Preventing Runs with Fees and Gates

Lukas Voellmy*

November 29, 2020

Abstract

I study whether redemption fees and gates as recently introduced into US and EU money market fund (MMF) regulations achieve their goal of eliminating the first-mover advantage, thereby making MMFs less susceptible to runs. I focus on purely self-fulfilling runs in a setting without fundamental risk, using a version of the Engineer (1989) model, which itself corresponds to a Diamond and Dybvig (1983) model with one additional time period. The results suggest that, when used correctly, redemption fees are a versatile and effective tool to eliminate the first-mover advantage. Notably, redemption fees have two distinct roles in preventing runs: (i) they ensure that redeeming investors internalize the cost of liquidating assets and (ii) they can be used to implement a redistribution among investors that incentivizes investors to remain in the MMF when others run. Given that MMFs can charge fees, gates only have a role in preventing runs if there are (regulatory) restrictions on the level of the fee that MMFs can charge.

JEL-Codes: G01, G18

Keywords: Bank Runs, Redemption Fees, Suspension of Convertibility

*Swiss National Bank. E-mail: lukas.voellmy@snb.ch

I thank Öslem Dursun-de Neef (discussant), Merwan Engineer, Todd Keister, Igor Letina, Cyril Monnet, John Moore, Klaus Neusser and Ernst-Ludwig von Thadden for helpful comments on various versions of the paper. I also thank two anonymous referees and an associate editor. All errors are mine. The views, opinions, findings, and conclusions or recommendations expressed in this paper are strictly those of the author. They do not necessarily reflect the views of the Swiss National Bank (SNB). The SNB takes no responsibility for any errors or omissions in, or for the correctness of, the information contained in this paper.

1. Introduction

Both in the United States and the European Union, recent regulatory reforms allow/mandate money market funds (MMFs) to charge redemption fees or suspend redemptions (also referred to as 'gates') if they run low on liquid assets. The goal of these measures is to make MMFs less susceptible to runs by eliminating the first-mover advantage in redemptions.[1] I provide a summary of the new regulatory provisions further below in this section.

In this paper, I analyze theoretically how (and whether) one can eliminate the first-mover advantage with fees and gates. I do this by using a version of the Engineer (1989) model, which itself is essentially Diamond and Dybvig (1983) model with an additional time period. The Engineer setting is well suited to study the effect of liquidity management tools (such as fees and gates) on the propensity to run because it allows for the realistic scenario that the liquidity management tools may itself be the source of runs. In the Engineer setting, investors may withdraw preemptively from a fund out of fear that the fund might impose a fee or a gate precisely at that point in time in the future when the investor needs her money. As an example, gates (suspension of convertibilty) eliminate runs in a Diamond Dybvig model without fundamental risk but the same is not true for the Engineer model (see Engineer (1989)).

In most traditional models of self-fulfilling runs (on banks or funds), the runs occur because redemptions lead to liquidation losses which are borne by investors who do not redeem. Fees and gates are two means to ensure that redemptions do not impose losses on investors who remain in the fund. Fees allow to impose liquidation losses on the investors who redeem rather than those who stay in the fund, and gates can stop any redemptions in situations where further redemptions would be very costly for the fund. However, in the Engineer model, ensuring that redemptions do not impose liquidation losses on investors who remain in the fund is *not* sufficient to eliminate runs. Fees and gates must additionally be employed in such a way that they do not cause investors to withdraw preemptively for fear of being hurt by a fee or a gate when remaining in the fund.

[1]The impetus for the reforms came from the financial crisis of 2007/08. Most notably, US money market mutual funds experienced a modern-day bank run in September 2008 - see Schmidt et al. (2016) for a detailed account of the episode.

I study runs on a single intermediary (MMF or simply 'fund') that distributes a consumption good to a continuum of investors. There are three time periods (date 1, 2 and 3) and, following Engineer (1989), there are three types of investors: those who want to consume most at date 1, date 2 and date 3 respectively. Liquidity needs are revealed in a staggered fashion. At date 1, investors learn (privately) whether they need to consume at date 1 or not. Those who do not want to consume at date 1 learn only at date 2 whether they want to consume at date 2 or date 3. When paying out investors, the MMF is subject to a sequential service constraint, meaning that investors arrive at the fund sequentially and need to be served on the spot. This captures the fact that MMFs generally allow investors to redeem at very short notice. I abstract from fundamental risk, that is, the MMF's assets pay a fixed amount at the final date (date 3) and investors' aggregate liquidity needs are known. If investors withdraw before the final date (i.e. at dates 1 or 2) the MMF needs to sell assets on a secondary market in order to pay out the redeeming investors. In every period, the secondary market can absorb a certain amount of assets at fundamental value. Once asset sales within a given period reach a certain level, the asset price drops below fundamental value ('fire sale') before it recovers next period. I consider both the case of a stable NAV and a floating NAV fund.[2] In a stable NAV fund, the amount that investors can withdraw is fixed and does not move with changes in the secondary market price of the fund's asset. In a floating NAV fund, the amount that investors can withdraw tracks in real-time the secondary market price of the MMF's asset.

As is often the case with intermediaries offering liabilities that are redeemable on demand, our model-MMF may be subject to a first-mover advantage which makes it susceptible to self-fulfilling runs. Indeed, if the MMF operates with a stable NAV, the use of liquidity management tools (fees or gates) once redemptions reach a certain level is necessary to prevent runs, unless the asset is perfectly liquid (i.e. the price never drops below fundamental value). Without fees or gates, redemptions at a fixed NAV impose liquidation losses on those who remain in the fund, leading to a run equilibrium à-la Diamond and Dybvig. In contrast, a floating NAV fund may not be susceptible to runs even without using liquidity management tools. The reason is that changes in the NAV incorporate the cost of liquidating assets and thus ensure that liquidation costs are borne by the

[2]In the US, the fees and gates provisions are geared towards floating NAV funds while in the EU they are geared towards stable NAV funds - see the discussion at the end of this section.

investors who redeem instead of those that remain in the fund. However, contrary to conventional wisdom, a floating NAV does not always eliminate run equilibria even if changes in the NAV fully account for the cost of paying out redeeming investors. The reason is that investors may withdraw preemptively out of fear that they might need to consume precisely at that point in time in the future when the NAV will be temporarily depressed. Roughly speaking, floating NAV funds that do not use liquidity management tools will be susceptible to runs whenever the MMF's assets are relatively illiquid, such that heavy redemptions in a given period cause a large (temporary) decrease in NAV.

Having established under which conditions an MMF (without fees or gates) will be susceptible to runs, I proceed to analyse how fees and gates can be used to eliminate run equilibria. I show that, conceptually, redemption fees can have two roles in preventing runs. First, they are a means to ensure that those who redeem bear the liquidation costs which the fund incurs by paying them out. This role of fees is relevant for stable NAV funds but not for floating NAV funds, since changes in the NAV already account for liquidation costs. Second, redemption fees are a means to implement a redistribution among investors in the fund. Specifically, the fee revenue collected from redeeming investors can be redistributed to those who remain in the fund. Redemption fees can be used as a redistributionary tool in such a way as to give investors an active incentive to remain in the fund precisely when others run. This role of redemption fees is relevant for both stable and floating NAV funds. I show how (at least in the setting studied here) both stable and floating MMFs can always prevent runs by charging properly calibrated redemption fees. In the case of stable NAV funds, this may necessitate making use of both functions of redemption fees as outlined above.

The fact that redemption fees alone are sufficient to eliminate runs raises questions about the role of gates in preventing runs. Gates are less versatile than redemption fees and as long as there are no regulatory constraints on the level of the redemption fee, anything that can be achieved with gates can also be achieved with fees. Gates can have a role in preventing runs at stable NAV funds if there is a regulatory upper bound on the redemption fee, notably when the upper bound prevents funds from setting the fee to a level that ensures that redeeming investors internalize liquidation costs. I cannot find a rationale to use gates in floating NAV funds, even in cases where there are regulatory (or other) constraints on the level of redemption fees. Furthermore, it is easy to come up with examples where gates *create* runs on floating NAV funds that would not be susceptible to runs without the use of liquidity management tools.

Institutional Background Money market funds issue liabilities which are redeemable on demand and promise a high degree of principal stability, providing investors with an alternative to traditional bank deposits. In recent years, both the US and the EU adopted new regulations with the goal of making MMFs less susceptible to runs. An important part of these reforms is to allow/mandate MMFs to impose redemption fees or gates under certain conditions. The US regulations adopted in 2014 allow (but do not mandate) US money market funds to impose fees or gates if (and only if) their liquid assets drop below a certain threshold. By default, the fees and gates provisions apply to floating NAV funds; stable NAV funds can voluntarily opt into them. In Europe, fees and gates provisions are part of new MMF regulations adopted in 2017 and, different to the US, they apply only to MMFs with stable NAV or 'Low Volatility NAV' (LVNAV) but not to floating NAVs. LVNAVs are allowed to maintain a fixed NAV as long as the deviation of the fixed NAV from the market value of the portfolio is not too large; if the deviation becomes too large they need to adjust the NAV. As in the US, EU MMFs can impose fees or gates once their liquid assets fall below a certain threshold.[3] Different to the US, EU MMFs are obliged to impose either a fee or a gate if liquid assets reach a very low level. Besides the fees and gates clauses, another important part of both the US and EU regulations is to restrict the use of stable NAV to MMFs that invest uniquely in government assets. However, EU MMFs invested in private assets can still offer something very close to stable NAV by operating as LVNAV. I refer to SEC (2014) and Morgan Stanley (2019) for more detailed summaries of the new MMF regulations in the US and the EU respectively.

Related Literature This paper is part of a (small) theoretical literature examining the effect of redemption fees and/or gates on investors' propensity to run on a bank or a fund. Cipriani et al. (2014) and Lenkey and Song (2016) study settings in which the fundamental return to the bank/fund's investment is uncertain, with some depositors being better informed about the true return than others. Lenkey and Song (2016) show that fees may either increase or decrease uninformed investors' propensity to run if they observe large withdrawals by other (informed) investors.

[3] In some cases, activation of fees and gates require additionally that daily net redemptions exceed a certain threshold.

Cipriani et al. (2014) demonstrate how redemption fees can cause informed investors to run preemptively. Different to Cipriani et al. (2014), preemptive runs in the present paper are unrelated to fundamental risk and can be eliminated by calibrating redemption fees in the right way. Ennis and Keister (2009a, 2010) show that imposing gates may not be optimal ex-post and the effectiveness of gates in preventing runs is severely limited if the bank or the regulatory authority cannot credibly commit to take measures that hurt depositors ex-post. While the papers above focus on the role of information asymmetries and limited commitment, the present paper studies how staggered revelation of liquidity needs over time affects investors' propensity to run under fees and gates. In this regard, the paper is also related to He and Xiong (2012) who show how staggered debt maturity (instead of staggered revelation of liquidity needs as here) can lead to run equilibria that extend over several periods, as do the run equilibria studied in this paper.

This paper also belongs to a literature that studies panic equilibria in Diamond Dybvig settings with riskfree investment returns. This literature has focused on the case with two types of depositors ('patient' and 'impatient') and stochastic aggregate liquidity needs. Classical papers in this tradition are Green and Lin (2003), Peck and Shell (2003) and Ennis and Keister (2009b). In this class of models with sequential service and unknown aggregate liquidity needs, the efficient payout schedule usually features decreasing payouts. One way to implement the efficient payout schedule would be to charge (progressively increasing) redemption fees. The role of the decreasing payout schedule (often referred to as *partial suspension*) in these models is fundamentally different than in the present paper however; in the present paper, redemption fees are a means to prevent multiple equilibria and, in the efficient allocation, are only charged off the equilibrium path.[4]

Finally, this paper contributes to a theoretical literature studying self-fulfilling runs on funds with floating NAV. Conventional thinking often suggests that a first-mover advantage is only present in stable NAV funds but not in floating NAV funds, which was one of the main reasons why the use of stable NAV in MMFs has been restricted by regulators (see above). Zeng (2017) already showed that floating NAV does not necessarily eliminate the first-mover advantage, since redemptions today may cause the fund to liquidate assets in future periods in order to replenish its cash reserves, and these liquidation costs will be borne by investors who remain in the fund. Relatedly, Chen et al.

[4] I thank an anonymous referee for pointing out this difference.

(2010) present a model where redemptions at the end-of-day NAV cause funds to execute costly asset sales in future periods. Compared to these papers, the present paper shows that floating NAV funds can be prone to runs even if changes in NAV track liquidation costs in real-time and investors that stay in the fund never bear liquidation losses. The prospect of *temporary* decreases in NAV in the future (even if they fully reverse) can be enough to entice investors to withdraw preemptively.

2. The Model

The economy lasts for three periods $t = 1, 2, 3$. There is a unit measure of ex-ante identical investors. Ex post, each investor turns out to be of type 1,2 or 3 with probability $\frac{1}{3}$ each. An investor of type t wants to consume most at date t. Payoffs equal $c_1 + \delta c_2 + \delta^2 c_3$ for type 1 investors, $c_2 + \delta c_3$ for type 2 investors and c_3 for type 3 investors, where c_t denotes consumption in period t and $\delta \in (0, 1)$ represents liquidity preference. (The lower δ, the higher liquidity preference.) At date 1, investors privately learn whether they are type 1 or not. At date 2, investors who are not type 1 privately learn whether they are of type 2 or 3. Denote by Θ_t the collection of information sets with regard to the own type at date t. We have $\Theta_1 = \{1, \text{not}1\}$ and $\Theta_2 = \Theta_3 = \{1, 2, 3\}$. In the aggregate, by a law of large numbers, $\frac{1}{3}$ of investors will be of type 1, $\frac{1}{3}$ of type 2 and $\frac{1}{3}$ of type 3.[5]

There is an intermediary ('fund' henceforward) that has access to a unit measure of assets whose returns it distributes investors. Each asset pays out 1 unit of the consumption good at date 3. When paying out investors at dates 1 and 2, the fund is restricted to follow sequential service, as will be described in detail below. Without loss of generality, I assume the fund distributes the return of the asset left in the fund at date 3 in a pro-rata fashion to the investors who are still in the fund at date 3. Investors can only communicate with the fund; they cannot communicate or trade with each other and they do not observe any actions taken by other investors. Since assets only pay out at date 3, the fund needs to liquidate assets in order to pay out investors at dates 1 and 2. The (secondary) market price of assets at a certain point in time within a given period (date 1 or 2) is a decreasing function of the total amount of assets liquidated within that period so far. Specifically, I assume

[5]There are well known technical problems with the law of large numbers in continuum economies. See Al-Najjar (2008) for a discussion of measurability issues in continuum-player games.

that

$$p_t(q) = \begin{cases} 1 & \text{if } q \leqslant \dfrac{1}{3} \\ \lambda & \text{if } q > \dfrac{1}{3} \end{cases} \quad \text{for } t = 1,2 \quad \text{with } \lambda \leqslant 1$$

where $p_t(q)$ is the market price of assets at a given point in time within date t, given that a measure q of assets have been sold so far in period t. The parameter λ can be interpreted as a measure for the liquidity of the fund's assets; the lower λ, the larger the price discount once a certain amount of assets have been sold.[6]

Investors who withdraw from the fund at dates 1 or 2 can store the consumption good across periods for later consumption (e.g. when a non-type 1 investor withdraws at date 1). When storing good for the first time, investors incur a fixed cost of $\kappa > 0$.[7]

First-best First-best is defined as the allocation that maximizes the expected payoff of investors subject to the economy's resource constraint, under the restriction that all investors of the same type must get the same payoff. It is not hard to see that the first-best allocation is to give each investor of type t one unit at date t and nothing in other periods.

Sequential Service Sequential service (in periods 1 and 2) is modelled in a similar fashion as in Ennis and Keister (2010), with some adaptations. The investors who choose to withdraw from the fund arrive at the fund sequentially and the fund needs to pay them out on the spot. Investors who choose not to withdraw from the fund stay at home in the given period and do not contact the fund. Showing up at the fund is therefore equivalent to withdrawing from the fund. Since payments need to be made on the spot, the payment to an investor that shows up at the fund can only be made contingent on the number of investors that showed up at the fund so far (within the same period and in previous periods) but not on withdrawal orders of investors who arrive later in the queue. The payout schedule specifies for date 1 a function $f_1 : [0, 1] \mapsto \mathbb{R}_+ \cup \{G\}$, where $f_1(z)$ is the payment to the z^{th} investor to show up at the fund in period 1. Setting $f(\cdot) = G$ means that the fund it

[6] The outside investors who buy the assets on secondary markets are not modelled explicitly and the liquidation losses incurred by the fund will be treated as a social loss. Fire sales can create social losses for a variety of reasons, for instance because outside investors forgo other, more socially useful investments (Stein (2012)) or because assets are bought by agents who are not well equipped to handle them (Gertler and Kiyotaki (2015)).

[7] κ could be seen as a fixed cost to access the storage facility. For simplicity, it needs to be paid only once and is independent of the amount stored. It captures the cost of withdrawing 'too early' from the fund, which may represent transaction- or search costs of changing to a different fund or bank, foregone return if funds are withdrawn early, etc.

imposed a gate (i.e. it suspended convertibility). An investor who shows up at the fund at date 1 with arrival point z receives $f_1(z)$ and cannot show up again in future periods. The investor then either consumes $f_1(z)$ or stores the good for future consumption, incurring the storage fixed cost κ. The only exception is if the fund imposed a gate: if an investor arrives at the fund with arrival point $\hat{z}$ and the payout schedule is such that $f_1(\hat{z}) = G$, then the investor receives nothing and can show up again in future periods.

Denote $\overline{z}_1$ as the total number (measure) of investors that withdraw at date 1. The number of investors left in the fund at date 2 equals $1-\overline{z}_1$. Analogous to date 1, the payout schedule specifies for date 2 a function $f_2(z;\overline{z}_1)$ which maps $[0, 1-\overline{z}_1]$ into $\mathbb{R}_+ \cup \{G\}$. As before, $f_2(z;\overline{z}_1)$ is the payment made to the z^{th} investor to show up in period 2. The date 2 payment schedule $f_2(z)$ can be made contingent on total date 1 withdrawals $\overline{z}_1$. The fund's entire payout schedule is then characterized by $f = (f_1, f_2)$. The payout schedule can in principle be such that the fund's resource constraint is violated if too many investors show up. If an investor shows up at the fund and the fund does not have the resources necessary to make the payment specified in the payout schedule, then the fund defaults and all investors who are still in the fund receive nothing. Investors know the structure of the game and are thus aware of the fund's resource constraint.

Stable and Floating NAV Funds The default payout schedule (i.e. without fees and gates) of a stable NAV fund is to pay out a fixed amount to every investor who withdraws, without any adjustment for changes in the market price of the fund's assets. Since implementing first-best requires that the fund pay one unit to all type t investors at date t, it is natural to model the default payout schedule of a stable NAV as having a fixed NAV of 1. With floating NAV, the share price reflects in real-time the market value of the fund's asset portfolio and thus adjusts downwards automatically if the liquidation price of the asset decreases.

Definition 2.1. *A stable NAV fund has default payout schedule* $f_1(z) = f_2(z,\cdot) = 1$ *for all* z. *A floating NAV fund has default payout schedule* $f_1(z) = f_2(z,\cdot) = p(z)$ *for all* z.

Withdrawal Game and Equilibrium Given the fund's payout schedule $f = (f_1, f_2)$, investors choose their withdrawal strategy in a non-cooperative game. At the beginning of periods 1 and 2, the investors who are still in the fund (those who did not withdraw in previous periods) wake up in a random order. The order of waking up at date t is captured by an index i_t that is randomly allocated to all investors who are still in the fund at date t. The date 1 indices are given by $i_1 \in [0, 1]$ and date 2 indices by $i_2 \in [0, 1 - \overline{z}_1]$. Each investor who is still in the fund at date t is allocated each position in the corresponding interval with identical probability, independent of the investor's type and independent of the indices in previous periods. Investors do not observe their own index i_t or any indices allocated to other investors.[8] Upon waking up, an investor sees how much she can withdraw from the fund, and she has two actions available: 'withdraw' (which is equivalent to 'go to fund') or 'not withdraw' (which is equivalent to 'stay at home'). Hence in any given period t, the investors with the lowest indices i_t are first to decide whether to withdraw, after which those with higher indices follow, etc. This implies that, in every period, the arrival point z of an investor will be weakly lower than her index i_t. If some investors with lower indices decide not to withdraw, the arrival point will be strictly lower than the index.

The position in the wake-up-order is similar to the 'position in the queue' common in models of runs. Indeed, the setting here is isomorphic to a setting where (i) investors first decide whether or not to go the fund; (ii) the investors who go to the fund arrive at the fund in a random order and (iii) once an investor arrives at the fund and sees how much she can withdraw, she decides whether or not to withdraw.

When making the withdrawal decision, an investor's information set consists of four elements: (i) the time period; (ii) the information about the own type; (iii) the amount that can be withdrawn from the fund and (iv) the observed history of events up to this point. Investors do not observe any actions taken by other investors. The observed history is the history of amounts which an investor could have withdrawn in previous periods. For instance, an investor who is still in the fund at date 2 knows how much she could have withdrawn at date 1. The observed history may in principle give investors information about how many other investors withdrew from the fund in previous periods, as well as information about their own indices i_t in previous periods.[9]

[8] It is possible that the fund could set its payout schedule such that the indices are revealed to investors. This will not play a role in the following analysis however.

[9] Note however that there is no risk on the aggregate level so that play on the aggregate level proceeds in a deterministic fashion for any strategy profile. The only information investors effectively learn from the observed history relates to the indices i_t they were allocated in previous periods. This is unrelated to future indices and thus does not convey any payoff-relevant information.

An investor's behavior strategy maps her information sets into probability distributions over the two possible actions 'withdraw' and 'not withdraw'. To streamline notation, I denote by $s_t(x, y)$ an investor's withdrawal strategy given that she finds herself in period t, the information about her own type is x, and the fund pays an amount y if she withdraws. The function s_t then prescribes behavior at date t for any observed history up to point t, given (x, y). As we shall see, this is without loss of generality (see also footnote 9). Denote $R(f_t)$ as the range of f_t, that is, $R(f_t)$ is the set of payments which the fund may potentially offer to an investor at date t. The date t strategy s_t is then given by

$$s_t : \Theta_t \times R(f_t) \mapsto \ \Delta\{\text{'withdraw', 'not withdraw'}\} \tag{1}$$

where Δ is the simplex of the pure strategy set. An investor's (behavior) strategy is given by $s = (s_1, s_2)$. Note that payoffs depend only on the aggregate behavior of other investors, not on actions by individual other investors. Since there is no risk on the aggregate level, play proceeds in a deterministic fashion on the aggregate level for any strategy profile s. Each player thus essentially plays against a deterministic continuum.

Throughout the paper, I will limit attention to symmetric equilibria in which all investors choose the same strategy s. An equilibrium is defined to be a strategy s^* such that (i) s^* maximizes each investor's expected payoff given that all others play s^* and (ii) s^* is sequentially rational in the sense that investors do not play strategies that are strictly dominated conditional on their information set having been reached. The equilibrium refinement in (ii) concerns behavior in information sets off the equilibrium path and corresponds to the restriction on off-equilibrium behavior imposed by the *weak perfect Bayesian equilibrium* concept. I will sometimes be a bit loose in the terminology and say that a strategy is 'strictly dominated' if it violates requirement (ii).

The remainder of the paper will be concerned with finding payout policies f that uniquely implement first-best; more precisely, payout policies f which are such that all equilibria s^* of the withdrawal game under payout schedule f implement first-best. Implementing first-best requires that investors withdraw from the fund only at the date that corresponds to their type. A *run equilibrium* denotes an equilibrium of the withdrawal game in which a strictly positive measure of investors withdraw from the fund at a date that does not correspond to their type.

3. Stable NAV Funds

It is not hard to see that a stable NAV fund (with a fixed NAV of 1) does implement first-best as an equilibrium of the withdrawal game. In particular, type 2 and 3 investors have no incentive to withdraw early if no other type 2 and 3 investors do so. However, stable NAV funds without fees or gates are susceptible to runs whenever assets are not perfectly liquid. If all investors play 'withdraw' at date 1 irrespective of their type, then the fund defaults during date 1 (it will run out of assets before everybody has shown up) and investors who do not withdraw at date 1 receive nothing. This leads to the well-known Diamond Dybvig type run equilibrium at date 1. We thus get the following result, stated without separate proof:

Proposition 3.1. *A stable NAV fund that uses neither fees nor gates uniquely implements first-best only if assets are perfectly liquid ($\lambda = 1$).*

3.1. Redemption Fees that Mirror Liquidation Losses

A stable NAV fund with neither fees nor gates is susceptible to runs because liquidation losses incurred by the fund (when paying out investors) are borne by investors who remain in the fund. In this subsection, I show that under some circumstances stable NAV funds can uniquely implement first-best by charging redemption fees that mirror the liquidation losses which the fund incurs by paying out investors. Such redemption fees ensure that redeeming investors internalize the liquidation losses which the fund incurs by paying them out. As an important corollary, this subsection also shows that such redemption fees are not always enough to prevent runs.

With a redemption fee τ (modelled as a haircut on withdrawals) that mirrors liquidation costs, redeeming investors only receive the liquidation price of λ units (instead of 1 unit) if they withdraw after the fund has depleted its liquid assets. The payout schedule, denoted f^τ, is given by:

$$f_1^\tau(z) = f_2^\tau(z;\cdot) = \begin{cases} 1 & \text{if } z \leqslant \frac{1}{3} \\ 1-\tau & \text{otherwise} \end{cases} \quad \text{with} \quad \tau = 1-\lambda \tag{2}$$

Note that payout schedule f^τ is the same as the payout schedule of a floating NAV fund. In the current setting, charging a redemption fee that tracks liquidation costs is equivalent to letting the share price adjust to changes in the secondary market price of the asset. Note further that with payout schedule f^τ the fund always pays 1 unit at date 3 to the investors in the fund at date 3, independent of how many investors withdraw at dates 1 and 2. In this sense, redemptions never impose liquidation losses on the investors who remain in the fund. It is straightforward that payout schedule f^τ does implement first-best as an equilibrium of the withdrawal game. Furthermore, the following result implies that when looking for run equilibria, we can limit attention to equilibria in which non-type 1 investors withdraw at date 1:

Lemma 3.1. *The withdrawal game under payout schedule f^τ uniquely implements first-best if and only if the withdrawal game does not exhibit an equilibrium in which non-type 1 investors withdraw at date 1.*

The proof of lemma 3.1 is given in appendix A. To see why runs at date 1 are possible under payout schedule f^τ we first need to consider what can happen at date 2. It will be convenient to denote B as the measure of type 1 and 2 investors left in the fund at the beginning of date 2. Since both type 1 and 2 investors prefer consumption at date 2 over consumption at date 3, B is also the measure of investors that try to withdraw 1 unit from the fund at date 2. (Since type 3 investors always receive 1 unit at date 3 under payout schedule f^τ it is never optimal for them to withdraw at date 2.) If no run occurs at date 1, then all type 1 investors (and only they) withdraw at date 1. In this case, B simply equals the measure of type 2 investors, that is, $B = \frac{1}{3}$.

Next, we denote by $P(B)$ the probability that an individual type 1 or 2 investor in the fund at date 2 manages to withdraw 1 unit at date 2. Under payout schedule f^τ the fund pays out 1 unit to a measure $\frac{1}{3}$ of investors at date 2 and then imposes the redemption fee. If more than a measure $\frac{1}{3}$ of investors want to withdraw 1 unit at date 1, then those arriving late in line at date 2 (those allocated indices $i_2 > \frac{1}{3}$) can either pay the redemption fee at date 2 (i.e. withdraw λ units at date 2) or wait until date 3. The probability $P(B)$ thus equals $\frac{1}{3}$ divided by the number of investors that wish to withdraw 1 unit at date 2: $P(B) = \min\left\{\frac{1}{3B}, 1\right\}$. The larger the number of type 1 and 2 investors in the fund at date 2 (the higher B), the lower the probability that an individual type 1 or 2 investor manages to withdraw 1 unit at date 2. Note that we have $P(\frac{1}{3}) = 1$ which means that, if no run occurs at date 1 (implying $B = \frac{1}{3}$), then all type 2 investors can withdraw 1 unit at date 2.

To examine whether runs constitute equilibria of the withdrawal game under payout schedule f^{τ} we need to study best responses of non-type 1 investors at date 1. Recall that a non-type 1 investor turns out to be of type 2 or 3 with probability $\frac{1}{2}$ each. Under payout schedule f^{τ}, type 3 investors always receive 1 unit at date 3, which gives them a payoff of 1. A type 2 investor can withdraw 1 unit at date 2 with probability $P(B)$; with probability $1 - P(B)$, the investor arrives late in line at date 2 and can either pay the fee on date 2 withdrawals, which yields a payoff of λ, or receive 1 unit at date 3, which yields a payoff of δ. If $\lambda \geqslant \delta$, paying the fee on date 2 withdrawals is the best response of type 2 investors at date 2. If $\lambda < \delta$, *not* paying the fee on date 2 withdrawals is the strictly best response for type 2 investors. Denote by $Z(P)_{[\lambda \geqslant \delta]}$ and $Z(P)_{[\lambda < \delta]}$ the expected payoff of *not* withdrawing at date 1 for a non-type 1 investor, given that the probability of being able to withdraw 1 unit at date 2 is P, and given that parameters satisfy $\lambda \geqslant \delta$ and $\lambda < \delta$ respectively:

$$\begin{aligned} Z(P(B))_{[\lambda \geqslant \delta]} &= \frac{1}{2} \overbrace{\left[P(B) + \lambda(1 - P(B))\right]}^{\text{expected payoff type 2}} + \frac{1}{2} \overbrace{1}^{\substack{\text{payoff} \\ \text{type 3}}} \\ Z(P(B))_{[\lambda < \delta]} &= \frac{1}{2} \underbrace{\left[P(B) + \delta(1 - P(B))\right]}_{\text{expected payoff type 2}} + \frac{1}{2} \underbrace{1}_{\substack{\text{payoff} \\ \text{type 3}}} \end{aligned} \tag{3}$$

Running at date 1 (i.e. trying to withdraw 1 unit at date 1) is the best response for a non-type 1 investor whenever the payoff of withdrawing 1 unit at date 1, which equals $1 - \kappa$, is higher than the expected payoff of remaining in the fund. For an individual non-type 1 investor, running at date 1 is therefore the best response iff $1 - \kappa \geqslant Z(P(B))$ where B itself is an equilibrium outcome that depends on other investors' behavior.[10] Running can only be optimal if $P < 1$, that is, run incentives for non-type 1 investors are always related to the possibility of turning out to be a type 2 investor who cannot withdraw 1 unit at date 2. Furthermore, as discussed above, P can only drop below 1 if some non-type 1 investors withdraw at date 1.

[10]Note also that paying the redemption fee at date 1 is never optimal for non-type 1 investors since $Z(P) > \lambda$ for all $P \in [0, 1]$.

In the next step, we need to determine which values of B (and hence P) can be part of an equilibrium. This depends on the behavior of type 1 investors at date 1. In particular, the number of type 1 and 2 investors in the fund at date 2 (B) will depend on the behavior of the type 1 investors who arrive late in line in a run at date 1 (those with indices $i_1 > \frac{1}{3}$). In order to study best responses of type 1 investors at date 1, we denote by $Y(P)$ the expected payoff for a type 1 investor of *not* withdrawing at date 1 given that the probability of being able to withdraw 1 unit at date 2 is P. That is, $Y(P)$ is defined analogously to $Z(P)$; the difference is that $Y(P)$ is for type 1 investors while $Z(P)$ is for non-type 1 investors. At date 2, incentives of type 1 investors are the same as those of type 2 investors. In particular, a type 1 investor arriving late in line at date 2 will prefer to pay the redemption fee at date 2 if $\lambda \geqslant \delta$ and will prefer to wait until date 3 if $\lambda < \delta$. We thus get:

$$\begin{aligned} Y(P(B))_{[\lambda \geqslant \delta]} &= \delta \left[P(B) + \lambda(1 - P(B)) \right] \\ Y(P(B))_{[\lambda < \delta]} &= \delta \left[P(B) + \delta(1 - P(B)) \right] \end{aligned} \tag{4}$$

As discussed before, $B > \frac{1}{3}$ (and hence $P < 1$) is only possible if some type 1 investors are still in the fund at date 1. Such a 'backlog' of type 1 investors in the fund at date 2 can occur if two conditions are fulfilled. First, some non-type 1 investors withdraw at date 1, so that not all type 1 investors can withdraw 1 unit at date 1. Second, the type 1 investors who arrive late in line in the run at date 1 (those with indices $i_1 > \frac{1}{3}$) prefer not to pay the redemption fee at date 1. An individual type 1 investor is better off not paying the redemption fee at date 1 if $Y(P(B)) \geqslant \lambda$; just as with non-type 1 investors, staying in the fund becomes less attractive for type 1 investors if the probability that they can withdraw 1 unit at date 2 decreases.

To illustrate under which conditions a run at date 1 constitutes an equilibrium, let us play through a scenario (that may or may not constitute an equilibrium) where all investors run on the fund at date 1. That is, we are considering a scenario where all investors play $s_1(\cdot, 1) =$'withdraw'. In this scenario, only investors in the first one-third of the line (those with indices $i_1 \leqslant \frac{1}{3}$) can withdraw 1 unit at date 1. Among those, by a law of large numbers, $\frac{1}{3}$ will be of each type 1, 2 and 3. Suppose further that *none* of the type 1 investors arriving late in line at date 1 withdraw at date 1. A measure $\frac{2}{3}$ of investors is then still in the fund at the beginning of date 2; among those, a fraction $\frac{1}{3}$ is of each type. The measure of type 1 and 2 investors in the fund at date 2 then equals $B = \frac{1}{3}\frac{2}{3} + \frac{1}{3}\frac{2}{3} = \frac{4}{9}$. It is not hard to see that this is an upper bound on B in equilibrium since the backlog of type 1 investors in the fund at date 2 cannot be higher than in the scenario just described. Note that if $B = \frac{4}{9}$, then we have $P = \frac{3}{4}$ which is thus a lower bound on P in equilibrium.

The run scenario described above can only be an equilibrium if two conditions are fulfilled. First, type 1 investors arriving late in line at date 1 must not find it optimal to pay the redemption fee, which causes a backlog of type 1 investors left in the fund at date 2. Second, non-type 1 investors must find it optimal to run at date 1 if they expect there to be a backlog of type 1 investors in the fund at date 2. Consider first the best response of type 1 investors arriving late in line in the run. *Not* paying the redemption fee at date 1 is their best response if the fee is high (λ is low) and liquidity preference is not very strong (δ is high). We have that:

$$\lambda \leqslant \frac{3}{4}\delta + \frac{1}{4}\delta^2 \;\Leftrightarrow\; \lambda \leqslant Y(\tfrac{3}{4})_{[\lambda<\delta]} \tag{5}$$

If condition (5) is fulfilled, then *not* paying the redemption fee at date 1 is optimal for type 1 investors even if the probability of being able to withdraw 1 unit at date 2 is at its lower bound ($P = \frac{3}{4}$). Let us assume for the moment that condition (5) is fulfilled, and consider best responses of non-type 1 investors at date 1. Given $P = \frac{3}{4}$, participating in the run is the best response for an individual non-type 1 investor iff:

$$1 - 8\kappa \geqslant \delta \;\Leftrightarrow\; \overbrace{1-\kappa}^{\substack{\text{payoff of withdrawing} \\ \text{1 unit at date 1}}} \geqslant \overbrace{Z(\tfrac{3}{4})_{[\lambda<\delta]}}^{\substack{\text{payoff of remaining} \\ \text{in fund given } P=0.75}} \tag{6}$$

It follows that, under payout schedule f^{τ}, there exists an equilibrium where all investors run on the fund at date 1 if both conditions (5) and (6) are fulfilled. We also get the converse result: if parameters satisfy condition (5), run equilibria do not exist if parameters do not satisfy condition (6). To see this, recall that $Z(P)_{[\lambda<\delta]}$ is strictly increasing in P and we have that $P \geqslant \frac{3}{4}$ in equilibrium. This implies that, if condition 6 is not fulfilled then, in equilibrium, non-type 1 investors are always better off not running.

Note that the run equilibria discussed above are not related to liquidation losses incurred by the fund. If conditions (5)-(6) are fulfilled, runs can occur despite the fact that investors never withdraw from the fund when the market price of the asset falls below 1, so that the fund never liquidates assets at a loss. Intuitively, the runs are due to the fact that, in periods where redemptions are high and the fund imposes a redemption fee, some of the investors who truly need to consume in the given period (i.e. type 1 investors at date 1) will wait with withdrawing until the next period, when the fund lifts the redemption fee. This implies that heavy redemptions today give rise to a 'backlog' of investors who want to withdraw the next period; as a result, redemptions will be again high the next period and the fund needs to impose the redemption fee again. For investors who do *not* need liquidity today, the fear of not being able to withdraw the next period without paying a fee (as a result of the above mentioned backlog) provides the incentive to withdraw preemptively today.

Above, we have discussed the case where type 1 investors do not find it optimal to pay the redemption fee at date 1 in case they arrive late in the queue at date 1. Consider now the opposite case, where type 1 investors who arrive late in the queue find it optimal to pay the redemption fee at date 1. The easiest case is the one where $\lambda \geqslant \delta$, in which case paying the redemption fee at date 1 is a dominant strategy for type 1 investors since receiving λ units at date 1 gives a higher payoff than receiving 1 unit at date 2. If a run occurs at date 1, all type 1 investors who arrive late in line at date 1 will pay the redemption at date 1. This implies that, no matter how many investors run at date 1, there will never be a backlog of type 1 investors left in the fund at date 2. We then have that $B \leqslant \frac{1}{3}$ and $P = 1$ in equilibrium, so that non-type 1 investors never have an incentive to run at date 1. Given $\lambda \geqslant \delta$, payout schedule f^τ thus uniquely implements first-best.

Note the non-monotonic effect of liquidity preference δ on the propensity to run under redemption fees as in payout schedule f^τ. On the one hand, higher liquidity preference (lower δ) makes redemption fees more effective since it is easier to induce type 1 investors to actually pay the fee if a run occurs at date 1. This in turn reduces non-type 1 investors' incentive to participate in a run since it reduces (or eliminates) the backlog of type 1 investors left in the fund at date 2. On the other hand, within the subset of the parameter space where investors are not willing to pay the redemption fee (that is, if λ is low relative to δ) higher liquidity preference makes it less likely that fees can prevent runs. If others run, non-type 1 investors correctly anticipate that they may not be able to withdraw the next period if they need to. (More precisely, they may only be able to withdraw at date 2 by paying a fee which they are not willing to pay.) The higher liquidity preference, the higher is the loss in payoff for type 2 investors if they consume only at date 3 and hence the higher the propensity to run at date 1. We conclude this discussion with the following proposition that summarizes and completes the results discussed above:

Proposition 3.2.

Payout schedule f^{τ} uniquely implements first-best if and only if:

(i) $\frac{\lambda}{\delta} \geqslant 1$ *or*

(ii) $\frac{\lambda}{\delta} \in \left(\frac{3}{4} + \frac{1}{4}\delta\,,\,1\right)$ *and* $1 - 2\,\kappa < \frac{\lambda}{\delta}$ *or*

(iii) $\frac{\lambda}{\delta} \leqslant \frac{3}{4} + \frac{1}{4}\delta$ *and* $1 - 8\,\kappa < \delta$

The proof of proposition 3.2 is given in appendix B, although much of it is contained in the discussion above. I elaborated on conditions (i) and (iii) above. Condition (ii) deals with the case where λ is within the (narrow) range $\lambda \in (\frac{3}{4}\delta + \frac{1}{4}\delta^2, \delta)$. In this case, some fraction of type 1 investors arriving late in a run at date 1 (on or off the equilibrium path) will pay the fee on date 1 withdrawals. The lower the value of λ within this interval, the smaller the fraction of late-arriving type 1 investors that pay the fee at date 1 and hence the higher the incentive to run for non-type 1 investors. Figure 1 depicts the set of parameters (shaded area) for which redemption fees that mirror liquidation losses prevent runs, given that $\kappa = 0.03$. Changing the cost of withdrawing early κ will make the shaded area uniformly bigger (increase in κ) or smaller (decreasing in κ) without changing its basic shape.

Intuitively, redemption fees whose sole purpose it is to make sure that redeeming investors internalize liquidation losses tend to be a suitable means to eliminate runs only if the fund's assets are sufficiently liquid. If assets are illiquid, the fee must be relatively high to ensure that liquidation losses are fully internalized by those who withdraw. Unless liquidity preference is very strong, investors are not willing to pay such a high redemption fee, which can lead to self-fulfilling run equilibria á-la Engineer (1989). In the next subsection, I show how a stable NAV fund can eliminate runs even if assets are very illiquid. This can be done by introducing another redemption fee

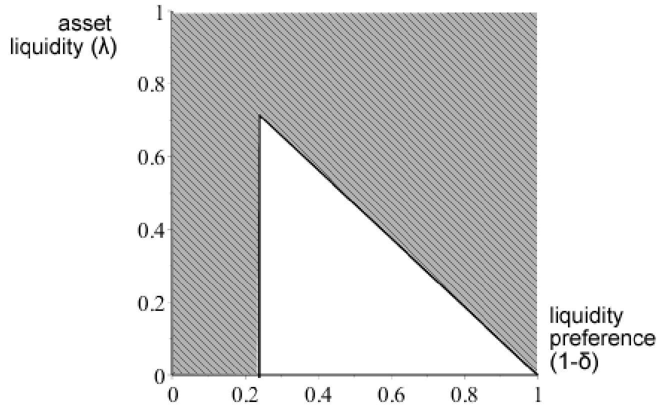

Figure 1: Set of parameters for which payout schedule f^{τ} prevents runs, given $\kappa = 0.03$.

that is not related to asset liquidations. The purpose of this additional redemption fee is to implement a redistribution among investors in such a way that non-type 1 investors are incentivized to stay in the fund if (off the equilibrium path) others run at date 1.

3.2. Two-Stage Redemption Fees

One result of the previous subsection is that redemption fees whose sole purpose it is to cover liquidation costs are not always enough to eliminate runs. In this subsection, I show how a stable NAV fund can prevent runs by adding a second redemption fee to its payout schedule.[11] As in the previous subsection, the fund charges redemption fees that mirror liquidation losses when it runs out of liquid assets. Additionally, if the fund has to activate the redemption fee at date 1 (i.e. if a run occurs at date 1) then (and only then) the fund will charge another redemption fee on *all* withdrawals at date 2. The fund then pays any fee revenue raised at date 2 to those who withdraw at date 3. Importantly, the fund charges the fee on date 2 withdrawals even if it does not liquidate any assets at date 2. The second-stage fee is not related to liquidation losses; instead, its purpose is to implement a redistribution among investors in the fund. Specifically, the second-stage redemption fee implements a redistribution from those who withdraw at date 2 (type 1 and type 2 investors) to those who withdraw at date 3 (type 3 investors).

[11]The payout schedule with two redemption fees as studied in this subsection is not necessarily the only type of payout schedule that uniquely implements first-best for the entire parameter space. However, it seems unlikely that there is another, comparably simple payout schedule achieves the same.

The additional fee charged at date 2 (after a run at date 1) is designed in such a way that a backlog of type 1 investors in the fund is used to the advantage of non-type 1 investors in the fund. The larger the backlog of type 1 investors left in the fund at date 2, the larger the amount of redemptions at date 2 and hence the larger the amount of fee revenue collected by the fund at date 2. This fee revenue is then paid to the investors who withdraw at date 3. Note that both type 1 and 2 investors will need to pay the fee on date 2 withdrawals after a run occurred at date 1. However, in the aggregate, the fee leads to a redistribution from type 1 investors in the fund to non-type 1 investors. The reason is that type 1 investors always incur the fee while non-type 1 investors incur the fee only if they turn out to be of type 2. If the redemption fee at date 2 is set to the right level (not too high and not too low) then the prospect of profiting from the fee revenue raised at date 2 incentivizes non-type 1 investors to remain in the fund if (off the equilibrium path) others run at date 1. In short, the additional redemption fee at date 2 ensures that the prospect of a backlog of type 1 investors left in the fund at date 2 gives non-type 1 investors an incentive to *remain* in the fund instead of giving them an incentive to withdraw preemptively as in subsection 3.1.

Expression (7) below describes formally the payout schedule with the two-stage redemption fee, denoted $f^{\tau\hat{\tau}}$. Payout schedule $f^{\tau\hat{\tau}}$ augments schedule f^{τ} of the previous subsection by an additional fee $\hat{\tau}$ at date 2 whenever a run occurred at date 1. Here we run into a technical problem with our set-up: since it is possible that nobody pays the redemption fee at date 1 if a run occurs (see the previous subsection) actual withdrawals at date 1 ($\overline{z}_1$) might be the same whether a run occurs or not. To circumvent this problem, and abusing the set-up a bit, I assume the fund can distinguish (ex post) between situations where it had to 'activate' the redemption fee at date 1 and situations where it had not to. I will say that $\overline{z}_1 = \frac{1}{3} + F > \frac{1}{3}$ if the measure of withdrawals at date 1 equals $\frac{1}{3}$ *and* the fund had to activate the redemption fee, that is, more than a measure $\frac{1}{3}$ of investors would have withdrawn 1 unit at date 1 had the fund not imposed the redemption fee.[12] If a run occurs at date 1 (and the fund thus needs to activate the redemption fee at date 1) then an additional redemption fee, denoted $\hat{\tau}$, will be charged on all date 2 withdrawals. Once the date 2 liquid assets are exhausted

[12] Another way to solve the problem would be to allow a measure $\frac{1}{3} + \varepsilon$ of investors to withdraw 1 unit at date 1 with $\varepsilon > 0$ but infinitely small. Whenever $\overline{z}_1 > \frac{1}{3}$, the fund then knows (ex post) that a run occurred at date 1. This approach would give the same results but would convolute the exposition.

(which is the case after a measure $\dfrac{1}{3(1-\hat{\tau})}$ of investors have withdrawn at date 2[13]) the fund again charges a fee that compensates it for liquidation losses. Formally, the payout schedule is thus given by

$$\text{Date 1: } f_1^{\tau\hat{\tau}}(z) = f_1^{\tau}(z)$$
$$\text{Date 2: } f_2^{\tau\hat{\tau}}\left(z; \overline{z}_1 \leqslant \tfrac{1}{3}\right) = f_2^{\tau}(z), \quad f_2^{\tau\hat{\tau}}\left(z; \overline{z}_1 > \tfrac{1}{3}\right) = \begin{cases} 1-\hat{\tau} & \text{if } z \leqslant \frac{1}{3}\frac{1}{1-\hat{\tau}} \\ \lambda & \text{otherwise} \end{cases} \quad \text{with } \hat{\tau} \in (0,1) \tag{7}$$

For each investor that incurs the redemption fee $\hat{\tau}$ at date 2, the fund only needs to sell $1-\hat{\tau}$ units of a project. If the fund activates the date 2 redemption fee $\hat{\tau}$ it can pay more than 1 unit to each investor left in the fund at date 3 whenever some investors withdraw at date 2 (that is, whenever $\overline{z}_2 > 0$). The reason is that the fund collects fee revenue at date 2 over and above any liquidation losses it incurs. Denoting $c_3^{\hat{\tau}}(\overline{z}_1, \overline{z}_2)$ as the amount paid out to each investor at date 3 after activating the fee $\hat{\tau}$, we get:

$$c_3^{\hat{\tau}}(\overline{z}_1, \overline{z}_2) = \frac{\overbrace{1-\overline{z}_1-\overline{z}_2+\min\{\overline{z}_2, \frac{1}{3}\frac{1}{1-\hat{\tau}}\}\,\hat{\tau}}^{\text{assets in fund at date 3}}}{\underbrace{1-\overline{z}_1-\overline{z}_2}_{\text{investors in fund at date 2}}} = 1 + \frac{\overbrace{\min\{\overline{z}_2, \frac{1}{3}\frac{1}{1-\hat{\tau}}\}\hat{\tau}}^{\substack{\text{fee revenue} \\ \text{raised at date 2}}}}{1-\overline{z}_1-\overline{z}_2} \tag{8}$$

As with the previous payout schedules, it is straightforward that payout schedule $f^{\tau\hat{\tau}}$ implements first-best as an equilibrium of the withdrawal game. Different to the previous payout schedules, payout schedule $f^{\tau\hat{\tau}}$ *uniquely* implements first-best as long as the redemption fee $\hat{\tau}$ on date 2 withdrawals is set to the right level (not too high and not too low). To gain some intuition about why the redemption fee $\hat{\tau}$ needs to be set to the right level for it to eliminate runs, it is useful to discuss the cases where $\hat{\tau}$ is set either too high or too low.

Consider first the case where the date 2 redemption fee $\hat{\tau}$ is set to a very high level. In this case, none of the investors in the fund at date 2 will be willing to pay the fee $\hat{\tau}$ if the fund activates it. Instead, all investors who are left in the fund at date 2 will wait to withdraw until date 3. Non-type 1 investors' propensity to run at date 1 would then be even higher compared to payout schedule f^{τ}.

[13]If a measure z withdraw at a redemption fee $\hat{\tau}$, the fund pays out an amount $z(1-\hat{\tau})$ of cash. Given that the fund has a measure $\frac{1}{3}$ of liquid assets at date 2, liquid assets are exhausted if $z(1-\hat{\tau}) = \frac{1}{3}$ or $z = \dfrac{1}{3}\dfrac{1}{1-\hat{\tau}}$.

To see this, note first that no fee revenue is collected at date 2 in case a run occurs at date 1 since nobody pays the fee at date 2. Furthermore, if a run occurs at date 1, *all* type 2 investors left in the fund at date 2 will consume only at date 3 as a result of the prohibitively high fee charged on date 2 withdrawals. Compared to fees that only offset liquidation losses (as in payout schedule 2) this exacerbates non-type 1 investors' fear of turning out to be a type 2 investor that cannot withdraw at date 2. Consider next the case where the date 2 redemption fee $\hat{\tau}$ is set very low. In this case, the number of (type 1 and 2-) investors who wish to withdraw $1-\hat{\tau}$ units at date 2 (after a run occurred at date 1) may exceed the fund's liquid assets at date 2. Those who arrive late in line at date 2 can then only withdraw at the (higher) redemption fee that compensates the fund for liquidation losses. We can then have the same type of run equilibrium as discussed in subsection 3.1. This can be seen most easily for the case where the fee on date 2 withdrawals is set arbitrarily small ($\hat{\tau} \to 0$) in which case the payout schedule with the two-stage redemption fee ($f^{\tau\hat{\tau}}$) becomes equivalent to payout schedule f^{τ}.

The discussion in the previous paragraph highlights that, in order to improve on a payout schedule where redemption fees only mirror liquidation losses (payout schedule 2) the second-stage redemption fee ($\hat{\tau}$) in payout schedule $f^{\tau\hat{\tau}}$ should be neither too high nor too low. On the positive side, proposition 3.3 (whose proof is given in appendix C) shows that the interval within which $\hat{\tau}$ should lie in order to eliminate runs is non-empty for the entire parameter space:[14]

Proposition 3.3. *Payout schedule $f^{\tau\hat{\tau}}$ with a second-stage redemption fee $\hat{\tau} \in \left(\frac{1-\delta}{2}, \frac{1-\delta}{1+\delta}\right)$ uniquely implements first-best.*

Intuitively, if the fee $\hat{\tau}$ is set within the interval given in proposition 3.3, it satisfies two criteria:

(i) The fee is not 'too low' in the sense that the number of investors that wish to withdraw at date 2 (at the redemption fee $\hat{\tau}$) never exceeds the fund's liquid assets at date 2.[15]

[14] The interval in proposition 3.3 gives a sufficient but not necessary condition to prevent run equilibria. Fees $\hat{\tau}$ outside the interval may also prevent run equilibria, depending on the cost of withdrawing early κ.

[15] To be precise, it may well be the case that not all type 1 and 2 investors left in the fund at date 2 will withdraw at date 2 after (off the equilibrium path) a run occurred at date 1. In this case, the number of investors who withdraw at date 2 ($\bar{z}_2$) will be such that type 2 investors are just indifferent between withdrawing $1-\hat{\tau}$ units at date 2 and withdrawing $c_3^{\hat{\tau}}(\bar{z}_1, \bar{z}_2)$ units at date 3. Hence all type 2 investors receive a payoff equal to $1-\hat{\tau}$ even if not all of them withdraw at date 2. A more detailed description is given in appendix C.

(ii) The fee is not ‘too high’ in the sense that sufficiently many type 1 and 2 investors left in the fund at date 2 will pay the fee on date 2 withdrawals instead of waiting until date 3 (if a run occurs at date 1). This generates fee revenue which can be paid to those who withdraw at date 3.

Note that both the lower- and the upper bound on the redemption fee $\hat{\tau}$ in proposition 3.3 are increasing in liquidity preference (i.e. decreasing in δ). The lower bound on $\hat{\tau}$ is related to the fact that the fund's liquid assets at date 2 are limited: if the fee is set below the lower bound, then the number of investors who wish to withdraw at date 2 after a run occurred at date 1 may be such that the fund's liquid assets at date 2 are not sufficient to pay them all out. The higher liquidity preference, the higher is the desire by type 1 and 2 investors to consume at date 2 rather than date 3. Higher liquidity preference therefore requires that the fund charge a higher fee on date 2 withdrawals to make sure that demand for date 2 withdrawals never exceeds liquid assets at date 2. The upper bound on $\hat{\tau}$ is related to type 1 and 2 investors' incentive constraint: the fee $\hat{\tau}$ must be such that a high enough number of type 1 and 2 investors are willing to pay the fee on date 2 withdrawals, should the fee be activated. The higher liquidity preference, the higher the willingness of type 1 and 2 investors to pay a fee on date 2 withdrawals instead of waiting until date 3. For this reason, the upper bound on $\hat{\tau}$ is increasing in liquidity preference as well.

Evidently, the second-stage fee $\hat{\tau}$ is not necessary to eliminate runs if parameters are such that runs are already eliminated if the fund charges fees that offset liquidation losses (i.e. if parameters satisfy the condition in proposition 3.2). On the other hand, proposition 3.3 tells us that the second-stage fee $\hat{\tau}$ will not do any harm either as long as it is set to the right level.

4. Floating NAV Funds

In a floating NAV fund, changes in the secondary market price of the asset lead to changes in the share price, which means that liquidation losses are always internalized by the investors who withdraw. Heavy redemptions in a given period lead to a temporary decrease in the share price, as a result of the temporary decrease in the secondary market price of the asset. In the current setting, the payout schedule of a floating NAV fund is equivalent to the one of a stable NAV fund that charges redemption fees covering liquidation losses; letting the NAV adjust to changes in secondary market prices and charging redemption fees that track liquidation costs are two equivalent ways to ensure that redeeming investors internalize liquidation costs. It follows immediately that floating NAV funds uniquely implement first best if and only if parameters satisfy the condition in proposition 3.2. Different to stable NAV funds, floating NAV funds may thus uniquely implement first-best without the use of redemption fees or gates. However, the potential for temporary decreases in the share price (that will fully reverse) can give rise to runs for the same reason as redemption fees: non-type 1 investors may withdraw preemptively at date 1 for fear of turning out to be a type 2 investor who can only withdraw at a depressed share price at date 2. The less liquid the fund's assets (the lower λ), the larger the decrease in the share price if liquid assets are exhausted and (keeping all else the same) the higher the propensity to run.

It also follows immediately from the discussion in subsection 3.2 that floating NAV funds can uniquely implement first-best by imposing a 'second-stage' redemption fee as in payout schedule $f^{\tau\hat{\tau}}$: whenever the share price at date 1 decreases, the fund charges a fee $\hat{\tau}$ on date 2 withdrawals. While redemption fees covering liquidation losses are not necessary in floating NAV funds (since changes in the NAV already incorporate changes in secondary market prices) redemption fees with a redistributionary purpose still have a role in floating NAV funds. We thus conclude this discussion with the following proposition which is stated somewhat informally:

Proposition 4.1.

(i) Floating NAV funds uniquely implement first-best without the use of fees or gates iff parameters satisfy the condition in proposition 3.2.

(ii) Floating NAV funds can uniquely implement first-best (for the entire parameter space) by charging a redemption fee $\hat{\tau}$ at date 2 after a decrease in the NAV at date 1 (as in payout schedule $f^{\tau\hat{\tau}}$) as long as the level of $\hat{\tau}$ is set within the interval given in proposition 3.3.

5. The Role of Gates in Preventing Runs

Sections 3 and 4 showed that both stable and floating NAV funds can eliminate runs by using redemption fees alone. Gates are therefore redundant as a tool to prevent runs as long as there are no regulatory (or other) restrictions on the redemption fees that funds can charge. This result is not very surprising, since gates can always be emulated by prohibitively high redemption fees that investors never pay in equilibrium.

Gates may have a role in preventing runs at stable NAV funds in case there is a regulatory upper bound on the redemption fee.[16] To see this, suppose there is an upper bound $\overline{\tau}$ so that any redemption fee τ the fund charges must satisfy $\tau \leqslant \overline{\tau}$. If $\overline{\tau} < 1 - \lambda$, then the fund cannot charge a redemption fee that is high enough to ensure that redeeming investors internalize liquidation losses. In this case, imposing a gate after a certain amount of withdrawals at date 1 is necessary (but not sufficient) for a stable NAV fund to prevent runs. If the fund never suspends payouts at date 1, it will be forced to pay out an amount that is higher than the liquidation price λ to investors that show up late in line at date 1. If all investors run on the fund at date 1, the fund will thus run out of assets before everybody showed up, giving rise to a run equilibrium á-la Diamond and Dybvig.[17]

Regulatory restrictions on the level of redemption fees can also be an issue for floating NAV funds. If regulatory restrictions on fees are such that redemption fees cannot be set within the interval given in proposition 3.3, floating NAV funds may not be able to eliminate runs with redemption fees alone. However, it is not obvious that gates are useful in this situation. In floating NAV funds, redemption fees are useful as a means to implement a redistribution among investors; the fee revenue collected from one set of investors is paid to another set of investors. Gates cannot fulfill such a redistributionary function, and I failed to come up with an example where regulatory restrictions on redemption fees are such that a floating NAV fund is unable to eliminate runs by using fees alone but can do so by using gates.

While it is hard to find an example where gates are useful in floating NAV funds, it is easy to see how gates can do harm. For instance, suppose parameters satisfy the condition in proposition 3.2

[16]The case with the regulatory upper bound is relevant because, in the US, redemption fees are limited to 2%. No such cap exists in EU MMF regulations.

[17]If the upper bound on the redemption fee is such that the fund cannot charge fees that fully compensate it for liquidation losses (i.e. $\overline{\tau} < 1 - \lambda$) but it is possible to charge redemption fees in the interval given by proposition 3.3, then the fund can uniquely implement first-best by combining gates with fees. More precisely, the fund can impose a gate whenever it runs out of liquid assets and, if the fund had to impose a gate at date 1, charge a redemption fee $\hat{\tau} \in \left(\frac{1-\delta}{2}, \frac{1-\delta}{1+\delta}\right)$ on withdrawals at date 2. Put differently, the fund can replace the 'first-stage' redemption fee in payout schedule $f^{\tau\hat{\tau}}$ with a gate. I leave out the proof of this, but it follows rather immediately from the proof of proposition 3.3. Note in particular that the result of proposition 3.3 does not rely on anybody actually paying the redemption fee at date 1 if (off the equilibrium path) a run occurs at date 1.

so that a floating NAV fund without fees and gates is *not* susceptible to runs. Suppose now the fund imposes gates, that is, it suspends payouts until the next period when liquid assets are exhausted. In terms of equilibrium outcomes, gates are equivalent to prohibitively high redemption fees that investors never pay in equilibrium. From the discussion in subsection 3.1 it follows that a floating NAV fund with gates will be susceptible to runs whenever $1 - 8\kappa \geqslant \delta$. (This is the condition for run equilibria to exist at a stable NAV fund that charges prohibitively high redemption fees once it runs out of liquid assets.) This shows that gates can lead to runs at floating NAV funds that would not have happened otherwise, in particular if investors' liquidity preference is strong (low δ).[18] The discussion in the previous paragraphs leads to the following proposition which is stated without separate proof:

Proposition 5.1.

(i) Gates do not have a role in preventing runs if there are no (regulatory or other) restrictions on the level of redemption fees that funds can charge.

(ii) Gates are necessary to prevent runs at stable NAV funds if there is an upper bound $\overline{\tau}$ on the redemption fee and the upper bound satisfies $\overline{\tau} < 1 - \lambda$.

6. Discussion of Model Assumptions

One simplifying assumption made in this paper is the abstraction from fundamental risk, both with regard to investment returns as well as with regard to investors' aggregate liquidity needs. This allows to focus on the themes of most interest for this paper while keeping the mathematical machinery to a mininum. In reality, runs often seem to be a combination of coordination failure and real investment losses, and it is hard to disentangle the two empirically (see for instance the survey in Goldstein (2013)).[19] The assumption of known liquidity needs at the fund level may be less controversial given the large size of most MMFs and the fact that liquidity shocks are plausibly

[18] The subset of the parameter space for which the condition in proposition 3.2 is fulfilled *and* $1 - 8\kappa \geqslant \delta$ (so that a fund with gates is susceptible to runs) is quite large.

[19] In terms of modelling strategy, a popular way to combine fundamental risk with coordination failure is the global games technique. Formulating the Engineer setting as a global game could be an interesting endeavour for future research. At the same time, it does not seem obvious that using global games would lead to additional qualitative insights regarding the effectiveness of fees and gates in preventing runs in an Engineer type setting, compared to the simpler approach taken here.

uncorrelated among an MMF's investors. Assuming that aggregate liquidity needs are unknown ex ante would change the analysis rather fundamentally, not least because the derivation of the benchmark allocation that is to be uniquely implemented (e.g. the best implementable allocation under a sequential service constraint) would be much more involved. An analysis of the Engineer setting with stochastic aggregate liquidity needs is left for future research.

Another simplifying assumption in this paper is that investors' preferences are linear. In the original Engineer (1989) model, ex-post preferences are of the type $u(c_1 + \delta c_2 + \delta^2 c_3)$ for type 1 investors, $u(c_2 + \delta c_3)$ for type 2 investors and $u(c_3)$ for type 3 investors, where $u(\cdot)$ is a strictly increasing and strictly concave utility function. Preferences of this form have also been used in the setting with two types, e.g. in Wallace (1988).[20] Assuming non-linear preferences of the form above would not change the analysis in a fundamental way, especially since the marginal rate of substitution between consumption at different dates is still constant. However, when facing a choice between a deterministic and an uncertain payoff, investors' choice would be tilted towards the deterministic option. The main effect of this is to increase investors' propensity to run at date 1. The decision whether or not to run at date 1 often entails a choice between a deterministic payoff that results from redeeming at date 1 and an uncertain payoff when remaining in the MMF. For instance, with redemption fees as in subsection 3.1, the ex-post payoff of a type 2 investor may depend on her position in the line at date 2. With two-stage redemption fees as in subsection 3.2, the ex-post payoff of a non-type 1 investor may be higher if the investor turns out to be of type 3 instead of type 2. This highlights that the thresholds for run equilibria to exist are sensitive to changes in the specification of investors' preferences. The advantage of using linear preferences is that it allows for a very tractable analysis of the main trade-offs involved, generating a number of qualitative insights regarding investors' propensity to run under fees and gates -clauses.

Finally, I assumed throughout the paper that the MMF can commit to its payout schedule. This assumption is important because policies that prevent run equilibria under commitment often entail measures that hurt (some) investors after (off-equilibrium) a run occurred, as highlighted by Ennis and Keister (2009a, 2010). For instance, all payout schedules studied in this paper are such that type

[20]In Diamond and Dybvig (1983), impatient consumers attach zero value to consumption in the last date. Jacklin (1987) studies a Diamond Dybvig setting with a non-constant marginal rate of substitution between consumption at different dates.

1 investors who arrive late in line in a run at date 1 receive a lower payoff than the type 3 investors left in the MMF. If the fund maximizes expected utility of all investors left in the fund at any point in time, such payout schedules may not be time consistent.

7. Conclusion

This paper derives a number of qualitative results regarding the effectiveness of fees and gates - clauses in preventing runs. The paper also shows how the Engineer (1989) model can provide a rich but still sufficiently simple framework for policy analysis. The results derived in the paper suggest that redemption fees in particular are a versatile and powerful tool to prevent runs if used correctly. Given that MMFs can charge redemption fees, the independent role of gates in preventing runs is somewhat unclear, although gates can have a role if there is a regulatory upper bound on the level of redemption fees.

Redemption fees are often treated solely as a means to ensure that liquidation costs are borne by redeeming investors. This paper suggests that the role of redemption fees should be seen more broadly; specifically, redemption fees can be used to implement a redistribution among investors in such a way as to incentivize investors to remain in the fund when others run. One question is to which extent current MMF regulations in the US and the EU would actually allow to use redemption fees in this manner. Notably, EU regulations state that redemption fees should "adequately reflect the cost to the MMF of achieving liquidity".[21] This could be an issue because, in order to implement a redistribution among investors, redemption fees need to be *higher* than liquidation costs.

A curious difference between US and EU MMF regulations is that, in the US, the fees and gates provisions are geared towards MMFs with floating NAV while in the EU they are geared towards MMFs with stable NAV. The results in this paper suggest that liquidity management tools such as fees and gates are necessary to prevent runs on stable NAVs (unless assets are perfectly liquid) but may not be necessary to prevent runs on floating NAVs. This would suggest that the fees and gates provisions should be geared towards stable NAVs as in the EU. Nevertheless, the discrepancy between US and EU regulations may be partly explained by differences in the liquidity of assets

[21] Quoted from article 34(1) of Regulation (EU) No 1131/2017.

held by US and EU stable NAVs. US stable NAVs need to be fully invested in US government assets which can plausibly be regarded as perfectly liquid ($\lambda = 1$) so that US stable NAV MMFs are not susceptible to runs even without fees and gates.[22] In the EU, stable NAVs can invest in European government debt which may not be perfectly liquid; furthermore, they can invest in private assets and still offer something similar to stable NAV by operating as Low Volatility NAVs (which are also subject to the fees and gates provisions).

As a final note, the new fees and gates -regulations have been 'put to the test' for the first time in the US in the recent covid-19 crisis. The somewhat sobering conclusion is that money market mutual funds seemed to be extremely reluctant to impose fees or gates despite massive redemptions and depletion of their liquid assets (Eren et al. (2020)). The Fed eventually provided (indirect) liquidity support to money market mutual funds. One interpretation of these events is that they point to serious issues of time consistency regarding the new fees and gates regulations.

[22]Indeed, US government MMFs have remained stable during the financial crisis (Schmidt et al. (2016)) as well as the recent covid-19 crisis (Eren et al. (2020)).

References

Al-Najjar, Nabil (2008) "Large Games and the Law of Large Numbers," *Games and Economic Behavior*, Vol. 64(1), pp. 1–34, DOI: http://dx.doi.org/10.1016/j.geb.2007.11.002.

Chen, Qi, Itay Goldstein, and Wei Jiang (2010) "Payoff complementarities and financial fragility: Evidence from mutual fund outflows," *Journal of Financial Economics*, Vol. 97(2), pp. 239–262.

Cipriani, Marco, Antoine Martin, Patrick McCabe, and Bruno M. Parigi (2014) "Gates, Fees and Preemptive Runs," Federal Reserve Bank of New York Staff Reports No. 670.

Diamond, Douglas W. and Philip H. Dybvig (1983) "Bank Runs, Deposit Insurance and Liquidity," *Journal of Political Economy*, Vol. 91(3), pp. 401–419, DOI: http://dx.doi.org/10.1086/261155.

Engineer, Merwan (1989) "Bank Runs and the Suspension of Deposit Convertibility," *Journal of Monetary Economics*, Vol. 24, pp. 443–454, DOI: http://dx.doi.org/10.1016/0304-3932(89)90031-7.

Ennis, Huberto and Todd Keister (2009a) "Bank Runs and Institutions: The Perils of Intervention," *American Economic Review*, Vol. 99(4), pp. 1588–1607, DOI: http://dx.doi.org/10.1257/aer.99.4.1588.

——— (2009b) "Run Equilibria in the Green-Lin Model of Financial Intermediation," *Journal of Economic Theory*, Vol. 144(5), DOI: http://dx.doi.org/10.1016/j.jet.2009.05.001.

——— (2010) "Banking Panics and Policy Responses," *Journal of Monetary Economics*, Vol. 57(4), pp. 404–419, DOI: http://dx.doi.org/10.1016/j.jmoneco.2010.04.005.

Eren, Egemen, Andreas Schrimpf, and Vladyslav Sushko (2020) "US Dollar Funding Market during the Covid-19 Crisis - the Money Market Funds Turmoil," BIS Bulletin No 14.

Gertler, Mark and Nobuhiro Kiyotaki (2015) "Banking, Liquidity, and Bank Runs in an infinite horizon economy," *American Economic Review*, Vol. 105(7), pp. 2011–43.

Goldstein, Itay (2013) "Empirical Literature on Financial Crises: Fundamentals vs. Panic," in Gerard Caprio ed. *The Evidence and Impact of Financial Globalization*: Elsevier, Amsterdam, pp. 523–534.

Green, Edward J. and Ping Lin (2003) "Implementing Efficient Allocations in a Model of Financial Intermediation," *Journal of Economic Theory*, Vol. 109(1), pp. 1–23, DOI: http://dx.doi.org/10.1016/S0022-0531(02)00017-0.

He, Zhiguo and Wei Xiong (2012) "Dynamic Debt Runs," *The Review of Financial Studies*, Vol. 25(6), pp. 1799–1843, DOI: http://dx.doi.org/10.1093/rfs/hhs004.

Jacklin, Charles J. (1987) "Demand Deposits, Trading Restrictions, and Risk Sharing," in Edward C. Prescott and Neil Wallace eds. *Contractual Arrangements for Intertemporal Trade*: Minneapolis: Univ. Minnesota Press.

Lenkey, Stephen and Fenghua Song (2016) "Redemption Fees and Information-Based Runs," Available at SSRN: https://ssrn.com/abstract=2796717.

Morgan Stanley (2019) "European Money Market Fund Regulatory Update," URL: https://www.morganstanley.com/im/publication/insights/regulatory/regulatory_mmfregulatoryupdate122017_us.pdf, (accessed October 3 2020).

Peck, James and Karl Shell (2003) "Equilibrium Bank Runs," *Journal of Political Economy*, Vol. 111(1), pp. 103–123, DOI: http://dx.doi.org/10.1086/344803.

Schmidt, Lawrence, Allan Timmermann, and Russ Wermers (2016) "Runs on Money Market Mutual Funds," *American Economic Review*, Vol. 106(9), pp. 2625–57, DOI: http://dx.doi.org/10.1257/aer.20140678.

SEC (2014) "SEC Adopts Money Market Fund Reform Rules.," SEC Press Release 2014-143, URL: https://www.sec.gov/news/press-release/2014-143, (accessed June 27 2020).

Stein, Jeremy C. (2012) "Monetary Policy as Financial Stability Regulation," *The Quarterly Journal of Economics*, Vol. 127(1), pp. 57–95.

Wallace, Neil (1988) "Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig Model With Sequential Service Taken Seriousiy," *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. Fall, pp. 3–15.

Zeng, Yao (2017) "A Dynamic Theory of Mutual Fund Runs and Liquidity Management," European Systemic Risk Board Working Paper No 42 / 2017.

Appendix

A. Proof of Lemma 3.1

First, it is useful to note that any strategy profile where all investors' strategies satisfy the following implements first-best:

$$\begin{aligned}
&\text{Date 1:} \qquad s_1(1,1) = \text{‘withdraw’} \quad s_1(not1,\cdot) = \text{‘not withdraw’} \\
&\text{Date 2:} \qquad s_2(2,1) = \text{‘withdraw’} \quad s_2(3,\cdot) = \text{‘not withdraw’}
\end{aligned} \tag{9}$$

The "only if" part in lemma 3.1 follows directly from the definition of first-best. To prove the "if"-part, we restrict attention to strategy profiles where non-type 1 investors play $s_1(\text{not}1, 1) = s_1(\text{not}1, \lambda) =$ ‘not withdraw’ with probability one and then eliminate strictly dominated strategies. First, withdrawing is the strictly dominant strategy for type 1 investors at date 1 (playing $s_1(1, 1) =$ ‘withdraw’ is strictly dominant). Furthermore, at date 2, withdrawing one unit is the strictly dominant strategy for type 2 investors (playing $s_2(2, 1) =$‘withdraw’ is strictly dominant). Since type 3 investors always receive 1 unit at date 3, not withdrawing at date 2 is the strictly dominant strategy for type 3 investors (playing $s_2(3, 1) = s_2(3, \lambda) =$‘not withdraw’ is strictly dominant). Given that we impose $s_1(\text{not}1, 1) =$‘not withdraw’, all strategy profiles that survive elimination of strictly dominated strategies therefore satisfy (9) and implement first-best. ■

B. Proof of Proposition 3.2

In order to proof proposition 3.2, it is useful to distinguish between three cases:

<u>Case 1: $\lambda \geqslant \delta$</u>

With $\lambda \geqslant \delta$, type 1 investors (weakly) prefer consuming λ units at date 1 to consuming 1 unit at date 2. This means that type 1 investors arriving late in line at date 1 prefer to withdraw at date 1 even if $P = 1$, that is, even if they know that they would be able to withdraw one unit at date 2 with certainty. In this case, we must have $P = 1$ in equilibrium, which can be shown with a proof by contradiction: Suppose $P < 1$. Since $\lambda > Y(P)_{[\lambda \geqslant \delta]}$ for any $P < 1$, type 1 investors are then strictly better of withdrawing λ units at date 1, which implies $B \leqslant \frac{1}{3}$ and $P = 1$ so that we arrive at a contradiction. Finally, since $Z(1)_{[\lambda \geqslant \delta]} > 1 - \kappa$, there is no equilibrium in which non-type 1 investors run at date 1. Combined with lemma 3.1 this leads to item (i) in proposition 3.2.

Case 2: $\lambda \in \left(\frac{3}{4}\delta + \frac{1}{4}\delta^2,\, \delta\right)$

As discussed in the main text, we have $P \in \left[\frac{3}{4}, 1\right]$ in equilibrium. We get that:

$$\lambda \in \left(\tfrac{3}{4}\delta + \tfrac{1}{4}\delta^2,\, \delta\right) \;\Leftrightarrow\; Y(\tfrac{3}{4})_{[\lambda<\delta]} < \lambda < Y(1)_{[\lambda<\delta]} \tag{10}$$

If condition (10) is fulfilled, then investors prefer *not* to pay the redemption fee at date 1 if they know that they can withdraw 1 unit with certainty at date 2 ($P = 1$); however, they *do* prefer to pay the redemption fee at date 1 if the probability that they can withdraw 1 unit at date 2 is at its lower bound ($P = \frac{3}{4}$). Consider now a scenario where all investors run at date 1 (investors play $s_1(\cdot, 1)$= 'withdraw'). As discussed in the main text, we have $B = \frac{4}{9}$ (and $P = \frac{3}{4}$) if none of the type 1 investors arriving late in the queue (those with indices $i_1 > \frac{1}{3}$) pay the redemption fee and withdraw at date 1. However, according to expression (10), type 1 investors then strictly prefer paying the fee at date 1. Conversely, if all type 1 investors pay the fee at date 1 (implying $B = \frac{1}{3}$ and $P = 1$) then, according to expression (10), type 1 investors strictly prefer not to pay the fee. It follows that, in the run scenario considered above, some fraction of the type 1 investors arriving late in line at date 1 will pay the redemption fee at date 1. The fraction of type 1 investors paying the redemption fee at date 1 will be such that type 1 investors are just indifferent between paying the fee at date 1 and not paying the fee. The larger the fraction of type 1 investors who pay the fee at date 1, the lower B. Since $Y(P(B))$ is continuous and strictly decreasing in B, there is a unique B^* such that $Y(P(B^*))_{[\lambda<\delta]} = \lambda$. To solve for B^*, we use expression (4) together with the fact that $P = \min\{\frac{1}{3B}, 1\}$, which gives $B^* = \dfrac{1}{3}\left(\dfrac{\delta - \delta^2}{\lambda - \delta^2}\right)$. Note that, since $B \leqslant B^*$ in equilibrium, we have $P \geqslant P(B^*) = \dfrac{\lambda - \delta^2}{\delta - \delta^2}$ in equilibrium. Consider now the best response of non-type 1 investors in the run scenario discussed above. Given that everybody else runs at date 1, running is the best response of each individual non-type 1 investor iff:

$$
\begin{aligned}
&Z(P(B^*))_{[\lambda<\delta]} \leqslant 1-\kappa \\
\Leftrightarrow\ & \frac{1}{2}\left[P(B^*)+\delta(1-P(B^*))\right]+\frac{1}{2} \leqslant 1-\kappa \\
\Leftrightarrow\ & \frac{1}{2}\frac{\lambda-\delta^2}{\delta-\delta^2}+\frac{\delta}{2}\left(1-\frac{\lambda-\delta^2}{\delta-\delta^2}\right)+\frac{1}{2} \leqslant 1-\kappa \\
\Leftrightarrow\ & \frac{\lambda}{\delta} \leqslant 1-2\kappa
\end{aligned} \tag{11}
$$

It follows that the withdrawal game exhibits an equilibrium where all investors run on the fund at date 1 whenever condition (11) is satisfied. The converse is also true: if condition (11) is not satisfied, the withdrawal game does not exhibit a run equilibrium, and first-best is uniquely implemented. To see this, recall first that $B \leqslant B^*$ in equilibrium, which implies that $Z(\cdot)$ never falls below $Z(P(B^*))$ in equilibrium. If condition (11) is not fulfilled, we thus have that $Z(\cdot) > 1-\kappa$ in equilibrium, which means that there is no equilibrium in which non-type 1 investors run at date 1. Combined with lemma 3.1, this leads to item (ii) in proposition 3.2.

Case 3: $\lambda \leqslant \frac{3}{4}\delta + \frac{1}{4}\delta^2$

Proof contained in the main text.

C. Proof of Proposition 3.3

We start with the following result which shows that, when looking for run equilibria, we can limit attention to equilibria where non-type 1 investors withdraw at date 1:

Lemma C.1. *The withdrawal game under payout schedule $f^{\tau\hat{\tau}}$ (with some fee $\hat{\tau} \geqslant 0$) uniquely implements first-best if and only if the withdrawal game does not exhibit an equilibrium in which non-type 1 investors withdraw at date 1.*

The proof of lemma C.1 is very similar to the proof of lemma 3.1 and is left out. The following result shows that, as long as the date 2 redemption fee $\hat{\tau}$ is high enough, withdrawals at date 2 (after a run occurred at date 1) are always strictly less than the fund's date 2 liquid assets:

Lemma C.2. *Suppose* $\hat{\tau} > \frac{1-\delta}{2}$. *Then there is no equilibrium in which non-type 1 investors run at date 1 and total withdrawals at date 2 satisfy* $\overline{z}_2 \geqslant \frac{1}{3(1-\hat{\tau})}$.

Proof of Lemma C.2: If non-type 1 investors run at date 1, the fund imposes the redemption fee $\hat{\tau}$ on date 2 withdrawals. Note first that, since type 3 investors always receive at least 1 unit at date 3 under payout schedule $f^{\tau\hat{\tau}}$, only type 1 and 2 investors may potentially withdraw at date 2. Furthermore, type 1 and 2 investors in the fund at date 2 strictly prefer *not* to withdraw at date 2 if $1 - \hat{\tau} < \delta c_3^{\hat{\tau}}(\cdot)$, where $c_3^{\hat{\tau}}$ are date 3 payouts given by expression (8). Note further that date 3 payouts $c_3^{\hat{\tau}}(\overline{z}_1, \overline{z}_2)$ are strictly increasing in $\overline{z}_1$ and $\overline{z}_2$. Since withdrawing at date 1 is a strictly dominant strategy for type 1 investors, we have $\overline{z}_1 \geqslant \frac{1}{3}$ in equilibrium. If follows that date 3 payouts are at least $c_3^{\hat{\tau}}(\frac{1}{3}, \overline{z}_2)$ for given $\overline{z}_2$. Some algebra yields that $1 - \hat{\tau} < \delta c_3^{\hat{\tau}}(\frac{1}{3}, \frac{1}{3(1-\hat{\tau})})$ is equivalent to $\hat{\tau} > \frac{1-\delta}{2}$. We can now proof lemma C.2 with a proof by contradiction. Suppose the date 2 redemption fee is activated and $\overline{z}_2 \geqslant \frac{1}{3(1-\hat{\tau})}$. Then $1 - \hat{\tau} < \delta c_3^{\hat{\tau}}(\cdot)$, so that all type 1 and 2 investors strictly prefer not to withdraw at date 2, which implies $\overline{z}_2 = 0$ and we arrive at a contradiction. ■

The next result shows that, if $\hat{\tau} > \frac{1-\delta}{2}$ and at least half of the investors in the fund at date 2 pay the redemption fee at date 2 after (off the equilibrium path) a run occurred at date 1, then the fee revenue collected at date 2 is large enough to take away the incentive for non-type 1 investors to run at date 1:

Lemma C.3. *Suppose* $\hat{\tau} > \frac{1-\delta}{2}$. *Then there is no equilibrium in which non-type 1 investors run at date 1 and total withdrawals at date 2 satisfy* $\overline{z}_2 \geqslant \frac{1-\overline{z}_1}{2}$.

Proof of Lemma C.3: If non-type 1 investors run at date 1, the fund imposes the redemption fee $\hat{\tau}$ on date 2 withdrawals. From lemma C.2 we know that, given $\hat{\tau} > \frac{1-\delta}{2}$, a type 2 investor in the fund at date 2 (after a run occurred at date 1) can always withdraw $1 - \hat{\tau}$ units at date 2 and thus will get a payoff of at least $1 - \hat{\tau}$.[23] Consider now the best response of a non-type 1 investor at date 1. Since a non-type 1 investor turns out to be of type 2 and 3 with probability $\frac{1}{2}$ each, the expected payoff of *not* running at date 1 equals $\frac{1}{2}(1 - \hat{\tau}) + \frac{1}{2}c_3^{\hat{\tau}}(\cdot)$. Withdrawing 1 unit at date 1

[23]This does not mean that after a run occurred at date 1 (off the equilibrium path) all type 1 and 2 investors in the fund will withdraw $1 - \hat{\tau}$ units at date 2. Since date 3 payouts $c_3^{\hat{\tau}}$ are increasing in date 2 withdrawals $\overline{z}_2$, it is possible that only a fraction of type 1 and 2 investors in the fund withdraw at date 2, with total date 2 withdrawals $\overline{z}_2$ being such that $1 - \hat{\tau} = \delta c_3^{\hat{\tau}}(\overline{z}_1, \overline{z}_2)$, so that type 1 and 2 investors are just indifferent between withdrawing $1 - \hat{\tau}$ units at date 2 and receiving $c_3^{\hat{\tau}}(\cdot)$ units at date 3.

gives a payoff of $1-\kappa$ for non-type 1 investors. Therefore, a sufficient condition for non-type 1 investors *not* to run at date 1 is that the expected payoff of not running is weakly larger than 1, that is, $\frac{1}{2}(1-\widehat{\tau})+\frac{1}{2}c_3^{\widehat{\tau}}(\overline{z}_1,\overline{z}_2) \geqslant 1$. Some algebra yields that this is equivalent to $\overline{z}_2 \geqslant \frac{1-\overline{z}_1}{2}$. Given $\overline{z}_2 \geqslant \frac{1-\overline{z}_1}{2}$, non-type 1 investors thus strictly prefer *not* to run at date 1, which leads to the result in lemma C.3. ■

Finally, the next result shows that, as long as the fee $\widehat{\tau}$ is not too high, at least half of the investors in the fund at date 2 will pay the fee $\widehat{\tau}$ after (off the equilibrium path) a run occurred at date 1:

Lemma C.4. *Suppose $\widehat{\tau} < \frac{1-\delta}{1+\delta}$. Then there is no equilibrium in which non-type 1 investors run at date 1 and total withdrawals at date 2 satisfy $\overline{z}_2 < \frac{1-\overline{z}_1}{2}$.*

Proof of Lemma C.4: If non-type 1 investors run at date 1, the fund imposes the redemption fee $\widehat{\tau}$ on date 2 withdrawals. Type 1 and 2 investors in the fund at date 2 strictly prefer to pay the redemption fee at date 2 (instead of being paid out at date 3) if $1-\widehat{\tau} > \delta c_3^{\widehat{\tau}}(\overline{z}_1,\overline{z}_2)$. Some algebra yields that $1-\widehat{\tau} > \delta c_3^{\widehat{\tau}}(\overline{z}_1,\frac{1-\overline{z}_1}{2})$ is equivalent to $\widehat{\tau} < \frac{1-\delta}{1+\delta}$. Since $c_3^{\widehat{\tau}}(\overline{z}_1,\overline{z}_2)$ is strictly increasing in $\overline{z}_2$, this means that, for any $\overline{z}_2 < \frac{1-\overline{z}_1}{2}$, all type 1 and 2 investors strictly prefer to withdraw at date 2. We can now proceed with a proof by contradiction. Suppose the date 2 redemption fee is activated and $\overline{z}_2 < \frac{1-\overline{z}_1}{2}$. Then all type 1 and 2 investors in the fund at date 2 are better off withdrawing at date 2. Since at least half of investors in the fund at date 2 are of either type 1 or 2[24] this implies $\overline{z}_2 \geqslant \frac{1-\overline{z}_1}{2}$, so that we get a contradiction. ■

From lemmas C.3 and C.4, it follows that there is no equilibrium where non-type 1 investors run at date 1, as long as the date 2 redemption fee $\widehat{\tau}$ is within the (non-empty) interval $\left(\frac{1-\delta}{2},\frac{1-\delta}{1+\delta}\right)$. Together with lemma C.1 this completes the proof of proposition 3.3.

[24]If all type 1 investors withdraw at date 1, half of the investors in the fund at date 2 will be of type 2, the other half of type 3. If some type 1 investors did not withdraw at date 1 (and are hence still in the fund at date 2), the combined fraction of type 1 and 2 investors in the fund at date 2 will be strictly more than half.