

# The fragile capital structure of hedge funds and the limits to arbitrage\*

Xuewen Liu<sup>†</sup>  
HKUST

Antonio S. Mello  
Wisconsin School of Business

Forthcoming in *Journal of Financial Economics*

## Abstract

During a financial crisis, when investors are most in need of liquidity and accurate prices, hedge funds cut their arbitrage positions and hoard cash. The paper explains this phenomenon. We argue that the fragile nature of the capital structure of hedge funds, combined with low market liquidity, creates a risk of coordination in redemptions among hedge fund investors that severely limits hedge funds' arbitrage capabilities. We present a model of hedge funds' optimal asset allocation in the presence of coordination risk among investors. We show that hedge fund managers behave conservatively and even abstain from participating in the market once coordination risk is factored into their investment decisions. The model suggests a new source of limits to arbitrage

JEL classification: G14; G24; D82

Keywords: Limits to arbitrage; Coordination risk; Fragile capital structure; Market liquidity

---

\*We thank Sudipto Bhattacharya, Enrico Biffis, Robert Kosowski, Tianxi Wang, seminar participants at the 2009 Imperial College hedge fund conference, and, in particular, Amil Dasgupta (the discussant) for helpful comments. We are very grateful to the referee for detailed comments and suggestions that significantly improved the paper.

<sup>†</sup>Corresponding author. E-mail address: Xuewenliu@ust.hk (X. Liu).

“One fund-of-funds manager says he rushes to be the first out if he suspects that others may desert a hedge fund. ... If managers [of hedge funds] worry that clients will bail out, they may try to raise cash in anticipation.”

*Economist*, October 25, 2008, pp. 87–88.

## 1 Introduction

In the financial crisis of 2007–2009, hedge funds reduced their exposures to risky investments and increased their cash holdings very quickly. The *Economist* estimates that between July and August 2008 alone, the industry’s cash holdings rose from \$156 billion to a record \$184 billion, equivalent to 11% of assets under management.<sup>1</sup> The reallocation toward cash seemed to be in anticipation of unprecedented pressure for redemptions from investors. In fact, the hedge fund industry experienced record levels of redemptions during the third and fourth quarters of 2008. In the first half of 2009, investors continued to pull money out of hedge funds. Ironically, redemptions appear to have been self-reinforcing. A report titled *Hedge Funds 2009* by the International Financial Services London Research (IFSL) says:

Hedge funds faced unprecedented pressure for redemptions in the latter part of 2008, with investors withdrawing funds due to dissatisfaction with the performance or to cover for even greater losses or cash calls elsewhere. This in turn led to forced selling and closures of positions by hedge funds causing a cycle of further losses and redemptions. Some funds were not able to meet withdrawal requests so were forced to suspend redemptions, as selling illiquid assets would have damaged the investors that remained.

In this paper we ask why hedge funds, with a reputation for being aggressive investors seeking high returns, hold significant amounts of cash in their balance sheets. And, why do they quickly increase their cash holdings (hoard liquidity) when a crisis strikes? One would think that it is precisely during times of high uncertainty and high volatility that arbitrage opportunities are greater.<sup>2</sup> Why then do hedge funds become much less aggressive in exploiting price distortions and

---

<sup>1</sup>“Hedge funds in trouble: the incredible shrinking funds,” *Economist*, October 25, 2008, pp. 87–88.

<sup>2</sup>Mitchell and Pulvino (2011) and Krishnamurthy (2010) provide convincing evidence that, during the recent liquidity crisis, arbitrage opportunities were pronounced in debt markets.

pass up profit opportunities?

In their quest for arbitrage gains, hedge funds also perform the important social role of enforcing market efficiency. From a social welfare point of view, it is not optimal that hedge funds allocate large fractions of their portfolios to riskless securities, instead of pursuing arbitrage opportunities in risky assets. Holding cash is synonymous with imposing limits on arbitrage. With insufficient arbitrage trading, financial markets can wander off erratically and risk falling into a vicious cycle that goes from lower price informativeness to less liquidity and back, a process of continued deterioration that can prolong a financial crisis. Thus, to analyze a modern financial crisis it is necessary to understand the reasons that hedge funds limit their arbitrage activities.

So far the finance literature has associated the limits to arbitrage with investor uncertainty about the fundamental value of the assets held by arbitrageurs. According to Shleifer and Vishny (1997), when investors do not understand or observe perfectly arbitrageurs' trading positions, they can react with their feet after observing poor performance. Understanding that poor performance can cause redemptions, hedge fund managers refrain from making investments that might lose money in the short term, even if a profit could be realized in the long run.<sup>3</sup>

In this paper, we offer an alternative explanation for the limits of arbitrage. We highlight investors' concerns about coordination risk (i.e., the uncertainty about what other investors might decide to do), instead of about fundamental risk. One key feature of the capital structure of hedge funds is the fragile nature of their equity. Equity capital in hedge funds can be redeemed at investors' discretion, a feature somewhat similar to demand deposit-debt in banks.<sup>4</sup> The fragile equity capital in hedge funds introduces the risk of coordination among hedge fund investors. The coordination risk arises because investors suspect that other investors might redeem, and to meet redemptions the hedge fund may be forced to liquidate positions at a loss. If investors suspect that, after

---

<sup>3</sup>Gromb and Vayanos (2002) build on the intuition of Shleifer and Vishny (1997) and show that margin constraints have a similar effect in limiting the ability of arbitrageurs to exploit price differences. Abreu and Brunnermeier (2002, 2003) use synchronization risk to explain market inefficiency.

<sup>4</sup>It might seem that, because of the lock-up period clause, the equity in hedge funds is not fragile. In practice, however, hedge fund managers are reluctant to use lock-up periods and early-withdrawal penalties because these might signal a lack of confidence in their trading strategies. More problematic, hedge funds with lock-ups in place often grant side deals to 'special investors' (see Brunnermeier, 2009; and Stein, 2005). Indeed, evidence shows that few hedge funds, only those with the strongest records, can lock money in for long periods.

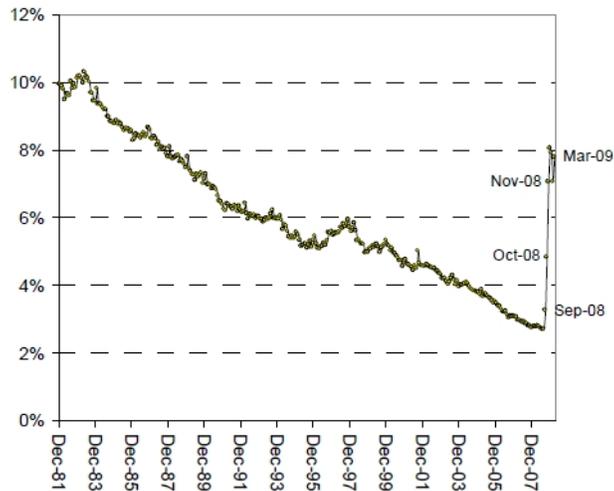
liquidating positions to satisfy early redemptions, the hedge fund would be left with insufficient equity, a clear advantage exists to being a first mover. Then, even long-horizon investors could decide to withdraw, resulting in a vicious cycle of redemptions and asset sales that can trigger a disorderly collapse of the hedge fund. This risk of coordination must be taken into account by prudent hedge fund managers, who limit their arbitrage activities and hoard cash both to honor redemptions at little cost and to reassure concerned investors.

We use global game methods to model the asset allocation decision of a hedge fund that is subject to the risk of a run by its investors. We start by considering the benchmark case with no coordination problem. A random number of investors is assumed to redeem early for exogenous reasons. The random amount of the early redemptions requires the hedge fund to decide how much cash it needs to hold *ex ante*. The hedge fund's trade-off is between lower potential liquidation costs if it holds more cash, and a higher return if it invests more in risky assets. This trade-off gives an optimal level of cash holdings. In the presence of coordination risk, however, we show that hedge funds choose to hold more (i.e., excess) cash. Coordination risk affects the optimal level of cash in a fund in two ways. First, the fear of a possible run makes a greater number of investors redeem even if they do not face liquidity shocks. To satisfy redemptions at minimum cost, hedge funds optimally choose to hold more cash. Second, cash holdings have a direct impact on investors' decision to withdraw and, consequently, on the probability of a run. Naturally, the probability of a run is decreasing in the cash holdings.

In sum, the source of the limits to arbitrage in our paper is market illiquidity and the unstable nature of the equity capital in hedge funds. We emphasize the friction of investors' uncertainty about the actions of other investors in the fund, i.e., the liability side of the balance sheet, while in Shleifer and Vishny (1997) investors are uncertain about the asset side of the balance sheet, i.e., asset quality. Limits to arbitrage due to coordination risk are particularly likely in financial crises, when market liquidity is low and coordination problems are severe. We believe that a fuller explanation for the limits to arbitrage requires that the two motives triggering investors' redemptions be taken together.

It is interesting to contrast the behavior of hedge funds with that of banks before and during the financial crisis of 2007–2009. Before the crisis, on average, hedge funds operated with much

lower leverage ratios than banks.<sup>5</sup> During the crisis, hedge funds also reduced their exposures to illiquid assets more quickly and more drastically. Fig. 1 shows that banks switched systematically to cash holdings only after the collapse of Lehman Brothers in September 2008, much later than hedge funds. We argue that the different investment and financing policies are due in part to differences in the nature of the capital structures of the two types of institutions. Equity in hedge funds is fragile, because it can be redeemed at the request of investors, while equity in banks is locked in permanently, because it cannot be tendered back to the bank for redemption. The fragile nature of the equity (and debt) in hedge funds creates an effective market-disciplining mechanism. It induces hedge funds to behave conservatively, by holding less risky portfolios and operating with less leverage. This difference may help to explain why hedge funds survived much better than banks when the initial mortgage default crisis produced *unexpectedly* large dislocations and turmoil in financial markets in 2007–2009. Many observers were surprised by the fact that hedge funds, typically regarded as the riskiest segment of the financial markets, were much more resilient than banks. Our paper helps to explain this phenomenon: The fragile capital structure of hedge funds encourages micro-prudence, but it also curtails arbitrage trading, leading to macro-unsoundness.



**Fig. 1. US commercial bank cash ratio increased drastically only after September 2008.**

Source of data is Federal Reserve H8.  
Figure is from Morris and Shin (2009).

Our work relates to a large literature on bank runs. Gorton and Winton (2003) provide an excellent survey. Gorton (2008) discusses in detail the panic runs in the recent crisis. The setup

<sup>5</sup>At the beginning of the crisis of 2007–2009, the hedge fund industry as a whole had a ratio of assets to equity of 1.8, compared with more than 20 for investment banks (*Economist*, October 25, 2008, pp. 87).

in our model is similar to that in Diamond and Dybvig (1983). The difference is that we assume that the size of early withdrawals is ex ante random, and this allows us to study the optimal cash holding problem of a hedge fund. In general, a bank-run game has multiple equilibria, but we resort to global game methods to solve for the unique equilibrium. In this respect, our paper is close to the work of Goldstein and Pauzner (2005), but the purpose and the scope of the two papers differ.<sup>6</sup> We study the optimal cash holdings in the balance sheets of hedge funds, while Goldstein and Pauzner study the optimal liquidity provision of banks.

A panic run can be driven by liquidity concerns or by concerns about solvency (see the discussion in Brunnermeier, 2009). In our model, the realization of large early withdrawals forces a hedge fund to liquidate positions in illiquid markets at reduced prices, and this triggers a run. In contrast, the run can be driven by fundamentals, as in Goldstein and Pauzner (2005). Bad fundamentals leave little asset value for investors who withdraw late, and this causes an immediate run. In the recent financial crisis, capital markets experienced many runs triggered for liquidity reasons, rather than for fundamental reasons. There were runs on many financial institutions that appeared to have adequate regulatory capital before they collapsed. Even so, investors ran, worried that the financial institutions had insufficient liquidity to meet massive early withdrawals and that assets would have to be liquidated at a deep discount.

Our work also relates to Chen, Goldstein, and Jiang (2010). They find that mutual fund outflows are significantly more sensitive to poor performance for funds that invest in illiquid assets. This insight is similar to our result that market illiquidity potentially generates a panic run on high-end financial intermediaries, such as hedge funds and mutual funds. Compared with their work, we explicitly model how hedge fund managers factor into their ex ante investment and financing decisions the ex post panic run. In the same vein as Shleifer and Vishny (1997), Stein (2005) argues that open-ended funds, as a result of competition, could be suboptimal, because their investments are limited by the occurrence of early withdrawals in response to poor performance. Our emphasis is on the mechanisms that curtail arbitrage. Our focus on coordination risk differs

---

<sup>6</sup>The global game methodology has been used in various contexts. For example, Rochet and Vives (2004) also use global games to study bank runs. Other applications include currency crises (Morris and Shin, 1998; and Corsetti, Dasgupta, Morris, and Shin, 2004), contagion of financial crises (Dasgupta, 2004; and Goldstein and Pauzner, 2004), and market liquidity (Morris and Shin, 2004; Plantin, 2009; and Plantin, Sapra, and Shin, 2008).

from the arguments presented in these papers. In a banking context, Morris and Shin (2008) also argue that cash held by a debtor bank can lower the threshold of coordination among creditors to not withdraw. They do not formalize the argument, however, which we develop further by studying the ex ante optimal asset allocation decision of hedge funds and the limits to arbitrage.

The paper is organized as follows. In Section 2, we first describe the model setup. Next, we solve the optimal asset allocation for a hedge fund, without and with the coordination problem facing the fund's investors. After that we analyze the model's implications and predictions. In Section 3, we generalize the model and discuss some broader issues of relevance to financial markets. In Section 4, we provide remarks and conclusions.

## 2 The model

We present the model in this section.

### 2.1 *The setup*

We use a model with four dates:  $T_0$ ,  $T_{\frac{1}{2}}$ ,  $T_1$ , and  $T_2$ . All agents are risk-neutral. There is no discount factor between  $T_1$  and  $T_2$ .

#### 2.1.1 *Hedge fund*

Consider a simple setup of a hedge fund that depends on investors for funding. The investors are the limited partners of the fund. The fund begins at  $T_0$  and ends at  $T_2$ . Investors have the right to redeem their investments at an intermediate date,  $T_1$ . We assume that the total amount of assets managed by the fund at  $T_0$  is 1, that there is a continuum of investors with measure 1, and that each investor contributes 1 unit of capital. We model the liquidity risk of the hedge fund's liabilities as follows: Each investor in the hedge fund is either an early investor or a late investor. Early investors are investors who, despite initially having the intention of staying for a long time with the fund, end up redeeming their investments at  $T_1$  for reasons that we leave unspecified but that could be related to consumption, portfolio allocation, or other financial emergencies. The probability that an investor is an early investor, depending on the state of the economy, is ex ante a random

variable, uniformly distributed with support  $\lambda \in [0, \bar{\lambda}]$ , where  $0 < \bar{\lambda} < 1$ . Then, in aggregate, the ex ante proportion of investors who end up as early investors is given by the random variable  $\lambda$ .

At  $T_0$ , the hedge fund needs to make an asset allocation decision: How much capital  $X$  should be invested in cash ( $C$ )? The remaining capital,  $1 - X$ , is invested in illiquid assets ( $A^L$ ). The cost of investment per unit of the illiquid assets at  $T_0$  is 1. We assume illiquid assets have a (gross) return of  $R$  at  $T_2$ , without any uncertainty, where  $R > 1$ . However, if illiquid assets are liquidated early, at  $T_1$ , they are sold at a discount, where the discounted price at fire sale is  $\alpha$  (per unit), with  $0 < \alpha < 1$ . Fig. 2 shows the hedge fund's balance sheet. Notice that there is no leverage, only equity. We discuss a levered hedge fund later.

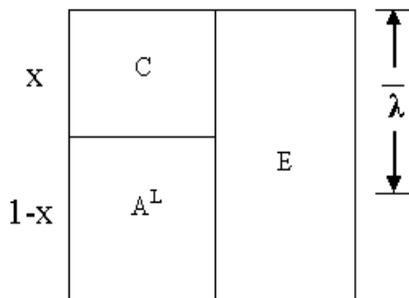


Fig. 2. The hedge fund's balance sheet

### 2.1.2 Investors

At  $T_{\frac{1}{2}}$ , uncertainty about the investors' status is resolved. At that time each investor knows whether he is an early investor or not, but no investor knows the status of other investors. Nevertheless, late investors receive signals about the state of the economy or the aggregate number of early investors ( $\lambda$ ). More specifically, late investor  $i$ 's signal is  $\lambda^i = \lambda + \epsilon^i$ , where  $\epsilon^i$  is a uniformly distributed variable with support  $[-\epsilon, \epsilon]$ .  $\epsilon^i$  is independent from  $\epsilon^j$  for  $i \neq j$ . In most of our analysis,  $\epsilon$  is assumed to be arbitrarily small:  $\epsilon \rightarrow 0$ , which is a typical setup in applications of global games. Further, at  $T_{\frac{1}{2}}$ , investors are perfectly informed about the cash position  $X$  as well as about the market depth,  $\alpha$ . The fund's position in cash can be known either because the fund discloses this in a regular letter to investors or because many investors in hedge funds are themselves professional investors with access to the management of the fund. Usually, a hedge fund is very cautious about

disclosing trading strategies or positions in specific assets, but it does not object to disclosing allocations in general asset classes.

At  $T_{\frac{1}{2}}$ , all investors need to decide whether to stay in the fund until  $T_2$  or to redeem their shares at  $T_1$ , in which case they inform the fund about their decisions. All early investors have to redeem. A late investor might decide to redeem if he thinks that many other investors may redeem, because too many withdrawals force the hedge fund to liquidate illiquid investments at a cost. It is the coordination problem among late investors that we want to study in this paper.

Let us consider the payoffs of an investor, whether an early or a late investor, from redemption and nonredemption. Suppose that the aggregate number of investors that redeem at  $T_1$  is given by  $s$ , where  $0 \leq s \leq 1$ .

The payoff function of an investor that redeems is

$$w^R(s) = \begin{cases} 1 & \text{if } 0 \leq s \leq X + (1 - X)\alpha \\ \frac{X + (1 - X)\alpha}{s} & \text{if } X + (1 - X)\alpha < s \leq 1 \end{cases} \quad (1)$$

Note that, in Eq. (1), investors withdrawing their shares at  $T_1$  receive a fraction of the hedge fund's net asset value. At  $T_{\frac{1}{2}}$ , when investors give the fund notice of their redemptions, the hedge fund's mark-to-market net asset value is 1, not the amount  $X + (1 - X)\alpha$  shown in Eq. (1). This is because the hedge fund possesses cash reserves, and as long as it has not started to liquidate illiquid assets, the mark-to-market value of the fund is 1.<sup>7</sup> Investors that withdraw are paid based on a proportion of this initial mark-to-market value of the fund, 1. However, if too many withdrawals occur, the hedge fund must liquidate illiquid assets after it exhausts its cash balance,  $X$ . The sale price of the illiquid assets is only  $\alpha$ , less than 1. When this happens, the realized value of an investor's share of the assets is lower than his claim. To fulfill early investors' claims, the fund has to liquidate part of the late investors' shares of the assets. The fund might have to sell all assets at  $T_1$ , in which case late investors are left with nothing. This first-mover advantage of early

---

<sup>7</sup>We can assume that, when the hedge fund sells its illiquid assets, the fund faces a special downward-sloping demand curve. Specifically, the market can absorb just a tiny amount of the assets at a price of 1. After absorbing that amount, the price drops to  $\alpha$  and the demand curve becomes perfectly elastic with constant price  $\alpha$ . Therefore, before the sale, the mark-to-market value of the fund is 1, but at liquidation it becomes  $X + (1 - X)\alpha$ .

withdrawals, due to discounted prices at fire sales and the first-come, first-served right to cash flows, is the fundamental reason for a run on hedge funds and on mutual funds (see the discussion in Brunnermeier, 2009).

The payoff function of an investor that does not redeem at time  $T_1$  (but stays) is

$$w^S(s) = \begin{cases} \frac{(X-s)+(1-X)R}{1-s} & \text{if } 0 \leq s \leq X \\ \frac{(1-X)-\frac{s-X}{\alpha}}{1-s} R & \text{if } X < s \leq X + (1-X)\alpha \\ 0 & \text{if } X + (1-X)\alpha < s \leq 1 \end{cases} \quad (2)$$

In Eq. (2), when the aggregate number of investors that redeem at  $T_1$  is less than  $X$ , the fund does not need to liquidate any long-term illiquid assets. Thus the value of the fund at  $T_2$  is  $(X - s) + (1 - X)R$ . However, if the aggregate number of investors withdrawing is higher than  $X$ , the fund must sell some of its illiquid assets at  $T_1$ . The number of units of the illiquid assets that need to be sold is  $\frac{s-X}{\alpha}$ , where  $s - X$  is the amount of cash short to satisfy redemptions. If the number of withdrawing investors exceeds the total liquidation value of the fund,  $X + (1 - X)\alpha$ , the hedge fund is completely liquidated and nothing is left after  $T_1$ , the third line in Eq. (2).

We define the difference between the payoffs of not withdrawing at time  $T_1$  and withdrawing as  $\Delta w(s) \equiv w^S(s) - w^R(s)$ . Fig. 3 shows the payoffs  $w^S(s)$  and  $w^R(s)$ , and Fig. 4 shows  $\Delta w(s)$ .

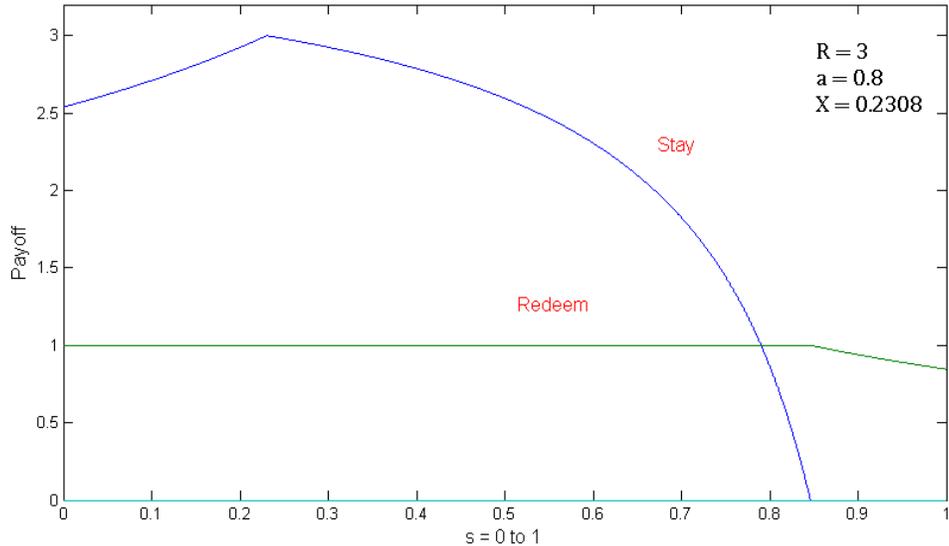


Fig. 3. The payoffs of staying and redeeming

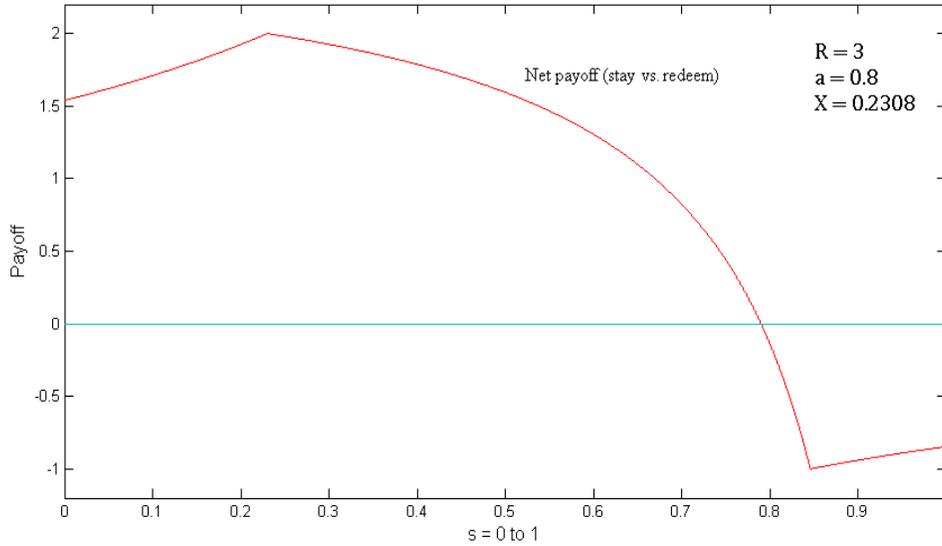


Fig. 4. The difference between the payoffs of not withdrawing and withdrawing

As in Diamond and Dybvig (1983) and Shleifer and Vishny (1997), we assume that the hedge fund is not able to replace redeeming investors with new investors.<sup>8</sup>

Fig. 5 summarizes the timeline.

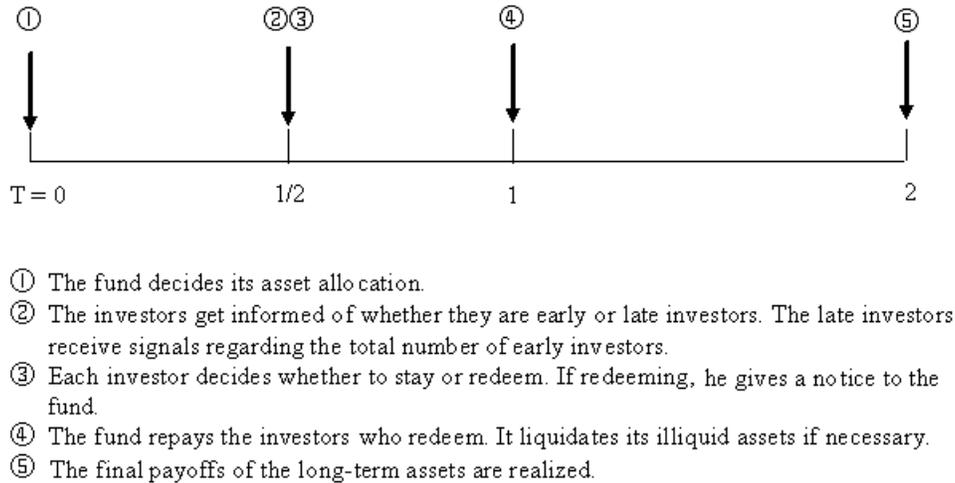


Fig. 5. The timeline

<sup>8</sup>This assumption can be justified with moral hazard or adverse selection arguments. In line with Diamond and Rajan (2000, 2001), only incumbent investors having one period of experience can monitor the fund manager or understand the true performance of the fund. The new investors cannot. Mitchell and Pulvino (2010) find empirical evidence supporting this assumption.

## 2.2 *Equilibrium*

In this subsection, we solve for the equilibrium and determine the optimal asset allocation for the hedge fund.

### 2.2.1 *Benchmark case: no coordination problem among hedge fund investors*

In the benchmark equilibrium, we assume that there is no coordination problem for late investors. All investors that inform the fund at  $T_{\frac{1}{2}}$  of their decision to redeem at  $T_1$  are genuine early investors. Late investors do not withdraw. We work out the optimal amount of cash holdings of the hedge fund by backward induction from  $T_{\frac{1}{2}}$  to  $T_0$ .

Following the setup in Shleifer and Vishny (1997), the hedge fund's objective is to maximize the fund's total equity value at  $T_2$ .<sup>9</sup> This objective is consistent with a hedge fund manager's maximizing the value of assets under management. The final value of the fund is an indicator of the fund's performance and affects its future capacity to raise funds. If the redemption shock  $\lambda$  is small and the cash holdings  $X$  are sufficient to cover investors' withdrawals at  $T_1$ , the total equity value of the fund at  $T_2$  is  $(X - \lambda) + (1 - X)R$ . By contrast, if the redemption shock  $\lambda$  is greater than  $X$ , the fund has to liquidate  $\min\{\frac{\lambda - X}{\alpha}, 1 - X\}$  units of the illiquid assets and the equity value at  $T_2$  is  $[(1 - X) - \frac{\lambda - X}{\alpha}]R$  if  $1 - X \geq \frac{\lambda - X}{\alpha}$  [or  $\lambda \leq X + \alpha(1 - X)$ ] and zero if  $1 - X < \frac{\lambda - X}{\alpha}$  [or  $\lambda > X + \alpha(1 - X)$ ]. Therefore, the expectation of the fund's equity value at  $T_2$  can be expressed as

$$W^I(X) = \frac{1}{\lambda} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_X^{\min[\bar{\lambda}, X + (1 - X)\alpha]} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\}. \quad (3)$$

The hedge fund manager chooses an  $X$  at  $T_0$  that maximizes the expected value of the equity at  $T_2$ :  $\text{Max}_X W^I(X)$ . In this maximization problem, the trade-off is between the return on investment and the cost of liquidating illiquid assets. A higher  $X$  reduces costly sales of illiquid positions in case the redemptions are unexpectedly large. Yet, a higher  $X$  wastes investment opportunities and

---

<sup>9</sup>Under the alternative assumption that the hedge fund's objective is to maximize fees on assets under management plus a performance incentive, "2%+20%," the results of the model do not change qualitatively, only quantitatively. But the model would be harder to solve analytically.

forgoes the high return  $R$  in case the redemption shock is small. The trade-off leads to a unique optimal value for the cash holdings,  $X^*$ .

**Theorem 1** *If  $0 < \bar{\lambda} \leq \frac{R-\alpha}{R(2-\alpha)-1}$ , the optimal amount of cash holdings is  $X^* = \frac{R-R\alpha}{R-\alpha}\bar{\lambda}$ . The fund survives any redemption shock  $\lambda \in [0, \bar{\lambda}]$ , where  $\bar{\lambda} \leq X^* + (1 - X^*)\alpha$ . If  $\frac{R-\alpha}{R(2-\alpha)-1} < \bar{\lambda} < 1$ , the optimal amount of cash holdings is  $X^* = \frac{R(1-\alpha)}{R(2-\alpha)-1}$ . The fund survives if the redemption shock lies within  $\lambda \in [0, X^* + (1 - X^*)\alpha] \subset [0, \bar{\lambda}]$ , where  $X^* + (1 - X^*)\alpha < \bar{\lambda}$ .*

Proof: See the Appendix.

## 2.2.2 Equilibrium when there is a coordination problem among hedge fund investors

When the cash holdings,  $X$ , do not exceed the realized redemption shock,  $\lambda$ , a costly sale of illiquid assets is unavoidable. The fire sale depresses the fund's net asset value and potentially causes a run by investors. The fund manager anticipates this possibility of a run and rationally takes it into account when deciding the fund's cash balances.

Again, we solve the equilibrium by backward induction, from  $T_{\frac{1}{2}}$  to  $T_0$ . In the first step, we work out investors' decisions at  $T_{\frac{1}{2}}$  for a given cash balance  $X$ . In the second step, we go back to  $T_0$  and solve for the fund's optimal cash holdings  $X$  by taking into account the investors' responses.

### 2.2.2.1. Investors' decision at $T_{\frac{1}{2}}$ given the cash holdings $X$

At  $T_{\frac{1}{2}}$ , late investor  $i$ 's decision rule can be characterized as a map:  $(\lambda^i, X) \mapsto (\text{Withdraw}, \text{Stay})$ , where  $(\lambda^i, X)$  is the information set of investor  $i$  and  $(\text{Withdraw}, \text{Stay})$  is his choice set.

We need to work out investors' decision rules that form an equilibrium. We first examine late investors' strategies when the state of economy  $\lambda$  is extremely low or extremely high. We have the following two extreme cases.

First, there exists a *lower dominance region* of  $\lambda \in [0, \lambda^L)$ , in which staying in the fund is the dominant strategy for late investors. To obtain the lower dominance region in the bank-run game, we follow Goldstein and Pauzner (2005) and modify the technology for a very low  $\lambda$ . Specifically, we

assume that when the proportion of early investors (the state of the economy)  $\lambda$  falls within  $[0, \lambda^L)$  (where  $\lambda^L$  is close to zero), there is no discount in the sale price of the illiquid assets, i.e.,  $\alpha = 1$ .<sup>10</sup> With this modification, we can conclude that if a late investor knows that  $\lambda$  is in the region  $[0, \lambda^L)$ , he does not run, no matter his beliefs about other investors' actions. This is because if there is no discount, early withdrawals do not erode the late investors' shares. Thus, late investors do not suffer any disadvantage if they wait until  $T_2$  to withdraw. In most of our analysis,  $\varepsilon$  is taken to be arbitrarily close to zero, so late investors are sure that  $\lambda$  is in  $[0, \lambda^L)$ . Staying is therefore their dominant strategy.

Second, using the result that  $s^* = X + (1 - X) \frac{\alpha(R-1)}{R-\alpha}$  solves the equation  $w^S(s) = w^R(s)$ , we define  $\lambda^U \equiv s^*$ . Then, an *upper dominance region* of  $\lambda \in (\lambda^U, \bar{\lambda}]$  exists, in which withdrawing is the dominant strategy for late investors. The reason is as follows: When  $\lambda > \lambda^U$ , there are a sufficient number of redemptions by early investors, and, consequently, even if all late investors choose to stay, the payoff of a late investor if he decides to redeem early is higher than the payoff if he decides to stay on. If a late investor is sure that  $\lambda > \lambda^U$ , he should redeem no matter his beliefs about the actions of other late investors. As  $\varepsilon$  is taken to be arbitrarily close to zero, late investors are informed when  $\lambda > \lambda^U$ . Therefore, if  $\lambda$  is in the interval  $(\lambda^U, \bar{\lambda}]$ , withdrawing is the dominant strategy for late investors. In our model, we show that it is not optimal for the hedge fund to choose cash holdings  $X$  such that  $X + (1 - X) \frac{\alpha(R-1)}{R-\alpha} \geq \bar{\lambda}$ . Therefore, the condition  $\bar{\lambda} > \lambda^U$  is always true in our model. That is, the upper dominance region  $(\lambda^U, \bar{\lambda}]$  exists.

For the intermediate region of  $\lambda \in [\lambda^L, \lambda^U]$ , late investors' decisions depend on their beliefs about the actions of other late investors. The signals regarding  $\lambda$  form their beliefs. Our main interest is in the threshold-strategy equilibrium in which a late investor's strategy depends on the signal he receives. Specifically, the threshold-strategy equilibrium is the equilibrium in which every late investor sets a threshold  $\lambda^*(X)$  and uses the threshold strategy  $(\lambda^i, X) \mapsto \begin{cases} \text{Withdraw} & \lambda^i > \lambda^*(X) \\ \text{Stay} & \lambda^i \leq \lambda^*(X) \end{cases}$ , such

---

<sup>10</sup>Using  $\lambda$  to represent the state of the economy has the following meaning: A low  $\lambda$  indicates that few investors are in need of withdrawing (funding liquidity) because the economy is doing well. Presumably, these are also times when the level of market liquidity is high and  $\alpha$  is consequently high. Implicitly, we are assuming that funding liquidity and market liquidity are positively correlated when the state of the economy is bad. This is consistent with the empirical evidence.

that, given that all other late investors set the threshold as  $\lambda^*(X)$ , it is optimal for a particular late investor to do that. We prove that there is a unique threshold equilibrium for our model.

We use global game methods to solve for the unique threshold equilibrium. Note that we do not consider equilibria other than the threshold strategy in this paper. With the assumptions made before about the existence of the two extreme dominance regions of  $\lambda$ , it is possible to provide a proof, as in Goldstein and Pauzner (2005), that the threshold equilibrium is the only equilibrium.<sup>11</sup>

The equilibrium result is summarized in Theorem 2.

**Theorem 2** *The model has a unique threshold equilibrium. In the equilibrium, a late investor redeems if, and only if, his signal is above the threshold  $\lambda^*(X)$ .*

We prove the theorem in several steps.

First, suppose that every late investor uses the threshold strategy and sets the threshold as  $\lambda^*(X)$ . We want to determine the total number of investors (both early and late investors) that decide to redeem at  $T_1$  for a given realization of  $\lambda$ . Given the realized  $\lambda$  and the threshold  $\lambda^*$ , the proportion of late investors that run equals  $\frac{\lambda + \epsilon - \lambda^*}{2\epsilon}$ . Hence, the total number of investors redeeming, denoted by  $s(\lambda, \lambda^*(X))$ , is

$$s(\lambda, \lambda^*(X)) = \begin{cases} \lambda & \text{if } \lambda \leq \lambda^*(X) - \epsilon \\ \lambda + (1 - \lambda) \frac{\lambda + \epsilon - \lambda^*(X)}{2\epsilon} & \text{if } \lambda^*(X) - \epsilon \leq \lambda \leq \lambda^*(X) + \epsilon \\ 1 & \text{if } \lambda \geq \lambda^*(X) + \epsilon \end{cases} \quad (4)$$

When  $\lambda^* - \epsilon \leq \lambda \leq \lambda^* + \epsilon$ ,  $s(\lambda, \lambda^*(X))$  has two components. The first component represents redemptions by early investors, and the second one represents redemptions by some late investors who change their minds and decide instead to redeem early.

Second, after the late investor receives the signal  $\lambda^i$  his posterior distribution of  $\lambda$  is uniform in the interval  $[\lambda^i - \epsilon, \lambda^i + \epsilon]$ . Therefore, the investor's expected net payoff of staying versus redeeming is

---

<sup>11</sup>In fact, it has become a standard result in the global game literature [see Morris and Shin (2003) for a discussion of Goldstein and Pauzner's result] that the threshold equilibrium is the only equilibrium for a bank-run game as long as the noise of the signal is uniformly distributed.

$$\pi(\lambda^i, \lambda^*) = \frac{1}{2\epsilon} \int_{\lambda^i - \epsilon}^{\lambda^i + \epsilon} \Delta w(s(\lambda, \lambda^*)) d\lambda. \quad (5)$$

Third, the marginal late investor who receives the threshold signal  $\lambda^*$  is indifferent between redeeming and staying. Therefore,  $\pi(\lambda^*, \lambda^*) = 0$ . That is,

$$\frac{1}{2\epsilon} \int_{\lambda^* - \epsilon}^{\lambda^* + \epsilon} \Delta w(s(\lambda, \lambda^*)) d\lambda = 0. \quad (6)$$

As previously pointed out, in our model we focus on the case of an arbitrarily small  $\epsilon$ :  $\epsilon \rightarrow 0$ . This assumption makes the model tractable and helps to highlight the main insights of the paper. When  $\epsilon \rightarrow 0$ , fundamental uncertainty disappears, while strategic uncertainty remains unchanged (see Morris and Shin, 1998). In our context, from the marginal late investor's perspective,  $\lambda$  falls within  $[\lambda^* - \epsilon, \lambda^* + \epsilon]$ . When  $\epsilon \rightarrow 0$ , the uncertainty associated with the first element of  $s(\lambda, \lambda^*)$  disappears while the uncertainty related to the second element does not. As a whole,  $s(\lambda, \lambda^*)$  is uniformly distributed in the interval  $[\lambda^*, 1]$ . Fig. 6 illustrates the idea. When  $\epsilon \rightarrow 0$ , the curve between points A and B, which depicts function  $s(\lambda, \lambda^*)$ , becomes a straight line.

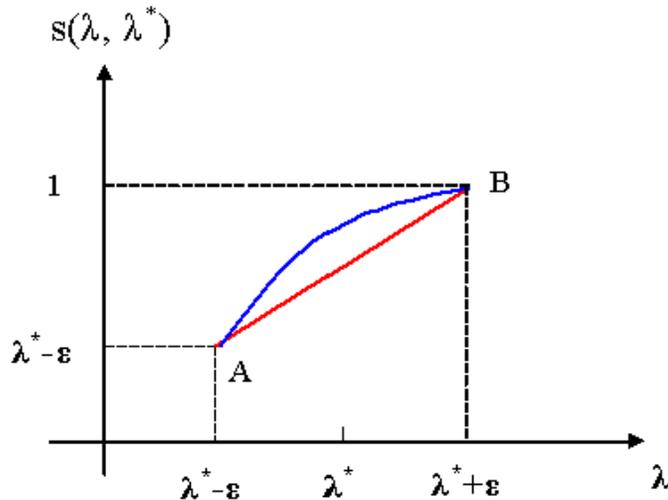


Fig. 6.  $s$  as a function of  $\lambda$

Given the uniform distribution of  $s(\lambda, \lambda^*)$ , we can rewrite Eq. (6) as

$$\int_{\lambda^*}^1 \Delta w(s) ds = 0. \quad (7)$$

From Eq. (7), we can solve for the unique  $\lambda^*$ . Geometrically,  $\lambda^*$  corresponds to the point in Fig. 7 that equates the shaded area below and above the x-axis.

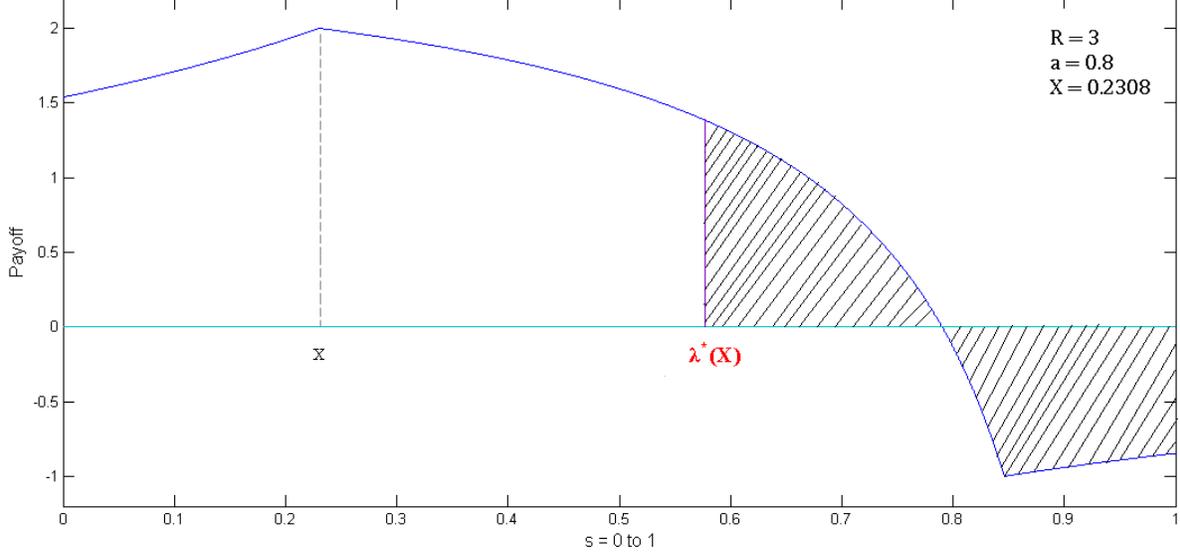


Fig. 7. The equilibrium

Finally, to show that  $\lambda^*(X)$  indeed forms a threshold equilibrium, we need to show that a late investor  $i$  receiving a signal higher (lower) than  $\lambda^*$  prefers to run (stay):  $\pi(\lambda^i, \lambda^*) < 0$  ( $\pi(\lambda^i, \lambda^*) > 0$ ). In fact, from late investor  $i$ 's perspective,  $\lambda$  is uniformly distributed in the interval  $[\lambda^i - \epsilon, \lambda^i + \epsilon]$ . If  $\lambda^i > \lambda^*$ , then as  $\epsilon \rightarrow 0$ ,  $s(\lambda, \lambda^*)$  is uniformly distributed in the interval  $[\lambda^i + (1 - \lambda^i) \frac{\lambda^i - \lambda^*}{2\epsilon}, 1]$  and has positive probability mass at 1. This is equivalent to taking a slice from the positive part of the shaded area above the x-axis in Fig. 7 and adding it to the negative part of the shaded area. Hence  $\pi(\lambda^i, \lambda^*) < 0$ . A similar argument applies when  $\lambda^i < \lambda^*$ .

In the proof of Theorem 2, we have shown how to find the threshold  $\lambda^*(X)$ . We now need to determine the parameter conditions that ensure that  $\lambda^*(X)$  exists. Theorem 3 establishes sufficient conditions for the existence of  $\lambda^*(X)$  as well as for  $\lambda^*(X)$  being on the right of  $X$  (i.e.,  $\lambda^*(X) > X$ ).

**Theorem 3** *If  $(R - \alpha) + \frac{R}{\alpha}(1 - \alpha) \log(1 - \alpha) > 1 - \alpha$  and  $\bar{\lambda} > X + (1 - X) \frac{\alpha(R-1)}{R-\alpha}$ , there exists a unique  $\lambda^*(X)$  and  $\lambda^*(X) > X$ .*

Proof: See the Appendix.

In addition, we want to make sure that  $\lambda^*(X)$  increases in  $X$ : The higher the cash balance,  $X$ , the higher is the threshold that late investors run. We find the necessary and sufficient conditions to guarantee that  $\lambda^*(X)$  increases in  $X$ . In our model, however,  $\Delta w(s)$  is not strictly monotonically decreasing in  $s$ . Finding the conditions is not trivial.

**Theorem 4**  $\lambda^*(X)$  is an increasing function with respect to  $X$  when  $\frac{R}{\alpha} \log(\frac{1-\alpha}{1-b}) < \log \alpha$  (where  $b$  solves  $(\frac{R}{\alpha} - 1)(\alpha - b) + \frac{R}{\alpha}(1 - \alpha) \log(\frac{1-\alpha}{1-b}) = -\alpha \log \alpha$ ).<sup>12</sup>

Proof: See the Appendix.

As we have obtained the unique threshold and because  $\epsilon \rightarrow 0$ , we conclude that all late investors run if  $\lambda > \lambda^*(X)$  and none runs if  $\lambda \leq \lambda^*(X)$ . Therefore, the total number of investors redeeming their investments from the hedge fund can be expressed as

$$s(\lambda) = \begin{cases} \lambda & \text{if } \lambda \leq \lambda^*(X) \\ 1 & \text{if } \lambda > \lambda^*(X) \end{cases}. \quad (8)$$

### 2.2.2.2. The fund's optimal cash holding decision at $T_0$

We now go back to  $T_0$  and look at the hedge fund's optimal cash holding decision when a run by investors might occur. To save space and simplify the analysis, we focus on the case  $\bar{\lambda} > \frac{R-\alpha}{R(2-\alpha)-1}$ . For the case  $\bar{\lambda} \leq \frac{R-\alpha}{R(2-\alpha)-1}$ , the main results do not change.<sup>13</sup>

We analyze two regions of  $X$ , in order:  $X \in \{X \mid X + (1 - X) \frac{\alpha(R-1)}{R-\alpha} < \bar{\lambda}\}$  and  $X \in \{X \mid X + (1 - X) \frac{\alpha(R-1)}{R-\alpha} \geq \bar{\lambda}\}$ . In the first region,  $X$  is low and the only equilibrium is the threshold equilibrium. In the second region of  $X$ , we make a weak assumption that investors coordinate and choose the Pareto efficient equilibrium—the staying-strategy equilibrium. That is, late investors ignore the signals received and stay:  $(\lambda^i, X) \mapsto \text{Stay}$ .<sup>14</sup>

<sup>12</sup>The parameter conditions in Theorems 3 and 4 are not restrictive at all. A wide set of parameters achieves the results. In fact, the reason for requiring parameter conditions is that a bank-run game does not have strictly global strategic complementarities, i.e., in our model,  $\Delta w(s)$  is not strictly monotonically decreasing in  $s$ . However, in Fig. 7, we see the intuition for why the parameter conditions are not restrictive: Toward the end of  $s = 1$ ,  $\Delta w(s)$  does not decrease but increases very slowly.

<sup>13</sup>In fact, we can assume that  $\bar{\lambda}$  must satisfy  $\bar{\lambda} > \frac{R-\alpha}{R(2-\alpha)-1}$ . That is, we can assume that only these are the parameter values that are relevant in this research. But it is easy to prove that the conclusions in this subsection are true in the alternative case of  $\bar{\lambda}$ . The economic idea of the proof is exactly the same. Only the mathematical technique changes.

<sup>14</sup>If investors still use the threshold equilibrium for the second region of  $X$ , the model result is strengthened, because we do not need to consider the second region of  $X$  separately.

We begin the analysis from the low region of  $X$ , i.e.,  $X \in [0, X^C]$ , where  $X^C + (1 - X^C) \cdot \frac{\alpha(R-1)}{R-\alpha} = \bar{\lambda}$ . In this region of  $X$ , investors use the threshold strategy in equilibrium. Because investors use the threshold strategy and set the threshold as  $\lambda^*(X)$ , the fund knows that if the realized redemption shock is greater than  $\lambda^*(X)$ , all investors redeem and desert the fund, driving its net asset value to zero. The fund survives if  $\lambda \in [0, \lambda^*(X)]$ .<sup>15</sup>

The expected value of the fund is

$$W^{II}(X) = \frac{1}{\lambda} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_X^{\lambda^*(X)} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\} \quad \text{for } X \in [0, X^C]. \quad (9)$$

Because  $W^{II}(X)$  is a continuous and bounded function with respect to  $X$  in the support  $[0, X^C]$ , the problem  $\underset{X}{Max} W^{II}(X)$  has a solution.<sup>16</sup> The first-order condition of Eq. (9) is

$$\left[ \left(1 - \frac{R}{\alpha}\right)X + \frac{1-\alpha}{\alpha} \lambda^*(X)R \right] + \left[ (X + (1 - X)\alpha - \lambda^*(X)) \right] \cdot \frac{R}{\alpha} \cdot \frac{d\lambda^*(X)}{dX} = 0. \quad (10)$$

The trade-off in Eq. (10) is as follows. The first term on the left-hand side is the effect of a higher  $X$  on the value of the fund, given the survival range  $[0, \lambda^*(X)]$ . This term captures the trade-off when there is no run. The trade-off is between higher returns ( $R$ ) by investing more in illiquid assets and a lower cost of liquidation by holding more cash. This trade-off was examined when we studied the benchmark equilibrium. As  $X$  increases, the first term becomes negative. Mathematically, this is because  $\lambda^*(X)$  increases at a lower rate than  $X$ . The economic intuition is that if  $X$  is too high, it distorts the optimal trade-off examined in the benchmark equilibrium. The second term in Eq. (10) is new. It represents the effect on the value of the fund from increasing the survival window. This term is always positive: Increasing the amount of cash  $X$  in the fund increases the survival range  $[0, \lambda^*(X)]$ . Overall, the sum of the two terms equals zero for some interior  $X$ . We denote this unique interior solution as  $X^{**}$ .

**Theorem 5** *There exists a unique optimal amount of cash holdings in the hedge fund's balance sheet,  $X^{**}$ , in the interval  $(0, X^C)$  that maximizes the hedge fund value  $W^{II}(X)$ .*

<sup>15</sup>By the properties of  $\lambda^*(X)$ , the inequality  $\lambda^*(X) < X + (1 - X) \frac{\alpha(R-1)}{R-\alpha} < \bar{\lambda}$  is valid and, therefore, the upper limit of the survival window is  $\min[\lambda^*(X), X + (1 - X)\alpha, \bar{\lambda}] = \lambda^*(X)$ .

<sup>16</sup>Rigorously,  $\underset{X}{Max} W^{II}(X)$  can be written as  $\underset{X}{Sup} W^{II}(X)$ .

Proof: See the Appendix.

Next, we analyze the second region of  $X$ , where  $X \in [X^C, 1]$ . In this region, cash holdings are very high, and all late investors decide to stay with the fund. The expected value of the fund is, therefore, equal to that in the benchmark case:

$$(11) \quad W^{II}(X) = W^I(X) = \frac{1}{\bar{\lambda}} \left\{ \int_0^X [(X-\lambda) + (1-X)R] d\lambda + \int_X^{\bar{\lambda}} [(1-X) - \frac{\lambda-X}{\alpha}] R d\lambda \right\} \text{ for } X \in [X^C, 1].^{17}$$

Because  $\bar{\lambda} > \frac{R-\alpha}{R(2-\alpha)-1}$ , using the proof of Theorem 1, it is easy to show that  $W^{II}(X)$  is decreasing in the interval  $X \in [X^C, 1]$ . Therefore,  $W^{II}(X)$  achieves a maximum at  $X = X^C$ . We need to prove that this maximum value is less than that in the region  $X \in [0, X^C]$ , or  $W^{II}(X^C) < W^{II}(X^{**})$ . This is true under general parameter values. Therefore, it is too costly for the hedge fund to choose cash holdings in the second region.

**Theorem 6** *For general parameter values, we have that  $W^{II}(X^C) < W^{II}(X^{**})$ .<sup>18</sup> That is, it is not optimal for the hedge fund to hold cash in the interval  $X \in [X^C, 1]$ . Globally, the optimal amount of cash holdings is  $X^{**}$ .*

Proof: See the Appendix.

The intuition for Theorem 6 is simple. A fund that holds a large amount of cash is better able to persuade late investors to stay on, eliminating the coordination problem and increasing its chances of survival. This policy is too costly, though, because the fund forgoes valuable investment opportunities. It is better that the fund holds less cash. Globally, the optimal amount of cash holdings is still  $X^{**}$ , which lies in the first region.

We wish to prove that the optimal amount of cash when a run is possible is higher than the optimal amount of cash when there is no risk of a run, i.e.,  $X^{**} > X^*$ .

---

<sup>17</sup>Because  $X + (1-X)\frac{\alpha(R-1)}{R-\alpha} > \bar{\lambda}$ , the upper limit of the fund's survival window is  $\min[\bar{\lambda}, X + (1-X)\alpha] = \bar{\lambda}$ .

<sup>18</sup>The need for parameter assumptions here is because we made a weak assumption; that is, investors coordinate and choose the staying-strategy equilibrium for the second region of  $X$ . If investors continue to use the threshold equilibrium, Theorem 6 is true automatically and does not require any parameter assumptions.

We look at what happens if the fund keeps the original cash holdings  $X^*$ . Because  $\bar{\lambda} > \frac{R-\alpha}{R(2-\alpha)-1}$ , from Theorem 1,  $X^* + (1 - X^*)\alpha < \bar{\lambda}$ . Because  $X^* + (1 - X^*)\frac{\alpha(R-1)}{R-\alpha} < X^* + (1 - X^*)\alpha$ , we have that  $X^* + (1 - X^*)\frac{\alpha(R-1)}{R-\alpha} < \bar{\lambda}$ . Therefore, we conclude that  $X^*$  indeed lies in the first region. Furthermore, by the properties of  $\lambda^*(X)$ , we have that  $\lambda^*(X^*) < X^* + (1 - X^*)\alpha$ . That is, the effect of the run causes the survival range to shrink from  $[0, X^* + (1 - X^*)\alpha]$  to  $[0, \lambda^*(X^*)]$ .

Because an increase in  $X$  widens the survival range [i.e.,  $\lambda^*(X)$  increases in  $X$ ], the hedge fund may need to increase the optimal  $X^*$  in the benchmark equilibrium to achieve a wider survival range and thus maximize the fund value. We evaluate the first-order derivative of  $W^{II}(X)$  at  $X = X^*$  to check the change in  $W^{II}(X)$  when  $X$  is changed:

$$\frac{dW^{II}(X)}{dX}\Big|_{X=X^*} = \left[ \left(1 - \frac{R}{\alpha}\right)X + \frac{1-\alpha}{\alpha}\lambda^*(X)R \right]\Big|_{X=X^*} + \left\{ \left[ (X + (1 - X)\alpha - \lambda^*(X)) \cdot \frac{R}{\alpha} \cdot \frac{d\lambda^*(X)}{dx} \right] \right\}\Big|_{X=X^*}. \quad (12)$$

An increase in  $X$  at  $X = X^*$  has two effects on  $W^{II}(X)$ . The second term in Eq. (12), the dominant effect, is positive. It captures the gains from increasing the survival range. The first term of Eq. (12) is negative. The intuition for the first term is as follows. At the *effective* maximum shock,  $X^* + (1 - X^*)\alpha$ , the optimal amount of cash is  $X^*$ . Because the effective maximum shock decreases to  $\lambda^*(X^*)$ , the optimal cash holdings should be lower. However, this effect is second order compared with the first effect under general parameter values. That is, overall,  $\frac{dW^{II}(X)}{dX}\Big|_{X=X^*} > 0$ , meaning that the fund can increase its value at  $T_2$  by increasing its positions in cash.

**Theorem 7** *For general parameter values, we have that  $X^{**} > X^*$ .<sup>19</sup> That is, the optimal amount of cash holdings when a run is possible is higher than when a run is not taken into account.*

Proof: See the Appendix.

Theorem 7 is a key result of the paper. It shows how runs on hedge funds affect hedge funds' optimal cash holdings. It is important to emphasize that ex post runs impact the optimal amount

---

<sup>19</sup>The need for parameter assumptions is purely for technical reasons. In fact, the first term in Eq. (12) is negative due to  $\epsilon \rightarrow 0$ . In general cases, when  $\epsilon$  is not arbitrarily small, the first term is also positive. Thus, the inequality that  $\frac{dW^{II}(X)}{dX}\Big|_{X=X^*} > 0$  is always true; it does not depend on particular parameter values. That is,  $X^{**} > X^*$  holds in general if we do not focus on  $\epsilon \rightarrow 0$ .

of cash ex ante in two ways. The first one is the *response* effect, represented by the first term on the left-hand side of Eq. (10). Because of the risk of a run, redemptions include not only early investors but also some late investors. The size of total redemptions is thus changed because of runs. In response to an increase in the size of redemptions, the hedge fund manager adjusts the optimal cash holdings upward. The second channel is the *signaling* effect and is captured by the second term on the left-hand side of Eq. (10). The probability of the run and, hence, the size of total redemptions is a function of the hedge fund's cash holdings. The fund manager holds more cash to 'signal' and reassure investors, and thus to reduce the likelihood of a run.

Before closing this subsection, we illustrate the results of the model with a numerical example. We set the parameter values to  $R = 3$ ,  $\alpha = 0.8$ , and  $\bar{\lambda} = 0.9$ . With these values, the optimal amount of cash holdings without considering a run is  $X^* = 0.23$ . The hedge fund survives as long as  $\lambda \in [0, X^* + (1 - X^*)\alpha] = [0, 0.85] \subset [0, 0.9]$ . With the cash holdings of  $X^* = 0.23$ , however, the effect of a run reduces the fund's survival range to  $\lambda \in [0, \lambda^*(X^*)] = [0, 0.59]$ . The fund value is  $W^{II}(X^*) = 1.27$ . Therefore, when a run is possible, it is optimal to increase the amount of cash holdings as this increases the range of values of  $\lambda$  at which the fund survives. In fact, at  $X^* = 0.23$ , the survival range of the fund changes at a rate of  $\frac{d\lambda^*(X)}{dX}|_{X=X^*} = 0.52$ , and the value of the fund changes at a rate  $\frac{dW^I(X)}{dX}|_{X=X^*} = 0.31$ . Therefore, the value of the fund can be increased by increasing the amount of cash holdings,  $X$ .

Indeed, if the cash holdings are increased to  $X^{**} = 0.34$ , the survival range widens to  $\lambda \in [0, \lambda^*(X^{**})] = [0, 0.64]$ . We have  $\frac{d\lambda^*(X)}{dX}|_{X=X^{**}} = 0.53$  and  $\frac{dW^{II}(X)}{dX}|_{X=X^{**}} = 0$ . Therefore, the fund achieves its highest expected value at  $X = X^{**}$ , where  $W^{II}(X^{**}) = 1.30$ . Note that  $X + (1 - X)\frac{\alpha(R-1)}{R-\alpha}|_{X=X^{**}} = 0.82 < \bar{\lambda}$  is true.

If the fund increases the cash holdings to  $X^C = 0.63$ , there is no coordination problem among investors. This provides the widest survival range of the fund,  $\lambda \in [0, \bar{\lambda}] = [0, 0.9]$ . But this policy is too costly and, therefore, it is not optimal. In fact, the maximum value of the fund within the region  $X \in [X^C, 1]$  is  $W^{II}(X^C) = 1.18$ , which is less than  $W^{II}(X^{**}) = 1.30$ .

Fig. 8 depicts how  $W^I(X)$  and  $W^{II}(X)$  change with  $X$ .  $W^I(X)$  is above  $W^{II}(X)$ , meaning that a run causes the value of the fund to go down.  $W^I(X)$  achieves its highest value at  $X = 0.23$ , while  $W^{II}(X)$  has a positive slope at  $X = 0.23$  and achieves a maximum at  $X = 0.34$ . For values of

$X$  above 0.63, there is no coordination problem any longer, and consequently  $W^I(X)$  and  $W^{II}(X)$  coincide. The optimal cash holdings without considering a run are  $X^* = 0.23$  and the optimal cash holdings when a run is possible are  $X^{**} = 0.34$ .

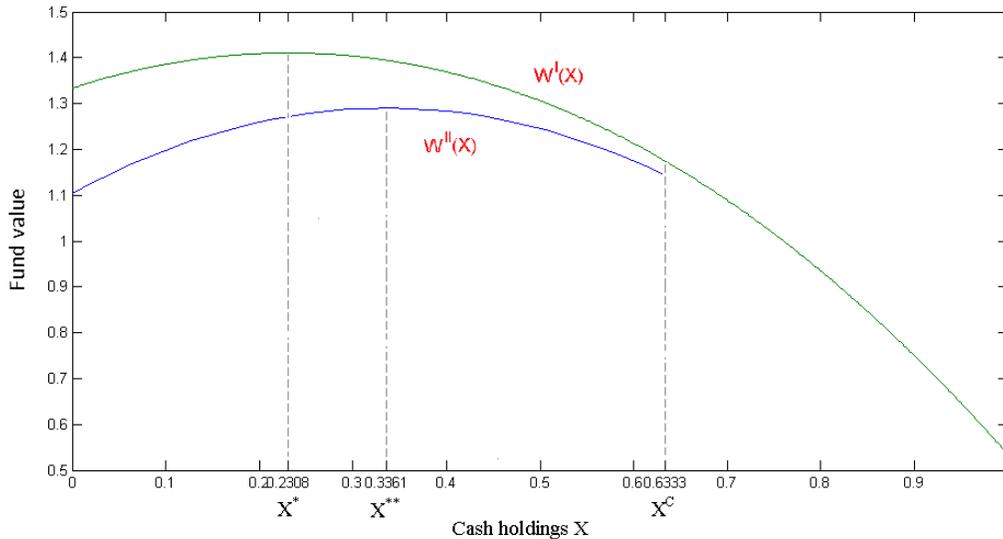


Fig. 8. The fund value as a function of cash holdings

## 2.3 Implications and predictions of the model

In this subsection, we conduct a comparative statics analysis and study the model's implications and predictions.

### 2.3.1 Implications of the model

From Theorem 7, a panic run forces a hedge fund to hold excess cash. This shows that coordination problems among investors induce a hedge fund to limit its exposure to illiquid assets.

Alternatively, the argument can be presented in the following way: Suppose that the reservation utility (opportunity cost) of the hedge fund manager is  $C$ . That is, only if the expected value of the fund at  $T_2$  is higher than  $C$  is the fund manager willing to participate in the market at  $T_0$ . We can prove that  $W^I(X^*, R)$  and  $W^{II}(X^{**}, R)$  are both increasing functions in the return,  $R$ . Because  $W^I(X^*, R) > W^{II}(X^{**}, R)$ , there exists an  $R^*$  and an  $R^{**}$  such that  $W^I(X^*, R^*) = C$  and  $W^{II}(X^{**}, R^{**}) = C$ , with  $R^* < R^{**}$ . Therefore, for returns  $R$  such that  $R^* < R < R^{**}$ ,

coordination problems prevent the hedge fund manager from taking profitable opportunities in illiquid positions. These lost opportunities in the interval  $[R^*, R^{**}]$  represent limits to arbitrage.

**Theorem 8** *The coordination risk leads hedge funds to give up participating in the market when the return  $R$  is moderately high (i.e.,  $R^* < R < R^{**}$ ).*

Proof: See the Appendix.

In Theorem 8, missing arbitrage opportunities ( $R^* < R < R^{**}$ ) occur not for the reason presented in Shleifer and Vishny (1997), but because of the coordination risk and the potential panic run on the hedge fund. Note that arbitrage opportunities are not taken only when returns are moderately high. When returns are very high, hedge funds still take arbitrage positions. In this sense, arbitrage activities are ‘limited’, not suspended completely, a result consistent with the empirical evidence.

Research by Mitchell and Pulvino (2011) and Krishnamurthy (2010) provides convincing evidence that, during the recent crisis, there were significant arbitrage opportunities in debt markets. Remarkably, Mitchell and Pulvino (2011) show that in many instances arbitrage failed because of hedge funds’ funding problems, instead of from problems related to the quality of the assets.

### 2.3.2 Predictions of the (ex post) probability of a run (at $T_1$ )

In our model, there are two factors that drive the probability of a run:  $\lambda$  and  $\alpha$ . From Eq. (8), we can derive the ex post probability of a run  $\Pr = \begin{cases} 0 & \text{if } \lambda \leq \lambda^*(X) \\ 1 & \text{if } \lambda > \lambda^*(X) \end{cases}$ .<sup>20</sup> It is easy to show that

$\frac{\partial \lambda^*(X)}{\partial \alpha} > 0$ . Therefore,  $\frac{\partial \Pr}{\partial \lambda} \geq 0$  and  $\frac{\partial \Pr}{\partial \alpha} \leq 0$ .

In our model,  $\lambda$  is the proportion of investors facing financial constraints, i.e., needing to withdraw money to cover for consumption, portfolio allocation, or other financial emergencies. Essentially, we interpret  $\lambda$  as the funding liquidity of hedge funds, i.e., the number of investors who are in need of withdrawing early to cover their own liquidity shocks. Further, a fall in  $\alpha$ , a sudden drying up of market liquidity, also increases the probability of a run.<sup>21</sup>

<sup>20</sup>Due to  $\epsilon \rightarrow 0$ , the function  $\Pr$  is discrete. In general,  $\Pr$  is continuous.

<sup>21</sup>See Brunnermeier and Pedersen (2009) for more discussions about market liquidity and funding liquidity.

We offer Prediction 1.

**Prediction 1:** Runs on hedge funds are more likely to occur either under poor market liquidity conditions or when large negative funding-liquidity shocks occur (i.e., a high  $\lambda$ ).

The impact of market liquidity is particularly relevant during a crisis, because markets typically become illiquid, often systemically across different asset classes.

### 2.3.3 Predictions of the (*ex ante*) optimal cash holdings

Using Eq. (8), we can derive the ex ante probability of a run:  $\overline{\text{Pr}} = 1 - \frac{\lambda^*(X)}{\lambda}$ . We have that  $\frac{\partial \overline{\text{Pr}}}{\partial \lambda} > 0$  and that  $\frac{\partial \overline{\text{Pr}}}{\partial \alpha} < 0$ . Thus, we have a time series prediction for hedge funds' cash holdings.

**Prediction 2 (time series of cash holdings):** Hedge funds increase their cash holdings when they expect either market liquidity or funding liquidity to worsen.

Also, because  $\frac{\partial \overline{\text{Pr}}}{\partial \alpha} < 0$ , we have a cross-sectional prediction.

**Prediction 3 (cross section of cash holdings):** Hedge funds that invest more in illiquid assets, and are consequently more likely to incur runs, are more likely to hold larger cash positions as a precaution.

## 3 Discussion

So far we have assumed that the entire capital of hedge funds comes from their investor-clients (equityholders), i.e., that funds carry no debt. While this is a gross simplification, it allows us to emphasize that the equity in hedge funds is fragile and subject to panic runs. This feature of hedge funds distinguishes them from investment banks and commercial banks, where a panic run cannot happen to the equity. In fact, equityholders of a bank cannot tender shares back to the bank for redemption but can only trade ownership in the secondary market.

In addition to equity, hedge funds use leverage, most of it provided by their prime-brokers. The debt of hedge funds is certainly susceptible to runs such as margin calls or sudden increases

in haircuts. Therefore, the whole right-hand side of the balance sheet of hedge funds is subject to runs. In particular, hedge fund runs triggered by debtholders or by equityholders can reinforce each other. Prime brokers calling the debt extended to a fund can scare the clients of the fund and trigger redemptions. The opposite is also true: A drop in the assets under management due to large scale redemptions induces prime brokers to impose stricter limits on leverage and reduce credit.

These arguments imply that hedge funds seem to have little or no capital in a strict sense, when capital is defined as “a long-term claim without a first-come-first-served right to cash flows” (see Diamond and Rajan, 2000). In other words, hedge funds have no real capital cushion on their balance sheets that can be used to mitigate a panic run by debtholders or equityholders. In contrast, banks have a capital cushion that can withstand the first loss if some debtholders start a run.

We argue that the fragile nature of the capital structure of hedge funds has critical implications both for individual hedge fund behavior and for capital markets.

### **3.1 *Capital fragility encourages micro-prudence***

Calomiris and Kahn (1991) and Diamond and Rajan (2001) argue that the coordination problem inherent in a bank run works as a commitment device, since the bank can commit not to engage in actions that dissipate value. Diamond and Rajan (2000) develop the argument further by adding that bank capital requirements sometimes can be harmful and reduce welfare, since they weaken the management’s commitment to behave appropriately by making the bank’s capital structure less fragile.

In the same spirit, we argue that hedge funds, with more fragile capital structures than banks, should be subject to greater market discipline. Tougher market discipline should lead to greater care in both investment and financing decisions.<sup>22</sup>

With respect to investment decisions, our paper formally shows that hedge funds optimally choose to hold more cash when a run is possible. Building on this logic, we expect that if a fund

---

<sup>22</sup>This conclusion is even stronger given that banks are protected by deposit insurance and hedge funds are not.

has some locked-up capital (i.e., some clients' funds are locked up until  $T_2$  and cannot be redeemed earlier), then the fund holds less cash and invests more in illiquid assets. To prove this formally, we adjust the original setup and assume that there is a tiny proportion ( $\delta$ ) of investors that is locked up. The proportion of investors that can run is thus  $1 - \lambda - \delta$ . We prove that the run threshold increases and, hence, that the optimal cash holdings decrease.

**Theorem 9** *If the capital that is locked up is positive, i.e.,  $\delta > 0$ , then the optimal amount of cash holdings,  $X^{**}$ , decreases (relative to when there is no capital that is locked up).*

Proof: See the Appendix.

Theorem 9 sheds some light on why hedge funds, with little or no locked-up capital, are less aggressive in trading illiquid assets than banks. The evidence shows that, as the recent crisis unfolded, hedge funds rebalanced their portfolios toward cash holdings more aggressively and faster than banks.

With respect to financing decisions, we can show that equity fragility makes hedge funds less aggressive than banks in the use of leverage, as long as the market is not perfectly liquid, i.e.,  $\alpha < 1$ . The idea is as follows. Any rollover problem in hedge funds' short-term debt can ignite a run not only by other debtholders but also by equityholders. Equityholders of banks, however, cannot run. A rollover problem by some lenders of a bank can trigger a run by other lenders but not by equityholders. The fact that equityholders cannot run provides a cushion for unconstrained lenders and reduces the incentive to run. This, in turn, encourages banks to operate with higher leverage ratios than hedge funds ex ante.<sup>23</sup>

Consistent with the above prediction, the data show that at the beginning of the crisis, on average, hedge funds operated with much lower leverage ratios than banks. We believe this evidence

---

<sup>23</sup>To see this, consider that both a hedge fund and a bank have the same debt-to-equity ratio and that the probability that the debt is rolled over is the same in both cases. Then, we can show that the ex post probability of a run triggered by the inability to roll over part of the debt is higher for the hedge fund than it is for the bank. Consequently, this encourages banks to operate with higher leverage ratios than hedge funds ex ante. This idea can be formalized with global game techniques. To do that, we need to explicitly model the payoff structure of both debtholders and equityholders, not that of equityholders alone as in the current model. Within such a framework, it is possible to model the optimal leverage ratio of financial institutions.

is in part explained by the argument that different degrees of capital fragility determine different leverage choices.

In sum, this article argues that hedge funds have strong incentives to run less risky portfolios and operate with lower leverage ratios because of their fragile capital structure. The relative conservatism of hedge funds helped them survive better in the recent financial crisis.

### ***3.2 Capital fragility limits arbitrage***

Our micro-prudence arguments have implications for the macro-soundness of capital markets. Despite their particular expertise, hedge funds are constrained by their fragile capital structure and at times are not able to fully exploit their trading capabilities. Liquidity risk effectively limits hedge funds' arbitrage activities with important implications for financial markets. Hedge funds must have spotted many interesting investment opportunities during the recent crisis. If they had been able to exploit these to the limit, markets would have been more informationally efficient. In reality, the opposite happened. Hedge funds drastically cut their exposures to risky and illiquid positions and, in so doing, they reduced their role as liquidity providers precisely when markets were in greater need of liquidity.

### ***3.3 Further extensions of the model***

The model can be extended to show that there is an additional feedback effect that can make things worse. Suppose that  $\alpha$ , the variable that captures market illiquidity, is a function of the total hedge fund capital invested in the market. So, if significant capital is invested by hedge funds in illiquid securities,  $\alpha$  is high; otherwise,  $\alpha$  is low. Once a crisis erupts, the tendency to hold cash among hedge funds increases and this feeds back into  $\alpha$ , making the crisis more severe. That is, there is an externality across hedge funds. Each hedge fund holds cash to protect itself against its own investors' coordination failures but, in taking care of its own problem, each hedge fund reduces market liquidity for all other hedge funds, in turn making investors' coordination failure more likely, necessitating even more cash hoarding, and so on.<sup>24</sup> In short, hoarding liquidity has spiraling implications that can exacerbate market paralysis during a crisis.

---

<sup>24</sup>See Boyson, Stahel, and Stulz (2010) for evidence along this line.

## 4 Concluding remarks

We have developed a model that highlights how the fragile nature of the capital structure in hedge funds combined with imperfect market liquidity hampers funds' ability to perform market arbitrage. In the absence of sophisticated investors, markets become less informationally efficient. In turn, less efficient markets reduce the participation of investors with spare funds. Limits to arbitrage can, therefore, cause a vicious cycle and prolong a financial crisis.

Our explanation for the limits of arbitrage differs from that in Shleifer and Vishny (1997). In their work, investors worry about the quality of a money manager's assets, and poor performance leads to early redemptions, which, in turn limits arbitrage. We instead emphasize that investors face a coordination problem with regard to short-term redemptions. The mere apprehension that other investors might pull out makes an individual investor consider pulling out as well. During a financial crisis, when market liquidity is low, a fire sale is costly and coordination problems are severe. These conditions contribute to the possibility of a panic run. It is this ex post potential panic run that makes hedge funds less willing to arbitrage markets ex ante.

Our model also helps explain one fact that seems to have surprised many observers of the recent financial crisis: Hedge funds showed a greater resilience than investment banks (and many commercial banks). We argue that the reason for this lies in part in the different nature of the financial claims in hedge funds relative to banks. Both the equity and the debt of hedge funds are fragile since they can be withdrawn on short notice by investors on a first-come, first-served basis. For a bank only the debt can be called; the equity is a nonredeemable long-term claim. The fragility of the whole liability side of the balance sheet of hedge funds implies that they are much more prone to runs. Hedge fund managers factor a possible run into their investment decisions, as well as into their financing decisions. The fragility induces hedge funds to be less aggressive than banks in investing in illiquid assets and in using leverage ex ante.<sup>25</sup> The relative conservatism of hedge funds helped them survive better in the recent financial crisis. An important implication of this work is that an unregulated sector of the financial markets can be safer than a regulated sector because of the particular nature of the financial contracts that exist in each case.

---

<sup>25</sup>In this paper, a market disciplining mechanism is used to explain why hedge funds behave more conservatively than banks. The moral hazard problem of "too big to fail", which encourages banks to take excessive risk ex ante, is perhaps another important factor behind banks' aggressiveness.

## Appendix

### A.1. Proof of Theorem 1

We can rewrite Eq. (3) as

$$W^I(X) = \begin{cases} \frac{1}{\bar{\lambda}} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_{X+(1-X)\alpha}^{\bar{\lambda}} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\} & \text{if } X + (1 - X)\alpha \geq \bar{\lambda} \\ \frac{1}{\bar{\lambda}} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_X^{X+(1-X)\alpha} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\} & \text{if } X + (1 - X)\alpha < \bar{\lambda} \end{cases} \quad (13)$$

Hence, we have

$$W^I(X) = \begin{cases} \frac{1}{\bar{\lambda}} \left[ \left( \frac{1}{2} - \frac{R}{2\alpha} \right) X^2 + \frac{1-\alpha}{\alpha} \bar{\lambda} R X + \left( \bar{\lambda} - \frac{\bar{\lambda}^2}{2\alpha} \right) R \right] & \text{if } X \geq \frac{\bar{\lambda} - \alpha}{1 - \alpha} \quad (14) \\ \frac{1}{\bar{\lambda}} \left[ \left( \frac{1}{2} - R + \frac{\alpha R}{2} \right) X^2 + R(1 - \alpha)X + \frac{\alpha R}{2} \right] & \text{if } X < \frac{\bar{\lambda} - \alpha}{1 - \alpha} \quad (15) \end{cases}$$

The first-order conditions of Eq. (14) and (15) are  $X^{1*} = \frac{(1-\alpha)\bar{\lambda}R}{R-\alpha}$  and  $X^{2*} = \frac{R(1-\alpha)}{R(2-\alpha)-1}$ , respectively.

Because  $X^{1*} - \frac{\bar{\lambda} - \alpha}{1 - \alpha} = \frac{(1-\alpha)\bar{\lambda}R}{R-\alpha} - \frac{\bar{\lambda} - \alpha}{1 - \alpha} = \frac{(R-\alpha) - [R(2-\alpha) - 1]\bar{\lambda}}{(R-\alpha)(1-\alpha)/\alpha}$  and  $X^{2*} - \frac{\bar{\lambda} - \alpha}{1 - \alpha} = \frac{R(1-\alpha)}{R(2-\alpha)-1} - \frac{\bar{\lambda} - \alpha}{1 - \alpha} = \frac{(R-\alpha) - [R(2-\alpha) - 1]\bar{\lambda}}{[R(2-\alpha) - 1](1-\alpha)}$ , the relation that  $(X^{1*} - \frac{\bar{\lambda} - \alpha}{1 - \alpha})(X^{2*} - \frac{\bar{\lambda} - \alpha}{1 - \alpha}) \geq 0$  is true. It means that  $X^{1*}$  and  $X^{2*}$  are on the same side of the vertical line  $X = \frac{\bar{\lambda} - \alpha}{1 - \alpha}$ .

If  $\bar{\lambda} \leq \frac{R-\alpha}{R(2-\alpha)-1}$ ,  $W^I(X)$  is an increasing function in the interval  $[0, \frac{\bar{\lambda} - \alpha}{1 - \alpha}]$ , while it increases first and then decreases in the interval  $[\frac{\bar{\lambda} - \alpha}{1 - \alpha}, 1]$ . Overall,  $W^I(X)$  achieves a maximum value at  $X^* = \frac{(1-\alpha)\bar{\lambda}R}{R-\alpha}$ , where  $X^* \in [\frac{\bar{\lambda} - \alpha}{1 - \alpha}, 1]$ .

If  $\bar{\lambda} \geq \frac{R-\alpha}{R(2-\alpha)-1}$ ,  $W^I(X)$  increases first and then decreases in the interval  $[0, \frac{\bar{\lambda} - \alpha}{1 - \alpha}]$ , while it decreases in the interval  $[\frac{\bar{\lambda} - \alpha}{1 - \alpha}, 1]$ . Overall,  $W^I(X)$  achieves a maximum value at  $X^* = \frac{R(1-\alpha)}{R(2-\alpha)-1}$ , where  $X^* \in (0, \frac{\bar{\lambda} - \alpha}{1 - \alpha}]$ .

### A.2. Proof of Theorem 3

As shown in Fig. A1, to guarantee that  $\lambda^*(X)$  is on the right of  $X$ , we need to make sure that the area of ADE (denoted as  $S_{ABC}$ ) is greater than the area of EFGH (denoted as  $S_{BDGF}$ ) for whatever  $X$ .

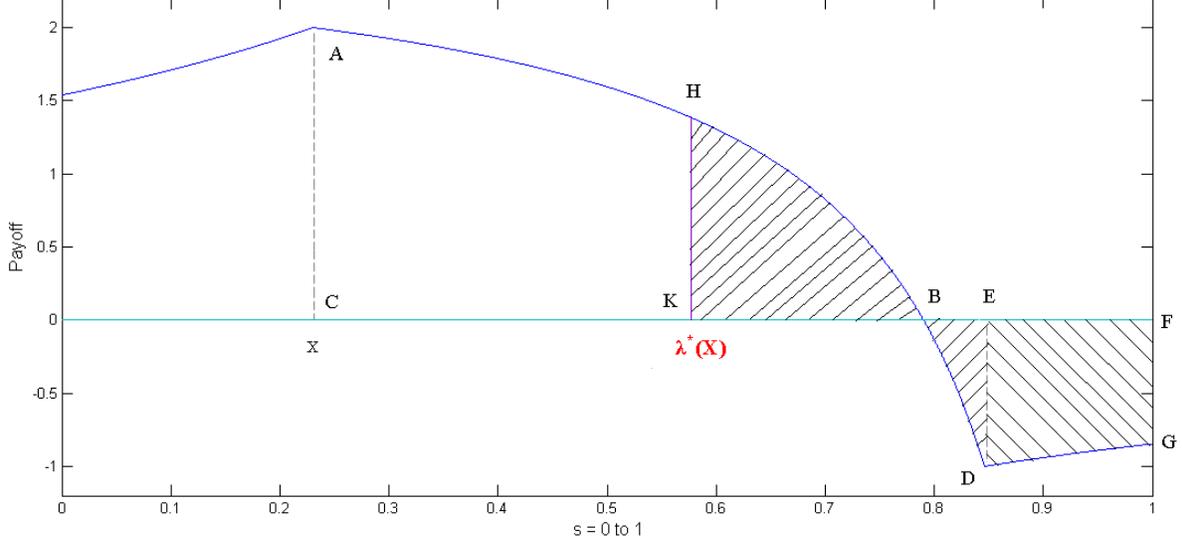


Fig. A1. The equilibrium illustration

It is easy to show that  $S_{ABC} - S_{BDE} = \int_X^{X+(1-X)\alpha} \left[ \frac{(1-X) - \frac{s-X}{\alpha}}{1-s} R - 1 \right] ds = (1-X) \left[ \frac{R}{\alpha} (1-\alpha) \ln(1-\alpha) + (R-\alpha) \right]$  and  $S_{DEFG} = \int_{X+(1-X)\alpha}^1 \frac{X+(1-X)\alpha}{s} ds = -[X+(1-X)\alpha] \ln[X+(1-X)\alpha] = (1-X)(1-\alpha) \left( \frac{-t \ln t}{1-t} \Big|_{t=X+(1-X)\alpha} \right)$ .

Therefore, we can work out the difference between the two areas:

$$\begin{aligned} \Delta S &= S_{ABC} - S_{BDGF} \\ &= S_{ABC} - (S_{BDE} + S_{DEFG}) \\ &= (1-X) \left\{ \left[ \frac{R}{\alpha} (1-\alpha) \ln(1-\alpha) + (R-\alpha) \right] - (1-\alpha) \left( \frac{-t \ln t}{1-t} \Big|_{t=X+(1-X)\alpha} \right) \right\}. \end{aligned} \quad (16)$$

Note that  $\frac{-t \ln t}{1-t}$  is increasing in the interval  $[0, 1)$ ,  $\frac{-t \ln t}{1-t} \Big|_{t=0} = 0$ , and  $\lim_{t \rightarrow 1} \frac{-t \ln t}{1-t} = 1$ . To guarantee that the area difference  $\Delta S$  is positive for whatever  $X$  is to make sure that  $\frac{R}{\alpha} (1-\alpha) \ln(1-\alpha) + (R-\alpha) > (1-\alpha) \left( \frac{-t \ln t}{1-t} \Big|_{t=1} \right)$ , that is,  $\frac{R}{\alpha} (1-\alpha) \ln(1-\alpha) + (R-\alpha) > 1-\alpha$ .

Also, we need  $\bar{\lambda}$  to be big enough so that  $\bar{\lambda} > \lambda^*(X)$ . The condition  $\bar{\lambda} > X + (1-X) \frac{\alpha(R-1)}{R-\alpha}$  suffices for this because  $X + (1-X) \frac{\alpha(R-1)}{R-\alpha} > \lambda^*(X)$  is definitely true by the properties of  $\lambda^*(X)$ .

### A.3. Proof of Theorem 4

We again use Fig. A1 to illustrate the proof.

We express  $\lambda^*(X)$  as  $\lambda^*(X) = X + b(1 - X)$ . Suppose  $X$  changes to  $X + \Delta X$ . We want to compare  $\lambda^*(X + \Delta X)$  with  $\lambda^*(X)$ . The idea of comparison is as follows. After  $X$  changes to  $X + \Delta X$ , the shape of the payoff curve of course changes because the parameter becomes  $X + \Delta X$  instead of  $X$ . We want to know under the new shape of curve whether the area difference  $S_{BKH} - S_{BDGF}$  becomes negative or positive if we still keep the threshold as  $\lambda^*(X)$ . If  $\Delta S$  becomes positive, it means that the above shaded area is bigger than the below one, that we need to remove part of the above shaded area to make the two shaded areas equal, and consequently that  $\lambda^*(X + \Delta X) > \lambda^*(X)$ .

We check how the area difference  $S_{BKH} - S_{BDGF}$  changes in two steps.

In the first step, we ask what the area difference  $S_{BKH} - S_{BDGF}$  under the new curve is if we keep  $b$  unchanged. That is, we want to find out the area difference  $S_{BKH} - S_{BDGF}$  under the new curve if the threshold is  $(X + \Delta X) + b[1 - (X + \Delta X)]$ .

It is easy to show  $S_{HBK} - S_{BED} = (1 - X)\{\frac{R}{\alpha}[(\alpha - \beta) + (1 - \alpha)\ln\frac{1-\alpha}{1-b}] - (\alpha - b)\}$  and  $S_{EFGD} = -[X + (1 - X)\alpha]\ln[X + (1 - X)\alpha] = (1 - X)(1 - \alpha)(\frac{-t\ln t}{1-t}|_{t=X+(1-X)\alpha})$ .

Because  $S_{HBK} - S_{BED} = S_{EFGD}$ , we have

$$\frac{R}{\alpha}[(\alpha - \beta) + (1 - \alpha)\ln\frac{1-\alpha}{1-b}] - (\alpha - b) = (1 - \alpha)(\frac{-t\ln t}{1-t}|_{t=X+(1-X)\alpha}). \quad (17)$$

Therefore,  $\frac{\partial(S_{HBK} - S_{BED})}{\partial X} = -\frac{R}{\alpha}[(\alpha - \beta) + (1 - \alpha)\ln\frac{1-\alpha}{1-b}] + (\alpha - b) = -(1 - \alpha)(\frac{-t\ln t}{1-t}|_{t=X+(1-X)\alpha})$ .

Also, we can work out  $\frac{\partial(S_{EFGD})}{\partial X} = -(1 - \alpha)(1 + \ln t)|_{t=X+(1-X)\alpha}$ .

Overall,  $\frac{\partial(S_{BKH} - S_{BDGF})}{\partial X} = -(1 - \alpha)[\frac{-t\ln t}{1-t} - (1 + \ln t)]|_{t=X+(1-X)\alpha}$ .

In the second step, we work out under the new curve how much the above shaded area  $S_{HBK}$  increases if the threshold reduces from  $(X + \Delta X) + b[1 - (X + \Delta X)]$  to  $X + b(1 - X)$ . It is easy to show that the area increases by  $(\Delta X)(1 - b)(\frac{\alpha - b}{1 - b} \cdot \frac{R}{\alpha} - 1)$ .

Aggregating the results in the above two steps, if the curve shape is changed with the new parameter  $X + \Delta X$  while the threshold is still  $\lambda^*(X)$ , the area difference  $S_{BKH} - S_{BDGF}$  becomes

$$(\Delta X)\{(1 - b)(\frac{\alpha - b}{1 - b} \cdot \frac{R}{\alpha} - 1) - (1 - \alpha)[\frac{-t\ln t}{1-t} - (1 + \ln t)]|_{t=X+(1-X)\alpha}\}. \quad (18)$$

Because  $-(1-\alpha)[\frac{-t \ln t}{1-t} - (1 + \ln t)]$  is positive for any  $t \in (0, 1)$ , we conclude that  $b$  must be decreasing with  $X$ . Hence the first term of Eq. (18) is increasing with  $X$ . Also, it is easy to check that the second term  $(1-\alpha)[\frac{-t \ln t}{1-t} - (1 + \ln t)]|_{t=X+(1-X)\alpha}$  is a decreasing function with respect to  $X$ . Therefore, the whole term of Eq. (18) is increasing with  $X$ . To guarantee that Eq. (18) is positive for any  $X \in (0, 1)$ , we need to have its value at  $X = 0$  to be positive. That is,  $(\alpha - b) \cdot \frac{R}{\alpha} - (1 - b) > -\alpha \ln \alpha - (1 - \alpha)(1 + \ln \alpha)$ , where  $b$  solves  $(\frac{R}{\alpha} - 1)(\alpha - b) + \frac{R}{\alpha}(1 - \alpha) \ln(\frac{1-\alpha}{1-b}) = -\alpha \ln \alpha$ . Alternatively, the above inequality can be simplified as  $\frac{R}{\alpha} \ln(\frac{1-\alpha}{1-b}) < \ln \alpha$ .

In sum,  $\lambda^*(X)$  is an increasing function with respect to  $X$  if parameters satisfy  $\frac{R}{\alpha} \ln(\frac{1-\alpha}{1-b}) < \ln \alpha$  (where  $b$  solves  $(\frac{R}{\alpha} - 1)(\alpha - b) + \frac{R}{\alpha}(1 - \alpha) \ln(\frac{1-\alpha}{1-b}) = -\alpha \ln \alpha$ ).

#### A.4. Proof of Theorem 5

Again, we express  $\lambda^*(X)$  as  $\lambda^*(X) = X + b(1 - X)$ . We have  $\frac{\partial \lambda^*(X)}{\partial X} = (1 - b) + (1 - X) \frac{\partial b}{\partial X}$ .

By Eq. (17), we have  $\frac{R}{\alpha}[(\alpha - \beta) + (1 - \alpha) \ln \frac{1-\alpha}{1-b}] - (\alpha - b) = (1 - \alpha)(\frac{-t \ln t}{1-t})|_{t=X+(1-X)\alpha}$ .

Differentiating on both sides of the above equation, we obtain  $\frac{R}{\alpha}(-\frac{\partial b}{\partial X} + \frac{1-\alpha}{1-b} \frac{\partial b}{\partial X}) + \frac{\partial b}{\partial X} = (1 - \alpha)^2 \frac{-(\ln t) - 1 + t}{(1-t)^2} |_{t=X+(1-X)\alpha}$ .

Thus,  $\frac{\partial b}{\partial X} = (1 - \alpha)^2 \frac{\frac{-(\ln t) - 1 + t}{(1-t)^2}}{-\frac{R}{\alpha} + 1 + \frac{R}{\alpha} \frac{1-\alpha}{1-b}} |_{t=X+(1-X)\alpha}$ .

Therefore,  $\frac{\partial \lambda^*(X)}{\partial X} = (1 - b) + (1 - X)(1 - \alpha)^2 \frac{\frac{-(\ln t) - 1 + t}{(1-t)^2}}{-\frac{R}{\alpha} + 1 + \frac{R}{\alpha} \frac{1-\alpha}{1-b}} |_{t=X+(1-X)\alpha}$ .

After we work out the explicit expression of  $\frac{\partial \lambda^*(X)}{\partial X}$ , we can also explicitly express  $\frac{dW^{II}(X)}{dX}$ , that is,

$$\frac{dW^{II}(X)}{dX} = [(1 - \frac{R}{\alpha})X + \frac{1-\alpha}{\alpha} \lambda^*(X)R] + [(X + (1 - X)\alpha - \lambda^*(X)) \cdot \frac{R}{\alpha} \cdot [(1 - b) + (1 - X)(1 - \alpha)^2 \frac{\frac{-(\ln t) - 1 + t}{(1-t)^2}}{-\frac{R}{\alpha} + 1 + \frac{R}{\alpha} \frac{1-\alpha}{1-b}} |_{t=X+(1-X)\alpha}]. \quad (19)$$

Based on Eq. (19), we can directly evaluate  $\frac{dW^{II}(X)}{dX}$  at any  $X$ .

It is true that  $\frac{dW^{II}(X)}{dX}|_{X=0} > 0$  and  $\frac{dW^{II}(X)}{dX}|_{X=X^c} < 0$  under general parameter values, which means that there is an interior solution of  $X^{**}$ .

#### A.5. Proof of Theorem 6

Because we know the expression of  $W^{II}(X)$ , the parameter conditions are the ones such that  $W^{II}(X^C) < W^{II}(X^{**})$ . Such parameters do exist. The numerical example in the text is one case.

#### A.6. Proof of Theorem 7

Based on Eq. (19), we can directly evaluate  $\frac{dW^{II}(X)}{dX}|_{X=X^*}$ . Therefore, the parameter conditions are ones such that  $\frac{dW^{II}(X)}{dX}|_{X=X^*} > 0$ .

#### A.7. Proof of Theorem 8

We need to prove only that  $W^I(X^*, R)$  and  $W^{II}(X^{**}, R)$  are both increasing functions of  $R$ . For

$$W^I(X) = \begin{cases} \frac{1}{\lambda} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_{\bar{\lambda}}^{\bar{\lambda}} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\} & \text{if } X + (1 - X)\alpha \geq \bar{\lambda} \\ \frac{1}{\lambda} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_X^{X + (1 - X)\alpha} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\} & \text{if } X + (1 - X)\alpha < \bar{\lambda} \end{cases},$$

both  $(X - \lambda) + (1 - X)R$  and  $(1 - X) - \frac{\lambda - X}{\alpha}]R$  are increasing in  $R$ . Therefore, when  $R$  increases,  $W^I(X^*, R)$  increases even if we do not adjust  $X^*$ . But, in fact, when  $R$  increases,  $X^*$  changes as well in maximizing  $W^I(X)$ . Therefore,  $W^I(X^*, R)$  certainly increases in  $R$ .

For  $W^{II}(X) = \frac{1}{\lambda} \left\{ \int_0^X [(X - \lambda) + (1 - X)R] d\lambda + \int_X^{\lambda^*(X)} [(1 - X) - \frac{\lambda - X}{\alpha}] R d\lambda \right\}$ , we need to prove that  $\lambda^*(X)$  is increasing in  $R$  first. In Fig. A1, as  $R$  increases,  $S_{EFGD}$  is unchanged while  $S_{BED}$  decreases. So, if keeping  $\lambda^*(X)$  unchanged,  $S_{BKH} - S_{BDGF} > 0$ . Therefore, to restore the equilibrium,  $\lambda^*(X)$  has to increase. After we prove that  $\lambda^*(X)$  is increasing in  $R$ , we use the argument similar to that in the case of  $W^I(X)$ , and we can prove that  $W^{II}(X^{**}, R)$  is increasing in  $R$ .

#### A.8. Proof of Theorem 9

In Fig. 7, the existence of a tiny proportion ( $\delta$ ) of locked investors is equivalent to cutting a  $\delta$ -width block at the right end of the below shaded area. So the below shaded area is reduced. To make the two shaded areas equal, the above shaded area should be reduced as well. So the threshold  $\lambda^*(X)$  has to increase. That is, for the same level of  $X$ , the survival window of the fund is bigger if there is locked capital than that if there is not. Therefore, in the first-order condition of Eq. (9), if we use the new increased threshold  $\lambda^*(X)$  to substitute for the original  $\lambda^*(X)$  while

keeping  $X$  the same, the first-order derivative becomes negative. That is, as the threshold  $\lambda^*(X)$  increases, the optimal cash holdings become less.

## References

- Abreu, D., Brunnermeier, M.K., 2002. Synchronization risk and delayed arbitrage. *Journal of Financial Economics* 66 (2–3), 341–600.
- Abreu, D., Brunnermeier, M.K., 2003. Bubbles and crashes. *Econometrica* 71 (1), 173–204.
- Allen, F., Gale, D., 2000. Financial contagion. *Journal of Political Economy* 108, 1–33.
- Boyson, N.M., Stahel, C.W., Stulz, R.M., 2010. Hedge Fund Contagion and Liquidity Shocks. *Journal of Finance* 65 (5), 1789–1816.
- Brunnermeier, M.K., 2009. Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives* 23 (1), 77–100.
- Brunnermeier, M.K., Pedersen, L.H., 2009. Market liquidity and funding liquidity. *Review of Financial Studies* 22 (6), 2201–2199.
- Calomiris, C., Kahn, C., 1991. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review* 81, 497–513.
- Chen, Q., Goldstein, I., Jiang, W., 2010. Payoff complementarities and financial fragility: evidence from mutual fund outflows. *Journal of Financial Economics* 97, 239–262.
- Corsetti, G., Dasgupta, A., Morris S., Shin, H.S., 2004. Does one soros make a difference? A theory of currency crises with large and small traders. *Review of Economic Studies* 71 (1), 87–113.
- Dasgupta, A., 2004. Financial contagion through capital connections: a model of the origin and spread of bank panics. *Journal of the European Economic Association* 2, 1049–1084.
- Diamond, D., Dybvig, P., 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91, 401–419.
- Diamond, D., Rajan, R., 2000. A theory of bank capital. *Journal of Finance* 55 (6), 2431–2465.
- Diamond, D., Rajan, R., 2001. Liquidity risk, liquidity creation, and financial fragility: a theory of banking. *Journal of Political Economy* 109, 287–327.
- Goldstein, I., Pauzner, A., 2004. Contagion of self-fulfilling financial crises due to diversification of investment portfolios. *Journal of Economic Theory* 119, 151–183.
- Goldstein, I., Pauzner, A., 2005. Demand deposit contracts and the probability of bank runs. *Journal of Finance* 60, 1293–1328.

Gorton, G., 2008. The panic of 2007. Federal Reserve Bank of Kansas City symposium. <http://www.kc.frb.org/publicat/sympos/2008/Gorton.10.04.08.pdf>.

Gorton, G., Winton, A., 2003. Financial intermediation. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *The Handbook of the Economics of Finance*. North Holland, Amsterdam, pp 431–552.

Gromb, D., Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics* 66, 361–407.

Johnson, W. T., 2006. Who monitors the mutual fund manager, new or old shareholders? <http://ssrn.com/abstract=687547>.

Krishnamurthy, A., 2010. How debt markets have malfunctioned in the crisis. *Journal of Economic Perspectives*, 24 (1), 3–28.

Mitchell, M., Pulvino, T., 2011. Arbitrage crashes and the speed of capital. *Journal of Financial Economics*, forthcoming.

Morris, S., Shin, H.S., 1998. Unique equilibrium in a model of self-Fulfilling currency attacks. *American Economic Review* 88, 587–597.

Morris, S., Shin, H.S., 2003. Global games: theory and application. In: Dewatripont, M., Hansen, L., Turnovsky, S. (Eds.), *Advances in Economics and Econometrics*. Cambridge University Press, Cambridge, England, pp. 56–114.

Morris, S., Shin, H.S., 2004. Liquidity black holes. *Review of Finance* 8, 1–18.

Morris, S., Shin, H.S., 2008. Financial Regulation in a System Context. Paper prepared for the Brookings Papers conference.

Morris, S., Shin, H.S., 2009. Illiquidity Component of Credit Risk. Unpublished working paper. Princeton University, Princeton, NJ.

Plantin, G., 2009. Learning by holding and liquidity. *Review of Economic Studies* 76, 395–412.

Plantin, G., Sapra, H., Shin, H.S., 2008. Marking-to-market: panacea or pandora's box? *Journal of Accounting Research* 46, 461–466.

Rochet, J.-C., Vives, X., 2004. Coordination failures and the lender of last resort: was Bagehot right after all? *Journal of the European Economic Association* 2, 1116–1147.

Shleifer, A., Vishny, R., 1997. The limits of arbitrage. *Journal of Finance* 52, 35–55.

Stein, J., 2005. Why are most funds open-end? Competition and the limits of arbitrage. *Quarterly Journal of Economics* 120, 247–272.