# Can Redemption Fees Prevent Runs on Funds?

Xuesong Huang
*Lingnan College*
*Sun Yat-sen University*

Todd Keister
*Rutgers University*

September 2024

# Motivation

- Recurrent phenomenon: runs on banks and related institutions
  - Spring 2023: *Silicon Valley Bank (SVB),* Signature, First Republic
  - Spring 2020: Money Market Mutual Funds (MMFs)
  - Fall 2008: investment banks, repo markets, MMFs, many more
- Much discussion and policy reforms on how to prevent runs
  - government guarantees, lender of last resort, capital requirements, liquidity regulation, etc.
- This paper looks at one approach: redemption fees
  - adjust payments based on redemption/withdrawal demand
  - recent reforms to MMFs in the U.S. provide a concrete laboratory
    - but the ideas potentially apply much more broadly

# Runs on MMFs

- Sept. 2008: runs on institutional prime MMFs
- July 2014: SEC modified the rules governing these MMFs
  - allowed to impose gates and redemption fees …
  - … when a fund's ratio of liquid to total assets falls below a threshold
- Interpretation: allow funds to operate as usual in normal times
  - but react to "unusually" high redemption demand
  - hope to put these events *off the equilibrium path of play*
- March 2020: runs on institutional prime funds again

$\Rightarrow$ the 2014 reform was ineffective

# Recent reforms

- July 2023: SEC finalized new rules
  - removed the liquid-asset threshold and the option to use gates
  - impose fees based on *current redemption demand*

  "*A mandatory fee is charged to redeeming investors when the fund has net redemptions above 5% of net assets*."

- Interpretation: apply redemption fees more often
  - on the equilibrium path (when no run is occurring)

  "*We estimate that an average of 3.2% of institutional funds would cross a 5% net redemption threshold on a given day.*"

- Will the new reform work?  What is the optimal fee policy?
  - our take: practice is ahead of theory on this issue

# This paper

- Develop a model to study MMF redemption-fee rules
- Show: using fees only in extraordinary times (~2014 reform) is often ineffective
  - model captures the problems that occurred in March 2020
- Derive the best **run-proof** fee policy
  - aim to understand the general principles for optimal policy
  - and find the best robust, simple (~2023 reform) policy
- Evaluate the 2023 reform
  - vulnerable when market liquidity may worsen
  - best policy has **smaller** fee that applies **more** often

# Related literature

- Existing models of *preemptive* bank runs
  - Engineer (1989), Cipriani et al. (2014), Voellmy (2021)
- Runs on MMFs and patterns of redemptions at mutual funds more broadly
  - Chen et al. (2010), Schmidt et al. (2016), Parlatore (2016), Goldstein et al (2017), Zeng (2017), Cipriani & La Spada (2020), Alvados & Xia (2021), Jin et al. (2022), Li et al. (2021), and others
- Policy papers on MMF reform
  - Ennis (2012), McCabe et al. (2013), President's Working Group Report (2020), Ennis, Lacker and Weinberg (2023), and others
- Our contribution: if the goal is to prevent runs …
  - what *principles* should determine MMF redemption fees?

# Outline

1) Model

2) Preemptive runs

- failure of the 2014 reforms

3) Run-proof policies

- using redemption fees in normal times as well
- evaluating the 2023 reforms

4) Concluding remarks

# Environment

- Investors: $i \in [0,1]$
  - endowed with one unit of good at $t = 0$, nothing later
- $t = 0,1,2,3$
- Technologies:
  - storage yields gross return of 1 in any period
  - investment at $t = 0$ yields: $\begin{Bmatrix} r_1 < 1 \\ r_2 < 1 \\ R > 1 \end{Bmatrix}$ at $\begin{Bmatrix} t = 1 \\ t = 2 \\ t = 3 \end{Bmatrix}$
    - $R$ is known
    - $r_2$ may be random
- Utility: $\begin{Bmatrix} u(c_1) \\ u(c_1 + c_2) \\ u(c_1 + c_2 + c_3) \end{Bmatrix}$ if investor is $\begin{Bmatrix} \text{type 1} \\ \text{type 2} \\ \text{patient} \end{Bmatrix}$ ← "impatient" (type 1 and type 2)
  - focus on: $u(c) = \ln(c)$

# Key assumptions

- Fraction of impatient investors (types 1 & 2) is known: $\pi$
- Fraction of type 1 investors is random: $\pi_1 \sim F[0, \pi]$
  - no uncertainty about *total* early redemption demand
  - but uncertainty about the *timing* of that demand
- Investors learn their type gradually
  - at $t = 1$, only learn whether or not they are type 1
- A fraction $\delta \in (0,1]$ of non-type 1 investors can redeem at $t = 1$
  - the remaining $1 - \delta$ are inattentive ("don't see the sunspot")
  - role: limits size of a potential run in period 1
  - assume $\delta$ is known (for now)

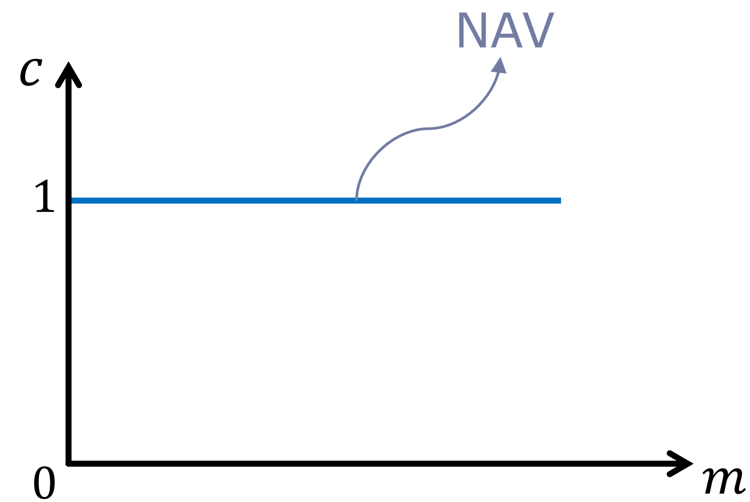
# Efficient allocation and Fund

- A planner with full information would set:
  - $c_1 = c_2 = 1$ and $c_3 = R$ (same efficient allocation as in DD)
  - Planner's portfolio: $\pi$ in storage, $(1-\pi)$ invested
- A *mutual fund* offers the following contract to investors:
  - follow planner's portfolio at $t = 0$: $(\pi, 1-\pi)$
  - allow investors to choose when to redeem ($\Rightarrow$ a game)
  - at $t = 1,2$ : fund observes redemption demand $m_t$
    - then pays all redeeming investors
  - contract specifies payment rules:
    - $c_1(m_1)$
    - $c_2(m_1, m_2)$
    - $c_3(m_1, m_2)$

no sequential service within a period

# MMFs without fees

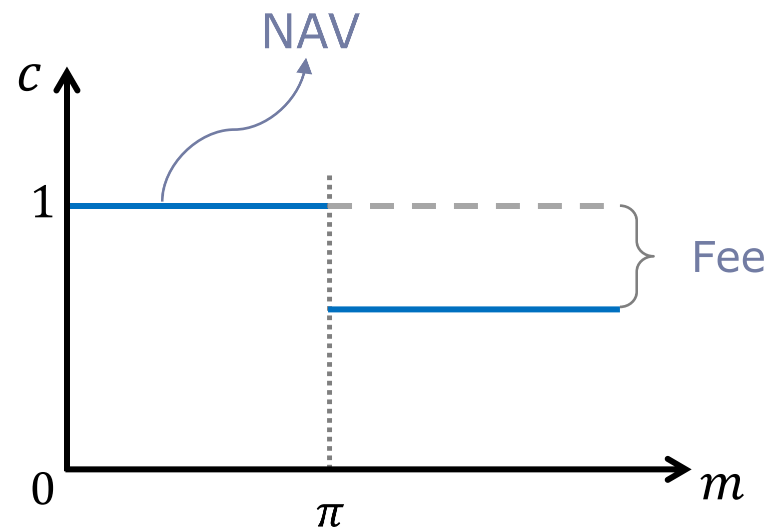
- Consider a MMF that adopts the following payment rule:
  - at $t = 1{,}2$, pay redeeming investors at net asset value (=1)
    - no uncertainty on R ⇒ NAV always =1
  - at $t = 3$, pay investors a pro-rata share of the remaining assets
    - No run ⇒ dividend $= R - 1$




- Like a demand deposit contract in banks
  - payments independent of redemption demand
  - can achieve efficient allocation in the no-run equilibrium
  - but, subject to the classical DD runs…

# Fee in extraordinary times

- Now suppose fund imposes a redemption fee …
  - … when net outflow is extraordinary ($m > \pi$)




- Goal: still achieve the efficient allocation
  - fee lies off equilibrium if no run
- Q: How to set the fee? $100\%$?
- Require the contract to satisfy *time consistency*
  - if redemption demand indicates a run …
  - … redemption fee must be ex-post efficient
    - in the spirit of Ennis and Keister (2009, 2010)

details

# Outline

1) Model

2) Preemptive runs

- failure of the 2014 reforms

3) Run-proof policies

- using redemption fees in normal times as well
- evaluating the 2023 reforms

4) Concluding remarks

# Preemptive run

- Suppose a non-type 1 investor expects a run in period 1
- Compares the expected utility of:

$m_1 = \pi_1 + \delta(1 - \pi_1)$

$m_2 = (1 - \delta)(\pi - \pi_1)$

redeem: $\int_0^{\pi} u\big(c_1(m_1)\big) f_n(\pi_1) d\pi_1$

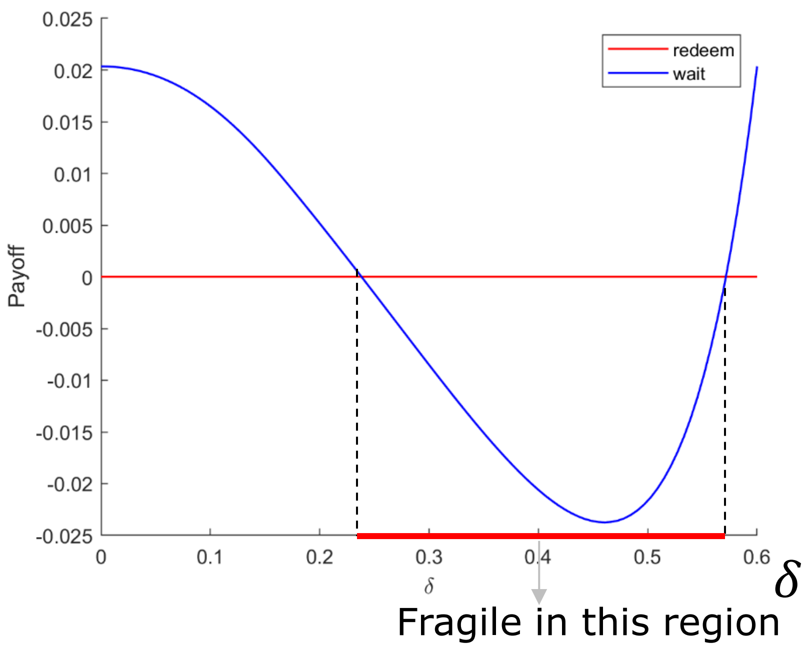
wait: $\int_0^{\pi} \big[p_n u\big(c_2(m_1, m_2)\big) + (1 - p_n) u\big(c_3(m_1, m_2)\big)\big] f_n(\pi_1) d\pi_1$

- If $\pi_1$ is large enough: $m_1 > \pi$ and run is detected immediately
  - fee imposed at $t = 1$ → no incentive to join the run
- Worry: if $\pi_1$ is small, run will not be detected until $t = 2$
  - a fee will be imposed then – and I might need to redeem
  - generates an incentive to redeem preemptively (today)

quotes

# An example

- Compare $EU(wait)$ and $EU(redeem)$ as $\delta$ varies




- Run equilibrium tends to exist …
  - … when $\delta$ is moderate
- When $\delta$ is large, a run is likely detected by the fund at $t = 1$
  - fee applied at $t = 1$ (and $t = 2$)
  - no incentive to redeem early
- When $\delta$ is small, a run is small $\Rightarrow$ fund is in good shape
- In between: a moderate-sized run may initially go undetected
  - in this region: incentive to redeem before the fee is imposed

# Implications

- Recall the key idea of 2014 reforms:
  - pay investors at NAV ($c_1 = 1$) in normal times (when $m \leq \pi$)
  - impose fees (or gates) in extraordinary cases (when $m > \pi$)
- This type of policy is susceptible to preemptive runs
  - due to the fear of getting stuck in the fees or gates
  - aligns with events in March 2020
    - President's WG on Financial Markets (2020), SEC (2022)
- Takeaway: When a run evolves gradually over time
  - fees only used in extraordinary cases may have a delayed reaction…
  - … which gives rise to the incentive to run preemptively

# Outline

1) Model

2) Preemptive runs

- failure of the 2014 reforms

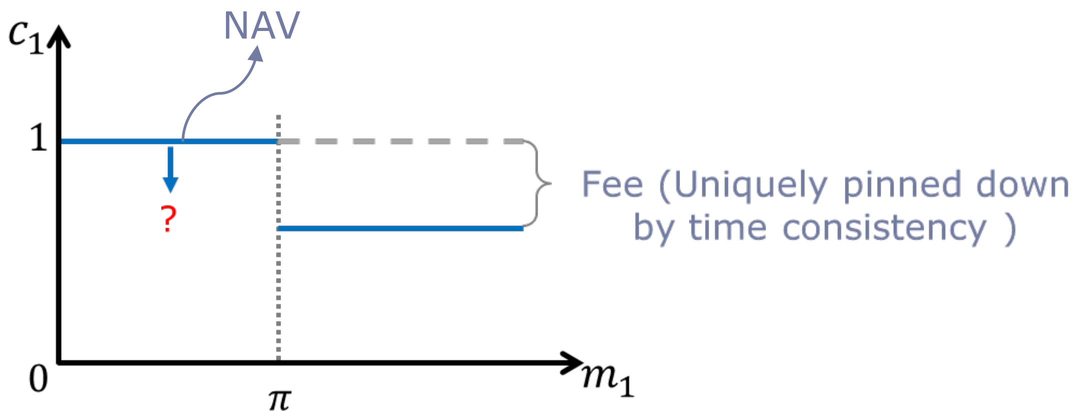
3) Run-proof policies

What should policymakers do?

- using redemption fees in normal times as well
- evaluating the 2023 reforms

4) Concluding remarks

# Fees in normal times




Choose the payment $c_1(m_1) \leq 1$ for $m_1 \leq \pi$ to solve:

expected utility with no run

$$\max_{\{c_1(m_1)|m_1 \leq \pi\}} \int_0^{\pi} \left\{ \begin{array}{c} \pi_1 u\big(c_1(\pi_1)\big) + (\pi - \pi_1) u\big(c_2(\pi_1, \pi_2)\big) \\ +(1-\pi) u\big(c_3(\pi_1, \pi_2)\big) \end{array} \right\} f(\pi_1) d\pi_1$$

- subject to the run-proof constraint:

redeem

$$\int_0^{\pi} u\big(c_1(m_1)\big) f_n(\pi_1) d\pi_1 \leq$$

if I expect all others to run

wait

$$\int_0^{\pi} \left[p_n u\big(c_2(m_1, m_2)\big) + (1 - p_n) u\big(c_3(m_1, m_2)\big)\right] f_n(\pi_1) d\pi_1$$

# Optimal fee

- Observation: The size of a run at $t = 1$ is at least $\delta$
  - when $m_1 < \delta$, fund is sure there is no run $\Rightarrow$ no fee imposed
- When $m_1 \in [\delta, \pi]$, optimal fee depends on:

likelihood of $m_1$:

conditional on no run

conditional on a run (if I am non-type 1)

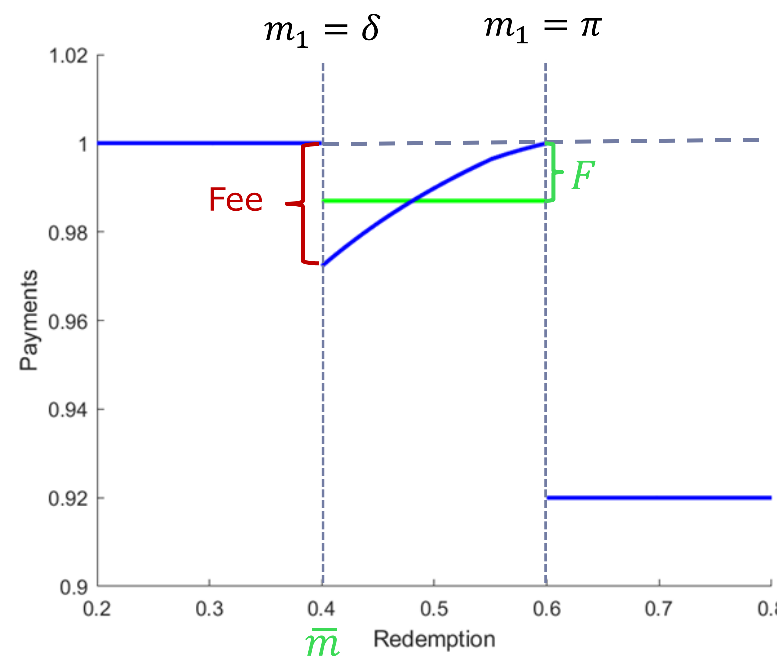
number of investors who pay the fee

$$\frac{f(m_1) \cdot m_1}{f_n\left(\frac{m_1 - \delta}{1 - \delta}\right)} \approx \frac{\text{decrease in objective}}{\text{benefit in meeting the constraint}}$$

- The optimal fee tends to <u>decrease</u> in $m_1$ in this region
  - intuition: costly to impose fees when many investors (truly) need the liquidity

# Back to the example

- Optimal fee:
  - positive for all $m_1 \geq \delta$
  - smaller than time-consistent fee
  - responsive to redemption demand
- May be difficult to implement in practice!




- Can restrict to "simple" policies (like 2023 rules)
  - zero fee below a threshold $\bar{m}$ and a constant fee $F > 0$ above it
  - time consistent fee for $m > \pi$ is unchanged
- Optimal simple policy is often $\bar{m} = \delta$ and $F$ is the "average"

# Robust simple policy

- So far, the optimal policy relies on knowing $\delta$ and $r_2$
  - which can easily change over time; difficult to monitor
- Robust approach: policy must be run-proof for …
  - all $\delta \in [0,1]$ and for $r_2 = \underline{r}$ (the worst case)
- A robust policy must guard against two possibilities:
  - large run $(\bar{\delta})$ $\Rightarrow$ fee very likely gets triggered
    - set fee high enough to remove incentive to run in this case
  - small run $(\underline{\delta})$ $\Rightarrow$ fee may or may not be triggered
    - set threshold so fee is imposed "often enough" in this case
- Optimal $(\bar{m}^*, c^*)$: run-proof condition is binding at $\bar{\delta}$ and $\underline{\delta}$

# Outline

1) Model

2) Preemptive runs

- failure of the 2014 reforms

3) Run-proof policies

- using redemption fees in normal times as well
- evaluating the 2023 reforms

4) Concluding remarks

# 2023 reforms

- July 2023 rules require:
  - one threshold: $\bar{m} = 5\%$
  - "vertical slice rule" to determine the fee:

"The size of the fee generally is determined by … costs the fund would incur if it were to *sell a pro rata amount of each security* in its portfolio to satisfy the amount of net redemptions.

- in our model (constant fee):

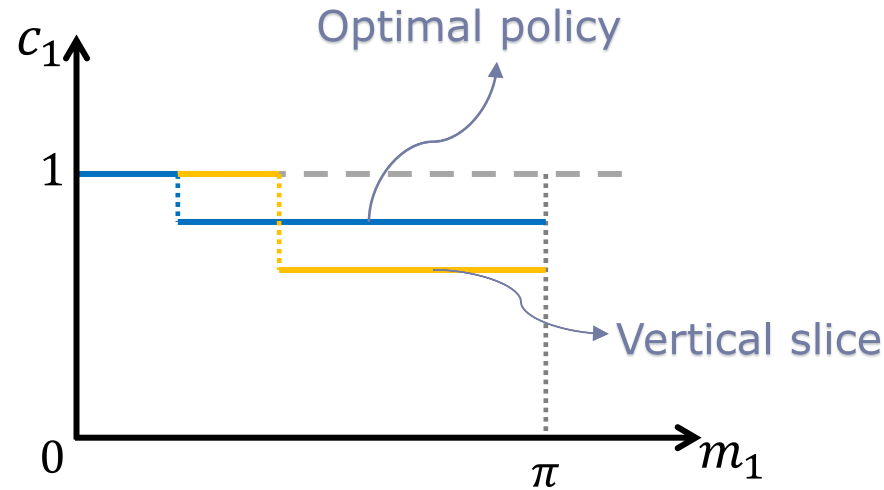
$$F = (1 - r_1)(1 - \pi) \quad \text{for } m_1 \in [\bar{m}, \pi]$$

- key: depends on the current liquidation value

- One vulnerability: not forward looking
  - $r_1 \approx 1$ and $r_2$ may get worse $\Rightarrow F = 0 \Rightarrow$ Not run-proof

# Evaluation

- When $r_1 = r_2 = \underline{r} < 1$
  - vertical slice rule is run-proof in this situation, but …
  - it is too harsh
    - large fee when many investors need to redeem




- Optimal fee is smaller
  - may require threshold to be smaller
  - fee is imposed more often, but fewer on investors
- Welfare cost of the vertical slice rule ↑ as $\underline{r}$ decreases

# Outline

1) Model

2) Preemptive runs

- failure of the 2014 reforms

3) Run-proof policies

- using redemption fees in normal times as well
- evaluate the 2023 reforms

4) Concluding remarks

# Concluding remarks

- We provide a framework for studying MMF redemption fees
- Indicates: 2023 rules are a step in the right direction, but…
  - vulnerable when market conditions may deteriorate
  - should impose a smaller fee, more often
- We also propose measures to implement our best robust policy
  - partial slice instead of vertical slice
- The framework could apply to other mutual funds as well:
  - prime example: corporate bonds funds
    - SEC has proposed “swing pricing” for these funds (~fees)
    - key difference: riskier assets → could be added to our model

# Appendix

# Time consistency in period 2

- What information does the fund have in period 2?
  - redemption demand in periods 1 and 2: $(m_1, m_2)$
  - remaining portfolio: $(s_2, i_2)$
- The time-consistent allocation $(c_2, c_3)$ solves

$$\max_{\{c_2, c_3\}} \quad m_2 u(c_2) + (1 - m_1 - m_2) u(c_3)$$

$$m_2 c_2 + e_2 = s_2 + r\ell_2 \qquad \mathrm{e}_2 \geq 0$$

$$(1 - m_1 - m_2) c_3 = R(i_2 - \ell_2) + e_2 \qquad \ell_2 \geq 0$$

  - solution has $c_2 \leq c_3 \Rightarrow$ no incentive to run in period 2
- When $m_1 \leq \pi$ and $m_1 + m_2 > \pi$, $s_2 = \pi - m_1$ and $i_2 = 1 - \pi$
  - solution has $c_2 < 1 \rightarrow$ fee imposed in period 2

# Time consistency in period 1

- If $m_1 > \pi$, the fund can forecast $m_2$
  - assumes a run is underway $\Rightarrow m_1 = \pi_1 + \delta(1 - \pi_1)$
    - observing $m_1 > \pi$ allows the bank to infer $\pi_1$
  - no run at $t = 2 \quad \Rightarrow m_2 = (1 - \delta)(\pi - \pi_1)$
- Time consistency at $t = 1$ requires $(c_1, c_2, c_3)$ to solve:

$$\max_{\{c_1, c_2, c_3\}} \quad m_1 u(c_1) + m_2 u(c_2) + (1 - m_1 - m_2) u(c_3)$$

$$m_1 c_1 + m_2 c_2 = s + r\ell$$

$$(1 - m_1 - m_2) c_3 = R(i - \ell) \qquad \ell \geq 0$$

  - solution has $c_1 = c_2 < c_3$ and $c_1 = c_2 < 1 \rightarrow$ fee imposed in period 1
- Note: redemption fee removes the incentive to run **if** the run is detected right away

return

# Quotes from SEC (2023)

- Investors make decisions without knowing $m_1$:

  "Investors generally will not know with certainty if the fund's flows for any particular day will trigger a liquidity fee since a fund's net flows are dependent on many investors' individual investment decisions, which are not knowable in advance and can be influenced by a multitude of different factors."

- Fees are mandatory:

  "The amended framework does not provide discretion to the board with respect to its application. Rather, the fund will be required to apply a fee if it crosses the net redemption threshold unless the fee amount is *de minimis."*

return

# Best run-proof policy

Choose the policy $c_1(m_1)$ for $m_1 \leq \pi$ to solve:

expected utility with no run

$$\max_{\{c_1(m_1) | m_1 \leq \pi\}} \int_0^{\pi} \left\{ \begin{matrix} \pi_1 u\big(c_1(\pi_1)\big) + (\pi - \pi_1) u\big(c_2(\pi_1, \pi_2)\big) \\ +(1-\pi) u\big(c_3(\pi_1, \pi_2)\big) \end{matrix} \right\} f(\pi_1) d\pi_1$$

- subject to the run-proof constraint:

redeem

$$\int_0^{\pi} u\big(c_1(m_1)\big) f_n(\pi_1) d\pi_1 \leq$$

if I expect all others to run

wait

$$\int_0^{\pi} \big[p_n u\big(c_2(m_1, m_2)\big) + (1 - p_n) u\big(c_3(m_1, m_2)\big)\big] f_n(\pi_1) d\pi_1$$

- where $c_2(m_1, m_2)$ and $c_3(m_1, m_2)$ are:

  $(i)$ feasibile

  $(ii)$ chosen optimally for $m_1 + m_2 = \pi$

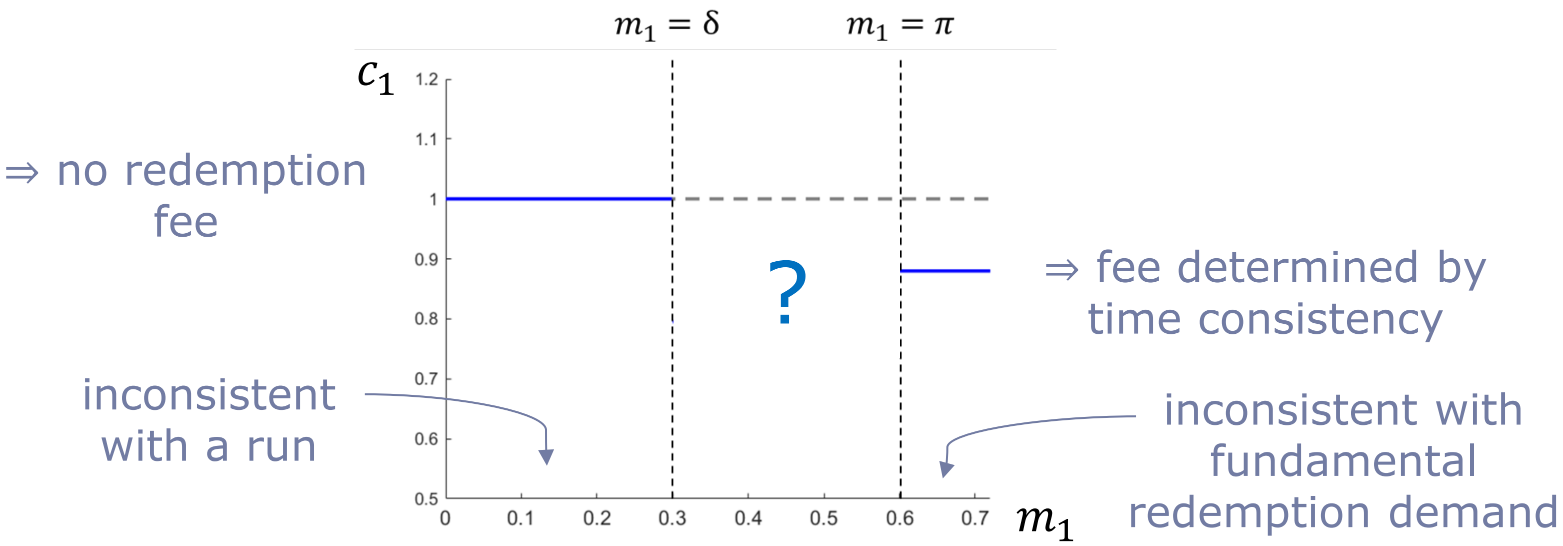
  $(iii)$ time consistent for $m_1 + m_2 > \pi$

same functions in objective and constraint

but evaluated at *different points*

# Three regions

- The best run-proof contract:




- When $m_1 < \delta$, fund is sure there is no run ⇒ no fee
- When $m_1 > \pi$, fund knows a run is underway

  ⇒ sets the time-consistent fee
- In between …

# Middle region

- Optimal fee in the middle region depends on the ratio:

likelihood of $m_1$:

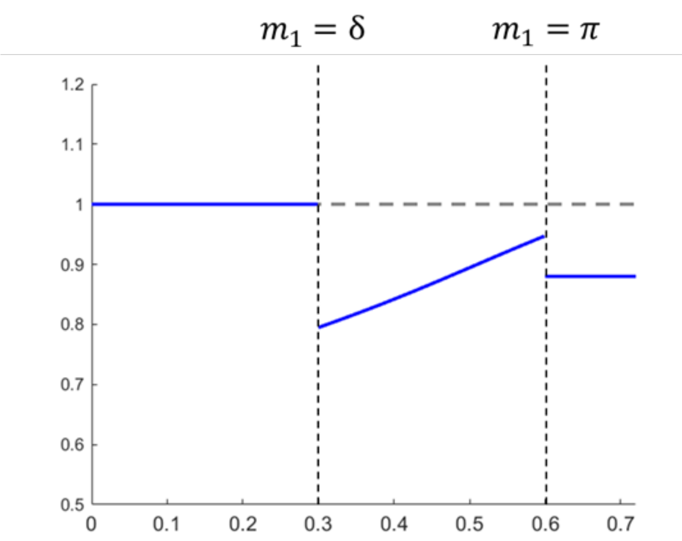
number of investors who pay the fee

conditional on no run

conditional on a run (if I am non-type 1)

$$\frac{f(m_1)\cdot m_1}{f_n\left(\frac{m_1-\delta}{1-\delta}\right)} \approx \frac{\text{decrease in objective}}{\text{benefit in meeting the constraint}}$$

- Overall shape of policy depends on $f$
- But fee tends to *decrease* in this region (counterintuitive)
  - costly to impose fees when many investors (truly) need the money




return

# Portfolio choice

- So far: fund follows planner's (i.e., first-best) portfolio
  - puts $\pi$ into storage and $1 - \pi$ into investment
- Should the fund hold a more liquid portfolio?
  - goal: smaller liquidation costs ⇒ smaller redemption fees

Result: If the $c_1(m_1) \geq \underline{c}$ for all $m_1 \leq \pi$, then holding excess liquidity lowers investors' expected utility.

- Intuition:
  - redemption fees reallocate resources across investors
  - holding more storage (↓ investment) decreases *total* consumption
    - similar in spirit to Ennis & Keister (2006)