$$\Leftrightarrow \quad \frac{1}{2}(1+\beta)R_i k_i \geq r \tag{7}$$

[11] Note that in steady state both types of collateral are available in equal quantity (I_i). The assumption is a simplification because loans from young and middle-aged investors can in principle differ in their collateral composition. For example, one can assume that loans from young investors are collateralized exclusively by fresh collateral, and loans from middle-aged investors exclusively by seasoned collateral. The most general formulation would have loans from young investors collateralized by a fraction λ_Y of fresh collateral (and $1-\lambda_Y$ of seasoned collateral) at an overall collateralization rate of k_Y, and analogously loans from middle-aged investors.

To simplify we assume $\lambda_Y = \lambda_M = 0.5$ and $k_Y = k_M = k$. Our whole analysis holds generally, but the exposition becomes more cumbersome.

We will abstract from more complicated considerations of default and ex post bargaining, and simply assume that collateral must satisfy the repayment constraint (7).[12]

2 Steady-state equilibrium

A steady state equilibrium is a collection of $(r_{1i}, r_{2i}, k_i, b_i, I_i, C_i)$ for each borrower i, where b_i is new funding, k_i collateral, C_i cash holding, and $I_i \leq \overline{I_i}$ investment per borrower, such that no borrower and investor would prefer another funding and investment policy, given the behavior of all others.[13]

In the appendix we provide a detailed characterization of the steady states of the model. Here we only discuss the intuition of our results. First, it is easy to see that all equilibrium interest rates must be equal across borrowers. If this were not the case, either investors would choose to finance borrowers paying a higher interest rate or, if these borrowers no longer accept additional funds, then one of the borrowers paying a high interest rate could lower the rate it offers to increase profits.

Second, a no-arbitrage argument for investors implies that the term structure of interest must be flat in equilibrium, i.e. that $r_2 = r_1^2$. The same result has been shown by Qi (1994) under the assumption that borrowers do not discount and make zero profits. We do not need these assumptions. Intuitively, since borrowers cannot distinguish between young and middle-aged investors, they cannot offer $r_2 < r_1^2$. But if the reverse inequality were true, then investors could agree on side trades at the borrowers' expense.[14]

We denote the one-period interest rate by r. Thus, repos offered by borrowers in our model last one period because of the possibility of arbitrage among investors.[15]

[12] 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.

[13] 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.

[14] For these side trades it is important to note that impatient investors do not "die" early (as the banking literature sometimes puts it casually), but that they can borrow in the middle-period of their lives in order to repay when old.

[15] In practice, the tenor of repos was typically very short. For example, FitchRatings

If a borrower has own funds (i.e. if $\pi_i > 0$), then she will invest them up to the capacity constraint, because each unit invested yields $\beta^2 R_i - 1 > 0$. For each unit of cash that the borrower invests at date t on behalf of investors, she expects to pay back αr in $t+1$, generate returns R_i in $t+2$, and pay back $(1-\alpha)r^2$ in $t+2$. Hence, her expected discounted profit on this unit of "external funds" is $\beta^2(R_i - (1-\alpha)r^2) - \beta\alpha r$. In equilibrium, all investor funds must be invested. Hence, the return on borrowed funds cannot be smaller than the return on own funds, if the borrower has funds of her own. Furthermore, competition will drive interest rates up to the point where the two rates of return are equal. This implies the following "dynamic participation constraint":

$$(1-\alpha)\beta^2 r^2 + \alpha\beta r = 1 \tag{8}$$

Basic algebra shows that its solution is

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

The dynamic participation constraint makes borrowers indifferent at the margin between attracting more cash from investors, which increases current borrower consumption, or attracting less and using their own cash to finance investments, which increases future consumption. Consequently, since the consumption value from investing own funds is positive, borrowers make positive profits in equilibrium.

The following proposition characterizes steady equilibria and highlights the role of positive profits.

Proposition 1 *There are steady state equilibria in which*

- *all investors roll over their loans according to their liquidity needs,*
- *all borrowers make strictly positive profits* π_i
- *investment in maximal:* $I_i = \overline{I_i}$*,*

Fund & Asset Manager Rating Group reports in its Money Market Funds U.S.A. Special Report (Oct. 4, 2010) that 81.8 percent of repo allocations by Fitch-Rated taxable money market funds as of August 31, 2010, were overnight. See also Krishnamurthy, Nagel and Orlov (2011).

- *borrowers do not hold cash:* $C_i = 0$*,*
- *all borrowers pay the same interest rate:* $r_i = \overline{r} = 1/\beta$*,*
- *all investor cash is lent:* $\sum_i b_i = N$*, and individual borrowing satisfies*

$$b_i \leq \frac{(1+\beta)}{1-\alpha+\beta} I_i \tag{9}$$

  and is otherwise indeterminate,

- *collateral* k_i *satisfies*

$$\frac{2}{\beta(1+\beta)R_i} \leq k_i \leq \frac{2\beta\overline{I_i}}{(1-\alpha+\beta)b_i} \tag{10}$$

  and is otherwise indeterminate.

In all other steady state equilibria, if they exist, all active borrowers make zero profits.

In the appendix, we provide a fuller characterization of steady state equilibria, of which Proposition 1 is a special case. The interesting equilibria are those with strictly positive profits. As argued above, at the interest rate $\overline{r}$, every borrower i is indifferent at any date t between using outside funds and using her own cash π_i for investment, and thus finds it indeed optimal to borrow any positive amount b_i satisfying (9).

Because $\overline{r} > 1$ and all borrowers 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 (3), there exist borrowing levels $(b_1, ..., b_M)$ satisfying (9) such that $\sum_i b_i = N$. Collateralization is indeterminate in equilibrium, as long as the repurchase constraint (7) is satisfied at the equilibrium rate, which is the first inequality in (10), and as long as the borrower has enough collateral, which is the second inequality in (10). The first inequality therefore is an incentive constraint on the borrower and the second inequality is a feasibility constraint derived from the borrower's equilibrium investment level. The interval in (10) is not empty because of (9).

The steady states described in proposition 1 all feature maximum investment and the same interest rate $\overline{r}$, but borrowers can differ in their reliance on outside funds b_i and the collateral k_i they post. In fact, subject to constraint (10), the exact amount of collateral plays no role in steady state 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 funding level b_i is only restricted by the requirement that the borrower has sufficiently profitable investment opportunities, as expressed by (9). This by itself implies that the borrower's steady state asset base is sufficient to collateralize her funding. Hence, our equilibrium characterization makes no exact prediction about the distribution of debt financing across borrowers. But our multiple equilibria are Pareto-ordered with respect to leverage. By the very definition of equilibrium, borrowers have no incentive to change their borrowing exposure. But interestingly, they may prefer other steady states. To see that, note that borrower equilibrium profits

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

are strictly decreasing in b_i. Hence, if a borrower could choose between a steady state with lower leverage and one with higher leverage, she would prefer the former. The multiplicity of steady states in Proposition 1 therefore is consistent with the notion that borrowers can be "trapped" in an equilibrium with high short-term funding and low profits. This does not matter as long as no shock is realized but, to the extent that period profits can act as a buffer against adverse shocks, as we show in the following sections, borrowers with larger exposure to short-term funding will be more fragile.

The steady states of proposition 1 always exist. In contrast, zero-profit steady states can exist, but need not. In the appendix, we characterize such equilibria. In such zero-profit equilibria, generally only some borrowers are active, borrow more than they can invest ($b_i > I_i$), investment is inefficient ($I_i < \overline{I}_i$), and the interest rate satisfies $r > 1/\beta$. The active borrowers in these equilibria would therefore like to reduce their borrowing, but are forced to keep it up in order to repay previous loans, because they have no other liquidity ($\pi_i = 0$).

The zero-profit equilibria therefore are a knife-edge case. Since they need not exist and are hardly convincing as a description of investment banking, we take the equilibria with $r = \overline{r}$ identified in Proposition 1 as a benchmark for the rest of the analysis. An important and novel feature of these equilibria is that condition (8) prevents competition from driving up interest rates to levels at which borrowers make zero profits. The reason why borrower profits are positive in these equilibria is intuitive (but not trivial): borrowers 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.

Positive profits from intermediation play an important role in our model. Borrowers facing the threat of a run may be able to repay investors that refuse to roll over their loans by using this inside liquidity. Hence, in contrast to other models, not all borrowers are susceptible to a run. In particular, as we show in the next two sections, our model generates testable implications, such as the fact that more leveraged borrowers, or less profitable ones, are more fragile.

3 Runs: Liquidity

In this section, we study the stability of borrowers 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 borrower 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 repo markets 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 borrower can mobilize to meet the repayment demands by middle-aged investors in such a situation. At the beginning of the period, a borrower, on the asset side of his balance

[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 infinite-horizon models.

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 borrower holds investor claims for dates t and $t+1$ on the liability side of his balance sheet. Remember that the borrower cannot sell his assets.[17]

The borrower'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 borrower if

$$(R_i - 1)\overline{I_i} \geq (\overline{r} + (1-\alpha)\overline{r}^2)b_i. \tag{11}$$

If (11) holds, a run would have no consequence whatsoever and all out-of-equilibrium investor demand would be buffered by the borrower'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 borrower who satisfies the run demand is

$$I_0 = R_i\overline{I_i} - (\overline{r} + (1-\alpha)\overline{r}^2)b_i. \tag{12}$$

Clearly, if $I_0 < 0$ the borrower does not have the liquidity to stave off the run and is bankrupt. If $I_0 \geq 0$, but (11) does not hold, the borrower must adjust his funding or investment in order to survive the run. Since after a run in $t+1$ the borrower 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, (13) in the following proposition, is obtained by simply writing out the condition $I_0 \geq 0$ from (12).

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

$$\beta^2 R_i\overline{I_i} \geq (1-\alpha+\beta)b_i. \tag{13}$$

Condition (13) is independent of the funding restriction (9) of Proposition 1, in the sense that (13) can hold or fail in steady state, depending on the

[17]The alternative extreme case where the borrowers' assets can be traded on perfect markets is treated in Martin, Skeie and von Thadden (2014).

parameters and on the borrower's exposure b_i. Hence, a borrower 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 (13) is the tighter,*

- *the higher is the borrower's short-term exposure* $b_i/\overline{I}_i$*,*
- *the smaller is the borrower's size* $\overline{I_i}$*,*
- *the lower is the borrower's profitability* R_i*.*

Proposition 3 shows that if borrowers have sufficient access to profitable investment ($\overline{I_i}$ large), if these investment opportunities are sufficiently profitable (R_i large), or if they have sufficiently low exposure to short-term outside funding ($b_i/\overline{I}_i$ small), then borrowers 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 borrowers out of equilibrium. If condition (13) is violated, a run bankrupts the individual borrower if it occurs.

It is interesting to compare the predictions of Proposition 3 to actual data before the crisis of 2008. Table 1 lists the size, short-term leverage, and profitability of the five big securities dealers in late 2007.[18] Of course, such a descriptive comparison has only limited empirical value. But consistent with the pecking order found in Proposition 3, Bear Stearns and Lehman, the two dealers that went under in 2008 after a precipitous loss of short-term funding, were the weakest of the dealers according to size and short-term leverage. In terms of profitability, the numbers don't fit the prediction of the proposition. But since profitability depends on a number of delicate accounting values,

[18]Size is measured as the quantity of assets. Short-term exposure is measured as short-term liabilities (the sum of short-term unsecured debt and collateralized borrowing, which includes repo financing, securities borrowed) as a fraction of all liabilities. One-year exposure is measured as the sum of short-term liabilities and debt in current liabilities (total amount of short-term notes; and the current portion of long-term debt, which is long-term debt due in one year) as a fraction of all liabilities. Profitability is measured as annualized total return on equity over the first two months of 2008. Source: SEC quarterly filings and the Buckingham Research Group estimates.

in particular the value of equity, the significance of this discrepancy is not clear.

Securities Dealer	Size ($billion)	Short-term exposure (%)	One-year exposure (%)	Profitability (%)
Goldman Sachs	1,120	29.0	71.7	-11.1
Morgan Stanley	1,045	35.0	69.7	-10.7
Merrill Lynch	1,020	35.8	79.8	-3.6
Lehman Brothers	691	40.4	95.8	-11.6
Bear Stears	395	35.2	80.7	-4.7

Table 1. Size, Exposure and Profitability by Dealer in 2007 Q4

4 Fragility

This section examines the microstructures of the tri-party and the bilateral segments of the U.S. repo market. We ask whether runs can occur in each of the institutional environments considered. 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 segment. In the appendix, we show how our analysis can be extended to other institutional settings, such as MMFs, ABCP conduits, and traditional bank deposits.

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 borrower i, so the borrower 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 borrower. The question is whether such beliefs can be self-fulfilling in a collective deviation from the steady state.

Our model allows us to contrast the fragility of the tri-party and the bilateral repo market. We show that for some parameter values, a repo market in which haircuts cannot move, as was the case in the tri-party market during the crisis, is more fragile than a market where haircuts can move, as was the case in the bilateral market. This is consistent with the fact that some dealers relying mainly on the tri-party market for their repo funding, such as Bear and Lehman, experienced dramatic declines in their repo funding similar to runs.

In section 4.1, we show that runs can occur even when there is no "first-come-first-serve" constraint, because of the intraday extension of credit from the clearing banks to the borrower called the "unwind". Eliminating the unwind removes the possibility of runs, as we show in section 4.2. Finally, we study the bilateral market, which has a "first-come-first-serve" constraint, in section 4.3. The bilateral market can suffer runs, as does the tri-party market with unwind. The key difference between the two markets is that haircuts increased considerably in the bilateral market, while they barely moved in the tri-party market. Our results show that flexible haircuts make the bilateral market less fragile.

Since the Law of One Price holds in the steady state of our model, a trivial coordination failure may induce all investors of a given borrower to switch to another borrower out of indifference. This looks like a "run", but is completely arbitrary. We will therefore assume that investors, if indifferent, lend to the borrower 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 borrower is liquid in the sense of Proposition 2.

Lemma 1 *If a borrower satisfies the liquidity constraint (13), there are no strict incentives to run on this borrower.*

The proof is simple. In a run on this borrower, all middle-aged patient investors would be repaid in full regardless of what young investors do and without affecting the borrower's asset position. Hence, patient middle-aged and young investors are indifferent between lending to the borrower 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 borrower has the money, such a check does not cause any real disruption, and the borrower may as well keep it until he invests in new securities.

4.1 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.[19] In particular, we show that a practice called the "unwind" of repos increases fragility in this market.

The clearing banks play several roles in the tri-party repo market. They take custody of collateral, which guarantees that a cash investor can have access to the collateral in case of a borrower 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 borrowers.[20]

In the US tri-party repo market, new repos are organized each morning, typically 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 borrowers 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 borrowers' to the investors' account and the securities from the investors' to the borrowers' account. They also finance the borrowers' 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.[21]

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 borrower only intraday. At each date, the clearing bank finances borrowers 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 borrower i this works as follows:

[19] 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.

[20] The reform proposed by the Task Force limits considerably the ability of the clearing banks to extend intraday credit once it is implemented (for more detail, see Task Force 2010 and the NY Fed's tri-party repo reform web page).

[21] Copeland, Duffie, Martin, and McLaughlin (2012) discuss the unwind process in more detail.

(a) The clearing banks sends the cash amount $b_i\left[\overline{r}+(1-\alpha)\overline{r}^2\right]$ to all investors of borrower i, extinguishing the investors' exposure to the borrower they have invested in.

(b) At the same time, the clearing bank takes possession of the assets the borrower has pledged as collateral.

(c) In the process, the clearing bank finances the borrowers temporarily, holding the assets as collateral for its loan.

2. $\bar{I}_i$ assets of a borrower mature (yielding $R_i\overline{I_i}$ in cash), allowing the borrower to repay some of its debt to the clearing bank.

3. Possibly a sunspot occurs.

4. The borrower 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 borrower.

6. If the borrower is unable to repay its debt to the clearing bank, he must declare bankruptcy. Otherwise, the borrower 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 collectively optimal for them.[22] For simplicity, we assume that the clearing bank extends the intraday loan to the borrower at a zero net interest rate. Also, since runs are zero probability events the clearing banks has no reason not to unwind repos.[23]

It is important to note that changing haircuts is operationally cumbersome in the tri-party repo market. In contrast to the bilateral market, haircuts in the tri-party market are not negotiated at the trade level but, instead, are established in the appendix of legal documents governing the arrangement between a clearing bank and each borrower and lender pair. One reason is

[22] 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 $(\overline{r}, Q_i, b_i, k_i)$).

[23] In the appendix, we consider the coordination problem between the clearing bank and the investors. This problem is also studied in Ennis (2011).

that the haircuts are an input to the collateral allocation tools that the clearing banks put at the disposal of the dealers. Changing haircuts is possible, of course, but might well require 24 or 48 hours, according to market participants. In practice, these changes 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.[24]

In contrast to other models of run, our time line implies that there is no "first-come-first-serve" (FCFS) constraint in the tri-party repo market. In step 1(a) the clearing bank gives all investors their cash back. A run occurs if middle-aged patient and young investors refuse to reinvest, or invest, respectively. The reason a run can occur in spite of the absence of a FCFS constraint is that the borrower can stay in business only if sufficiently many investors provide funding. There is, thus, a coordination problem among investors because investment has a high payoff if sufficiently many invest and a low payoff otherwise.

In response to the contract offer by the borrower, individual investors must compare their payoff from investing with the borrower in question to that from investing with another borrower. The latter decision yields the common market return $\overline{r}$,[25] the return from the former depends on what the other investors do. Table 2 shows the payoffs of the two decisions for the individual investor (rows) as a function of what the other investors do (columns), if the borrower is potentially illiquid (i.e. if the liquidity constraint (13) is violated). If the investor re-invests her funds with the borrower, the clearing bank will accept the cash, since it reduces its intraday exposure to the borrower, 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 borrower 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 $\gamma_i^t R_i k_i$ at date $t + 1$ if the borrower defaults.

[24] 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.

[25] 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	$\widehat{r}_i$	$\gamma_i^t R_i k_i$
don't	$\overline{r}$	$\overline{r}$

Table 2: Payoffs in tri-party repo with unwind

Hence, investors will finance the borrower in case of a run iff[26]

$$\overline{r} \leq \gamma_i^t R_i k_i \tag{14}$$

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 (14) does not hold, the collective decision not to lend to the borrower 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 borrower i would also choose not to invest. In our model, steady state collateral is not unique, but clearly, if constraint (14) is violated for the maximum possible amount of collateral consistent with (10) (denoted κ_i in (34)), then it cannot hold in any case.

Combining the above results with those of the previous section and writing out condition (14) for κ_i 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 borrower i can occur and bankrupt the borrower if and only if the borrower's liquidity constraint (13) and his collateral constraint*

$$\beta^2 R_i \overline{I_i} \geq \frac{1-\alpha+\beta}{\gamma_i^t(1+\beta)} b_i \tag{15}$$

are both violated.

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

Condition (15) is implied by the steady-state borrowing constraint (9) 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 borrowers ("good" collateral in "normal" times), the collateral constraint is slack, and the borrower is run-proof. The collateral constraint becomes relevant only when the perceived value of collateral falls such that there are larger differences in valuation between investors and borrowers.

Furthermore, condition (15) is independent of the liquidity constraint (13). 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 (15) is the tighter,*

- *the lower is the value γ_i^t of collateral to investors*
- *the higher is the borrower's short-term leverage $b_i/\overline{I}_i$,*
- *the smaller is the borrower's size $\overline{I_i}$,*
- *the lower is the borrower'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 (13) of Proposition 1 and (15). Both constraints are relaxed if borrowers have sufficient access to profitable investment ($\overline{I_i}$ large), if these investment opportunities 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 borrowers if they occur out of equilibrium. In the opposite case, a run can be a self-fulfilling prophecy and bankrupt the borrower.

4.2 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. This case is similar to the tri-party repo markets in Europe. It is also similar to what the US tri-party repo market will become once the recommendation of the Task Force are implemented.[27]

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

1. Possibly a sunspot occurs.
2. The borrower 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 borrower.
4. If the borrower is unable to repay his debt to last period's repo investors, he must declare bankruptcy. Otherwise, the borrower continues.

One might think that removing the unwind would lead to a FCFS constraint, since the clearing bank no longer returns cash to all investors in the morning. This does not occur because repos settle simultaneously. An investor who does not roll over his repo is repaid $\overline{r}$ if and only if the borrower can repay all investors who request their cash back - otherwise the borrower is bankrupt and repays investors an amount smaller than the contractual payment.

In contrast to the case with unwind, young and middle-aged investors are in a different situation: Young investors hold cash while middle-aged investors hold a repo with the borrower, until the borrower is able to repay. Hence, in case of bankruptcy, all middle-aged investors keep their collateral.[28] It follows that, in case of bankruptcy, the payoff of middle-aged investors is independent of whether an individual investor has demanded to be repaid or has agreed to roll over his loan. If the borrower does not go bankrupt, a patient middle-aged investor can get her cash back, but in this case she would prefer to have invested. Hence, patient middle-aged investors weakly prefer to reinvest. 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 borrower.

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

[28] Depending on the bankruptcy rules they may also obtain additional cash as unsecured creditors for the difference between the market value of their collateral and the face value of their repo

From Lemma 1 it is again enough to consider the case in which the borrower is illiquid after a run.

Lemma 2 *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 borrower is liquid, because

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

by (9). The borrower therefore has enough assets that will mature in the future to satisfy all future claims by young agents who invest today. ∎

As noted above, 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 borrowers.*

4.3 Bilateral repos

In this section, we apply our model to bilateral repos. 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.[29] 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 that funds "Fedwire-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

[29] Also, a dealer 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 investor liquidity needs, this cannot happen in our model. He and Xiong (2012) analyze the consequences of (exogenously determined) staggered short-term debt for the stability of financial institutions.

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 FCFS constraint. In the case of a run, investors who send the securities they hold as collateral early are more likely to receive cash than investors who send their securities late. This contrasts with the tri-party market, both with and without unwind, were the FCFS constraint was absent.

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 investors decide whether to engage in new repos with a borrower.
4. Patient middle-aged investors are repaid in the order in which they send back their collateral, until the borrower runs out of cash. From that point on, investors receive their collateral and any investor 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 them to increase the collateral per unit borrowed. Other than the FCFS constraint, the ability to change margins is the key difference with the tri-party market. In order to withstand the run, the borrower must at least cover the missing amount

$$m_i \equiv (\overline{r} + (1-\alpha)\overline{r}^2)b_i - R_i\overline{I_i} \ , \tag{16}$$

which is the difference between its promised payout to investors and revenues. 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

$$\overline{k}_i = \overline{I_i}/m_i. \tag{17}$$

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

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

Table 3: Payoffs to young investors in bilateral repos

Table 3 gives the payoff to an individual young investor as a function of the collective behavior of all other investors. The payoffs are as in Table 2, 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

$$\gamma_i^t R_i \widehat{k}_i \geq \overline{r} \tag{18}$$

Now, if the funding shortfall m_i is small, the borrower can increase his collateralization beyond the bound κ_i in (34), and this condition is weaker than (14) in the tri-party context.

Note that the borrower can attract as many young investors as necessary to fund the shortfall m_i if he has the collateral, because he can compete away investors from other borrowers if his offer is sufficiently attractive. Inserting m_i from (16) into (17) yields the collateral constraint of the following proposition.

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

$$\beta^2 R_i \overline{I_i} \geq \frac{1 - \alpha + \beta}{1 + \gamma_i^t \beta} b_i \tag{19}$$

is violated.

Proof. Condition (19) is (18) 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 investors will fund the shortfall if it holds. In order to prove necessity, we must examine the incentives of middle-aged patient investors to roll over their existing repos.

Suppose therefore that condition (19) is violated. From (16), only a fraction

$$\varphi \equiv \frac{R_i \bar{I}_i}{b_i \left[\overline{r} + (1-\alpha)\overline{r}^2\right]} \in (0,1) \tag{20}$$

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

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

Table 4: Payoffs to middle-aged patient investors in bilateral repos

Table 4 differs from Table 3 in the lower right cell, which reflects the different positions of young and patient middle-aged investors. The outcome (don't, don't) is strictly optimal for the individual patient middle-aged investor if and only if

$$\gamma_i^t R_i \widehat{k}_i < \varphi\overline{r} + (1-\varphi)\gamma_i^t R_i k_i \tag{21}$$

This condition holds for all k_i and $\widehat{k}_i$ iff it holds for $\widehat{k}_i = \overline{k}_i$ from (17) and $k_i = 1/\beta^2 R_i$ from (10). Re-writing (21) for these two extreme values and setting $d_i = \overline{I}_i/b_i$ yields

$$\gamma_i^t R_i \frac{d_i}{\overline{r} + (1-\alpha)\overline{r}^2 - R_i d_i} < \varphi\overline{r} + (1-\varphi)\gamma_i^t R_i \frac{1}{\beta^2 R_i} \tag{22}$$

$$\Leftrightarrow \qquad \frac{\gamma_i^t \beta^4 d_i R_i}{1-\alpha+\beta-\beta^2 d_i R_i} < \frac{\beta^3 d_i R_i}{1-\alpha+\beta} + \gamma_i^t \frac{1-\alpha+\beta-\beta^2 d_i R_i}{1-\alpha+\beta} \tag{23}$$

Since (19) is violated, we have

$$1-\alpha+\beta-\beta^2 d_i R_i > \gamma_i^t \beta^3 d_i R_i \tag{24}$$

Hence, (23) is equivalent to

$$(1-\alpha+\beta)\gamma_i^t\beta^4 d_i R_i < \left[\beta^3 d_i R_i + \gamma_i^t(1-\alpha+\beta-\beta^2 d_i R_i)\right](1-\alpha+\beta-\beta^2 d_i R_i) \tag{25}$$

Suppose first that $\gamma_i^t > \beta$. By (24), it is enough to show that

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

which is implied by (24).

Now suppose that $\gamma_i^t \leq \beta$. (25) is linear in γ_i and holds for $\gamma_i^t = 0$ and for $\gamma_i^t = \beta$. Hence, it holds for all $\gamma_i^t \leq \beta$. ■

As condition (15) in the tri-party case, condition (19) is implied by the steady-state borrowing restriction (9) 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. If firms take this into account strategically, this reduces the value of repo financing in times of crisis.

This is consistent with the findings of Auh and Sundaresan (2013) who test a model of optimal liability structure with long-term unsecured debt and short-term super-senior secured debt that is subject to runs. They consider the onset of the financial crisis in 2007 as an exogenous shock that decreased the value and increased the riskiness of collateral and find that leverage and short-term debt decreased, and fell much more rapidly for financial firms than non-financial firms due to the greater exposure of financial firms to runs.

Furthermore, and differently from the tri-party case, condition (19) is strictly weaker than the liquidity constraint (13). Hence, if it is violated, (13) is violated as well. This means that (19) is necessary and sufficient for the stability of bilateral repos.

Finally, the bilateral collateral constraint is strictly weaker than the tri-party constraint (15). 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 – a direct consequence of the "liquidity buffer" provided to bilateral repo borrowers through the option to collect additional funding liquidity in the face of a liquidity shock. In particular, in contrast to the tri-party repo market, in the bilateral market the condition $k_i = \widehat{k_i}$ need not hold. 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 those in the bilateral market, hence have a lower γ_i^t.[30]

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

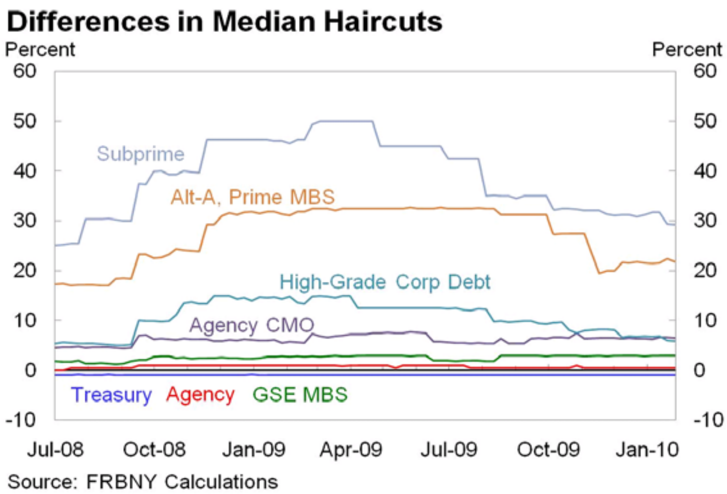

Figure 1: Differences in median haircut between bilateral and tri-party repos per asset class

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 (2012) of sharply rising haircuts during the crisis of 2008.[31] However, the behavior of haircuts was very different in the tri-party and bilateral repo markets. Figure 1 provides some graphical evidence of 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.

In addition, Lehman, experienced a sudden decrease in the size of its tri-party repo funding shortly before its downfall, with hardly any adjustment in haircuts. Anecdotal evidence suggests that Bear Stearns had a similar experience. We are not aware of similar sudden losses of funding in the bilateral repo market. Instead, institutions in this market saw more gradual increase in haircuts that reduced the amount of funding they could obtain (Gorton and Metrick, 2012). Our results in Sections 4.1 and 4.3 are consistent

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

with these two different developments in the bilateral and tri-party repo markets. As such, our model can account for the differences in the fragility of different segments of the repo market that were revealed by the financial crisis of 2007-2009.

5 Conclusion

We have developed an infinite-horizon equilibrium model to study how the fragility of short-term funding markets depends upon the particular microstructures, liquidity, and collateral arrangements that may lead to runs at various types of financial institutions. The value of collateral and the endogenous liquidity of intermediaries then become crucial for their stability. Runs can be forestalled by mobilizing sufficient liquidity and having sufficiently valuable collateral. 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".

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 infinite-horizon models such as Kiyotaki and Moore (1997), Brunnermeier and Pedersen (2009) or Uhlig (2010). On a behavioral note, this approach is consistent with Gennaioli, Shleifer, and Vishny's (2012) "neglected risk" model.

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 in the crisis of 2008. 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. The rigidity of haircuts and the "unwind" in the tri-party market are an example of an institution that is useful in normal times (because it simplifies frequent transactions), but potentially destabilizing in stress times.

Our framework can be used to consider a number of policy questions related to the fragility of short-term funding markets. In its 2010 White Paper, the Federal Reserve Bank of New York highlighted three weaknesses of the tri-party repo market: 1) The market's reliance on intraday credit from the clearing banks, as a consequence of the unwind, 2) Inadequate risk management practices, and 3) The lack of plans to support orderly liquidation of a defaulted dealer's collateral. Our model provides relevant insights for the first and the third items.

As already noted, our model suggests that the current reform efforts aiming to eliminate the unwind in the tri-party repo market are likely to enhance stability in that market. Another important aspect of reform is related to the risk of fire sales in the repo market. Lehman's demise highlighted the risks associated with the lack of plans to support orderly liquidation of a defaulted dealer's collateral, although these risks did not fully materialize in part because of exceptional measures taken by the public sector. 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 liquidate the positions of a defaulting borrower, or other similar mechanisms. Along similar lines, the model can be enhanced to consider the risk of fire sales. Martin, Skeie, and von Thadden (2012) use the model of this paper to analyze asset sales as a way for a dealer to meet some of its liquidity needs in the face of a run.

6 Appendix:

6.1 Characterization of Steady State Equilibria

Lemma 1: *In steady-state borrowers do not hold cash:* $C_i = 0$ *for all* i.
Proof. Since $\beta < 1$, and $C_i > 0$ does not affect the borrower's budget constraint, each borrower does strictly better by consuming C_i. ■

Lemma 2: *For each borrower* 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 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 borrower one period later and repay $q\widehat{r}$. This deviation is feasible if (i) $q\widehat{r} = r_{2i}$. It makes both parties weakly better off iff (ii) $q \geq r_{1i}$, (iii) $\widehat{r} \geq r_{1i}$, and (iv) $\widehat{r}r_{1i} \geq r_{2i}$. (i), (ii), and (iv) imply $\widehat{r}r_{1i} = r_{2i}$. Hence, (iii) holds with strict inequality, and there are gains from trade.[32] ■

Lemma 3: $r_{1i} = r_{1j}$ *for all borrowers* 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 borrowers 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 borrower $k \in \mathcal{J}_i$ can deviate to $r_{1k} - \varepsilon$ for $0 < \varepsilon < r_{1k} - r_{1i}$ and strictly increase his profit. ■

Hence, there is a single interest rate $r = r_1 = \sqrt{r_2}$ in steady state and no precautionary cash holding. Let $(r, b_1, ..., b_M, I_1, ..., I_M, k_1, ..., k_M)$ be a steady state equilibrium.

Define $\widetilde{r}_i > 1$ by

$$\beta(1-\alpha)r^2 + \alpha r - \beta R_i = 0 \tag{26}$$

[32] This lemma is based on a no-arbitrage argument and does not assume concavity of the utility function. Alternatively, the lemma can be proved in an environment in which investors cannot trade with each other, if utility functions are concave with relative risk-aversion greater than 1.

$\widetilde{r}_i$ is the break-even rate from investing one unit of borrowed funds for borrower i. By construction, for all $i = 1, ..., M$, $\widetilde{r}_i > \overline{r} = 1/\beta$ as defined by (8).

Lemma 4: $r \leq \widetilde{r}_i$.
Proof. Take any i with $b_i > 0$ and suppose that $r > \widetilde{r}_i$. Deviation: reduce borrowing and investment by a marginal \$. This yields an expected gain of

$$\beta(1-\alpha)r^2 + \alpha r - \beta R_i > 0$$

by the definition of $\widetilde{r}_i$. Hence, $r \leq \widetilde{r}_i$. ■

Lemma 5: *If* $\pi_i = 0$, *then* $b_i > I_i$.
Proof. If $b_i = 0$ then trivially $\pi_i > 0$ by (4). Hence, $b_i > 0$.
The definition of $\widetilde{r}_i$ in (26) and Lemma 4 imply

$$\beta(R_i - (1-\alpha)r^2) - \alpha r \geq 0$$

Suppose that $b_i \leq I_i$. Then

$$\begin{aligned} \pi_i &= (R_i - 1)I_i - ((1-\alpha)r^2 + \alpha r - 1)b_i \\ &\geq (R_i - (1-\alpha)r^2 - \alpha r)I_i \\ &> 0 \end{aligned}$$

because $\beta < 1$. ■

Lemma 6: *If* $r < \overline{r}$ *then* $\pi_i = 0$ *for all* $i = 1, ..., M$.
Proof. Suppose that $\pi_i > 0$ for one i. Deviation: Borrow a sufficiently small amount X at (gross) interest $r + \varepsilon < \overline{r}$ and consume it (repayment will come out of future π_i's). This yields in expectation

$$(1 - \beta\alpha(r+\varepsilon) - \beta^2(1-\alpha)(r+\varepsilon)^2)X > 0,$$

by the definition of $\overline{r}$. ■

Lemma 7: *If* $r > \overline{r}$ *and* $\pi_i > 0$, *then* $b_i = 0$.
Proof. Suppose that $b_i > 0$. Deviation: Reduce borrowing by a marginal \$ without changing investment (by reducing π_i^t), which has an expected gain of

$$\beta\alpha r + \beta^2(1-\alpha)r^2 - 1 > 0$$

by the definition of $\overline{r}$. ■

Lemma 8: *If $r > \overline{r}$, then there exists a set $\mathcal{M} \subset \{1, ..., M\}$, $\mathcal{M} \neq \varnothing$, such that*

- $b_i = 0$ *for all* $i \notin \mathcal{M}$,
- $\pi_i = 0$ *and* $b_i = \frac{(R_i-1)I_i}{\sum_{j\in\mathcal{M}}(R_j-1)I_j}N$ *for all* $i \in \mathcal{M}$.

Proof. By Lemma 7 and market clearing, not all i can have $\pi_i > 0$. Let $\mathcal{M} := \{i : \pi_i = 0\}$. $\pi_i = 0$ means

$$
\begin{aligned}
(R_i - 1)I_i - ((1-\alpha)r^2 + \alpha r - 1)b_i &= 0 && (27)\\
\Leftrightarrow \qquad b_i &= \frac{(R_i - 1)I_i}{(1-\alpha)r^2 + \alpha r - 1} && (28)
\end{aligned}
$$

(28) and market clearing imply that

$$
\begin{aligned}
\sum_{j\in\mathcal{M}} b_j &= N \\
&\Leftrightarrow \sum_{j\in\mathcal{M}}(R_j - 1)I_j = N((1-\alpha)r^2 + \alpha r - 1) && (29)\\
&\Leftrightarrow (1-\alpha)r^2 + \alpha r = A_{\mathcal{M}} && (30)
\end{aligned}
$$

where

$$A_{\mathcal{M}} := 1 + \frac{1}{N}\sum_{j\in\mathcal{M}}(R_j - 1)I_j \qquad (31)$$

(28) and (30) now imply

$$b_i = \frac{(R_i - 1)I_i}{\sum_{j\in\mathcal{M}}(R_j - 1)I_j}N \qquad (32)$$

for all $i \in \mathcal{M}$, as claimed in the lemma. ■

Lemma 9: *Suppose that $r > \overline{r}$. Then for all $i \in \mathcal{M}$,*

$$R_i > A_{\mathcal{M}} > 1 + \frac{1}{\beta^2}(1-\beta)(1+\beta-\alpha) \qquad (33)$$

Proof. The first inequality follows from (32) and Lemma 5. The second follows, after some simple algebra, from solving (30) for r and from $r > \overline{r}$. ■

Remark: The upper bound on $A_{\mathcal{M}}$ in Lemma 9 is simple to obtain, but not tight. Lemma 4 yields the stricter condition

$$\beta^2(1-\alpha)(R_i - A_{\mathcal{M}})^2 + \beta\alpha^2(1-\beta)(R_i - A_{\mathcal{M}}) - \alpha^2(1-\beta)^2 A_{\mathcal{M}} \geq 0$$

However, condition (33) is sufficient to construct examples of non-existence.

Lemma 10: *If* $\pi_i > 0$*, then* $I_i = \overline{I_i}$*.*
Proof. Suppose the lemma is wrong. The borrower can then increase investment at any date t by using his own cash. By condition (5), this yields a strict increase in discounted profits. ■

Lemma 11: If there exists a borrower i with $\pi_i > 0$ and $b_i > 0$ then $r = \overline{r}$.
Proof. For each unit of cash that borrower 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 borrower receives funds from investors in steady state ($b_i > 0$) and has funds of his own ($\pi_i > 0$), this cannot be strictly better. Hence, $(1-\alpha)\beta^2 r^2 + \alpha\beta r \leq 1$, which means $r \leq \overline{r}$. The strict inequality is impossible because of Lemma 6. ■

6.2 Proof of Proposition 1

As argued in the main text, at the interest rate $\overline{r}$, every borrower 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 (9). (9) also implies that borrower profits

$$\begin{aligned} \pi_i &= (R_i - 1)\overline{I}_i + (1 - \alpha\overline{r} - (1-\alpha)\overline{r}^2)b_i \\ &= \frac{1}{\beta^2}\left[(\beta^2 R_i - 1)\overline{I}_i + (1-\beta)((1+\beta)\overline{I}_i - (1+\beta-\alpha)b_i)\right] \end{aligned}$$

are strictly positive.

Because $\overline{r} > 1$ and all borrowers 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 (3), there exist borrowing levels $(b_1, ..., b_M)$ satisfying (9) such that $\sum_i b_i = N$. For $r = \overline{r}$, the repayment condition (7).is equivalent to the first inequality in (10). For the second inequality in (10), remember that in steady state the borrower has $\overline{I}_i$ units of fresh and $\overline{I}_i$ units of seasoned collateral to offer. This total of $2\overline{I}_i$ units of collateral can be pledged for the total amount of funds provided by investors per period, which is $b_i \left[1 + (1-\alpha)\overline{r}\right] = b_i \left[1 - \alpha + \beta\right]/\beta$. It follows that the maximum amount of collateral per unit that the borrower can offer is

$$\kappa_i \equiv \frac{2\beta \bar{I}_i}{b_i \left[1 - \alpha + \beta\right]}. \tag{34}$$

The second inequality in (10) is the condition $k_i \leq \kappa_i$. Both inequalities in (10) are compatible because of (9). ■

6.3 Fragility: 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 4.1, just before step 1 the clearing bank comes to believe that at step 5 all investors will refuse to engage in repos with borrower i. In this case, the clearing bank will refuse to unwind if the loan it makes to the borrower, $b_i \left[\bar{r} + (1-\alpha)\bar{r}^2\right]$, exceeds the proceeds it could obtain from the assets, $R\bar{I}(1 + \beta\gamma)$.[33] This condition can be written as

$$\beta^2 R_i \bar{I}_i \geq \frac{1 + \beta - \alpha}{1 + \gamma\beta} b_i. \tag{35}$$

This condition is the same as the collateral condition for bilateral repos, (19).

The flip side of this coordination problem is that investors may choose not to invest with borrower i if they believe that the clearing bank will refuse to unwind that borrower's repos the next morning.[34] 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.

[33] Here we assume that the clearing bank faces the same γ as the investors.

[34] 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.

6.4 Extension: Money market mutual funds

In this extension, we adapt our model to the case of money market mutual funds (MMFs) 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. MMFs 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 MMFs similar to banks since, under most circumstances, investors can obtain their funds on demand at a fixed price. However, MMFs do not hold a capital buffer and do not have access to the discount window. MMFs 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, MMFs 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.[35] In contrast to repo investors, MMF investors do not have a claim on a specific piece of collateral.

In our framework, a MMF 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 MMFs do not invest capital of their own, the argument used to establish the dynamic participation constraint (8) cannot be applied in this context. However, this characterization ignores the important role played by MMFs' parent institutions. A MMF 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 MMF experiences difficulties. Support by parent institutions has been an important source of stability for MMFs 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 the proof of Proposition 1 shows that the MMF'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.

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

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

1. Possibly a sunspot occurs.
2. The MMF offers a new contract $(\hat{r}_i, \hat{Q}_i)$.
3. New and patient middle-aged investors decide whether to withdraw from the MMF.
4. The parent institution decides whether to inject liquidity into the MMF.
5. Investors who redeem their shares get cash until the MMF runs out. At that time, the MMF 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.[36]

Compared to our lead example of Section 3, the liquidity of MMFs therefore differs in two respects. First, MMFs do not invest beyond the level of their short-term funding b_i. This reduces their liquidity and thus tightens their liquidity constraint (13). Second, however, MMFs 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

[36] 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).

analyze the run scenario, we therefore assume that the liquidity constraint is violated and that the parent does not inject liquidity.[37] 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 (20), 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

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

Note that the denominator is again $m_i \equiv (\bar{r}+(1-\alpha)\bar{r}^2-R_i)I_i$. 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	$\widehat{r}_i$	$\gamma_i^t R_i \mu_i$
don't	$\bar{r}$	$\varphi\bar{r}+(1-\varphi)\gamma_i^t R_i \mu_i$

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

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

$$\beta^2 R_i \geq \frac{1+\beta-\alpha}{1+\gamma_i^t\beta}. \tag{36}$$

Interestingly, this condition is the same as (19), evaluated at $b_i = \overline{I}_i$. Note, however, that condition (36) is independent of fund size (which is

[37] 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.

equal to outside funding). This is consistent with the observation that the crisis of MMFs in the wake of the Lehman bankruptcy hit funds across the board, regardless of their size. Again, if (36) is violated, so is the liquidity constraint. Hence, if (36) is violated, the survival of the fund depends on whether the parent company has the cash m_i necessary to stabilize the fund. As (16) shows, this cash shortfall depends on the size of the fund.

6.5 Extension: Asset-backed commercial paper conduits

In this extension we briefly describe the structure of ABCP conduits that Acharya, Schnabl, and Suarez (2013) 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 (forthcoming) focus on the turmoil of the ABCP market in the second half of 2007, which marked the onset of the Great Financial Crisis. As shown in their paper, ABCP had 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 (forthcoming) argue that the precipitous drop in outstanding ABCP of roughly $190 billion in August 2007 had many characteristics of a traditional run.

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 (forthcoming) 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 MMFs.

We do not provide a formal model of ABCP conduits, which would be similar to that of MMFs sketched previously, and only report the findings of Covitz, Liang, and Suarez (forthcoming) 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).

6.6 Extension: Traditional banks

The investors in traditional banks, depositors, are different from the money market participants whom we have considered up to now. But structurally, the analysis for traditional banks is similar to the analysis for MMFs. 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 MMF investors, bank depositors do not get a claim to a specific piece of collateral, but rather a claim on the bank's assets in case of bankruptcy. The major difference between a MMF 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. The bank offers a new deposit contract (r_i, Q_i).
2. New and patient middle-aged investors decide whether to deposit (again) with the bank.
3. 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 MMF, 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 (19). 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 (13) 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.

References

[1] Acharya, V., D. Gale, T. Yorulmazer (2011). "Rollover Risk and Market Freezes." *Journal of Finance* 66, 1175-1207.

[2] Acharya, V., P. Schnabl, and G. Suarez (2013). "Securitization without Risk Transfer." *Journal of Financial Economics* 107, 515-536.

[3] Auh, J. K. and S. Sundaresan (2013). "Bankruptcy Code, Optimal Liability Structure and Secured Short-Term Debt." Working paper.

[4] 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)

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

[6] Bolton, P., Santos, T., Scheinkman, J. (2011). Outside and Inside Liquidity. *Quarterly Journal of Economics* 126-1, 259-321.

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

[8] 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* no. 477.

[9] Copeland A., I. Davis, E. LeSueur, and A. Martin (2012) "Mapping and Sizing the U.S. Repo Market," *Liberty Street Economics*, June 25. (http://libertystreeteconomics.newyorkfed.org/2012/06/mapping-and-sizing-the-us-repo-market.html).

[10] Copeland A., D. Duffie, A. Martin, and S. McLaughlin (2012). "Key Mechanics of the U.S. Tri-Party Repo Market," Federal Reserve Bank of New York *Economic Policy Review* 18, 17-28.

[11] Covitz, D., N. Liang, and G. Suarez (2014). "The Evolution of a Financial Crisis: Collapse of the Asset-Backed Commercial Paper Market." *Journal of Finance*, forthcoming.

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

[13] Duffie, D. and D.A. Skeel (2012). "A Dialogue on the Costs and Benefits of Automatic Stay for Derivatives and Repurchase Agreements." Scholarship at Penn Law Paper 397.

[14] Ennis, H. (2011). "Strategic Behavior in the Tri-Party Repo Market." Federal Reserve Bank of Richmond *Economic Quarterly* 97, 389-413.

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

[16] Gennaioli, N., A. Shleifer and R. Vishny (2012). "Neglected Risks, Financial Innovation, and Financial Fragility." *Journal of Financial Economics* 104, 452-468.

[17] Gorton, G. and A. Metrick (2012). "Securitized Banking and the Run on Repo." *Journal of Financial Economics* 104, 425-451..

[18] Gorton G., and A. Metrick (2013) "Who ran on repo," Manuscript, Yale University.

[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 (2012). "Dynamic Debt Runs." *Review of Financial Studies* 25, 1799-1843.

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

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

[24] Martin, A., D. Skeie, and E. von Thadden (2014). "The Fragility of Short-Term Secured Funding Markets." *Journal of Economic Theory*, forthcoming.

[25] Pozsar, Z., T. Adrian, A. Ashcraft, and H. Boesky (2012). "Shadow Banking." Federal Reserve Bank of New York *Staff Report* No. 580.

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

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