Financial Intermediation and Crises

Itay Goldstein

Wharton School, University of Pennsylvania

Introduction

Financial Systems

- Financial systems are crucial for the efficiency of real activity and resource allocation
    - Vast empirical evidence: e.g., Levine (1997), Rajan and Zingales (1998)
- Different roles performed by the financial sector:
    - Transmission of resources from savers/lenders to investors/borrowers
    - Risk sharing possibilities, encouraging more risk taking
    - Information aggregation guiding investment decisions

Not always working perfectly…

Financial Crises

- Financial markets and institutions are often subject to crises:
  - Failure of banks, and/or the sharp decrease in credit and trade, and/or the collapse of an exchange rate regime, etc.
  - Generate extreme disruption of these normal functions of financial and monetary systems, thereby hurting the efficiency of the economy
- Many examples:
  - East-Asian crisis of late 90s
  - Global financial crisis of 2007-2009 and its aftermath

Three Branches of Theories of Financial Crises

Banking Crises and Panics

- Banks provide liquidity transformation allowing people to benefit from the fruits of illiquid long-term investments even if they need early liquidity
- This exposes banks to the risk of bank runs and coordination failures
    - Bryant (1980) and Diamond and Dybvig (1983)
- Policies designed to reduce the risk of bank runs e.g., deposit insurance
- Phenomenon manifested itself in other institutions and markets recently
    - Schmidt, Timmermann, and Wermers (2016), Covitz, Liang, and Suarez (2013)

Credit Frictions and Market Freezes

- Basic frictions like moral hazard and adverse selection affect the financial sector preventing smooth flow of credit and trade
  - Akerlof (1970), Stiglitz and Weiss (1981)
- Link to crises: shocks in the financial system or in the real economy are amplified due to financial frictions, leading to a vicious circle
  - E.g., Holmstrom and Tirole (1997)
- Much literature in macroeconomics studying the effect of frictions on business cycles
  - E.g., Kiyotaki and Moore (1987)

Currency Crises

- Governments try to maintain a fixed exchange rate regime which is inconsistent with other policy goals such as free capital flows and flexible monetary policy
- First generation models: speculators force devaluation
  - Krugman (1979)
- Second generation models: government is making an active choice between exchange rate stability and other policy goals
  - Obstfeld (1996)
- Link to models of sovereign debt crises

Interactions between Different Branches of Models

- Over time, we see that crises are not isolated, but rather the different types of crises interact with each other and amplify each other
- Twin Crises: banking crises and currency crises are strongly related
  - Kaminsky and Reinhart (1999)
  - Mechanisms where banking crises amplify currency crises and vice versa
- Borrowing moral hazard interacts with banking crises and currency crises
- Integration of different theories; mostly following the 1990s crises

Financial Fragility and Coordination Failures

- A primary source for fragility is: c**oordination failures**
- A coordination failure arises when economic agents take a destabilizing action based on the expectation that other agents will do so as well. The result is a **self-fulfilling crisis**
- The key ingredient for this to arise is **strategic complementarities**: agents want to do what others do
- The result is often described as **panic**

Crises: Fundamentals vs. Panic

- Key question in the literature on financial crises is whether they reflect pure fundamentals or they are a result of panic
- Many economists support the panic view:
  - Crises are sudden and unexpected; hard to predict with fundamentals: Friedman and Schwartz (1963) and Kindleberger (1978)
- Large empirical evidence supporting link between fundamentals and crises:
  - For example, Gorton (1988)

- This issue is important not only for understanding the nature of crises but also for policy reasons
  - It is often believed that policy should aim to prevent panic, but not necessarily stop crises that are driven by bad fundamentals
- Global Games Approach connects the two views
  - There is an element of panic in crises, but panic is triggered by fundamentals
  - Carlsson and van Damme (1993), Morris and Shin (1998), Goldstein and Pauzner (2005)
- Empirical evidence: Chen, Goldstein, and Jiang (2010)

Banks, Liquidity Transformation, and Runs

Risk Sharing and Bank Runs: Diamond and Dybvig (1983)

- Diamond and Dybvig provide a seminal model of financial intermediation and **bank runs**.
- Banks Create liquid claims on illiquid assets using **demand-deposit contracts**.
    - Enable investors with early liquidity needs to participate in long-term investments. Provide **risk sharing**.
    - Drawback: Contracts expose banks to panic-based bank runs.

Model (Extended based on Goldstein and Pauzner (2005))

- There are three periods (*0*, *1*, *2*), one good, and a continuum *[0,1]* of agents.
- Each agent is born at period *0* with an endowment of *1*.
- Consumption occurs only at periods *1* or *2*.
- Agents can be of two types:
    - Impatient (probability λ) – enjoys utility $u(c_1)$,
    - Patient (probability $1\text{-}\lambda$) – enjoys utility $u(c_1 + c_2)$.

- Types are i.i.d., privately revealed to agents at the beginning of period *1*.

- Agents are highly risk averse. Their relative risk aversion coefficient:

$$-\frac{cu''(c)}{u'(c)} > 1 \text{ for any } c \geq 1.$$

  - This implies that $cu'(c)$ is decreasing in c for $c \geq 1$, and hence $cu'(c) < u'(1)$ for $c > 1$.

  - Assume $u(0) = 0$.

- Agents have access to the following technology:

- *1* unit of input at period *0* generates *1* unit of output at period *1* or *R* units at period *2* with probability $p(\theta)$.
- θ is distributed uniformly over *[0,1]*. It is revealed at period *2*.
- $p(\theta)$ is increasing in θ.
- The technology yields (on average) higher returns in the long run:

$$E_\theta[p(\theta)]u(R) > u(1).$$

Autarky

- In autarky, impatient agents consume in period 1, while patient agents wait till period 2. The expected utility is then:

$$\lambda u(1) + (1-\lambda)u(R)E_{\theta}[p(\theta)]$$

- Because agents are risk averse, there is a potential gain from transferring consumption from impatient agents to patient agents, and letting impatient agents benefit from the fruits of the long-term technology.

- We now derive the first-best and see how it can be implemented.

First-Best Allocation (if types were verifiable)

- A social planner verifies types and allocates consumptions.
- Period-1 consumption of impatient agents: c_1.
- Period-2 consumption of patient agents is the remaining resources: $c_2 = \frac{(1-\lambda c_1)}{1-\lambda}R$ (with probability $p(\theta)$).
- Planner sets c_1 to maximize expected utility:

$$\lambda u(c_1) + (1-\lambda)u\left(\frac{(1-\lambda c_1)}{1-\lambda}R\right)E_\theta[p(\theta)]$$

- First order condition:

$$u'(c_1^{FB}) = Ru'\left(\frac{(1-\lambda c_1^{FB})}{1-\lambda}R\right)E_\theta[p(\theta)]$$

- Suppose that $\mathrm{c}_1^{\mathrm{FB}} = 1$: $\mathrm{u}'(1) > Ru'(\mathrm{R})\mathrm{E}_\theta[\mathrm{p}(\theta)]$.
- Since the LHS is decreasing and the RHS is increasing in c_1^{FB}, we get that: $c_1^{FB} > 1$.
- The social planner achieves risk sharing by liquidating a larger portion of the long-term technology and giving it to impatient agents. The benefit of risk sharing outweighs the cost of lost output.

The Role of Banks

- The main insight of Diamond and Dybvig is that banks can replicate the first-best allocation with demand-deposit contracts.
    - Hence, they overcome the fact that types are not verifiable.
- Banks offer a short-term payment r_1 to every agent who claims to be impatient.
- By setting $r_1 = c_1^{FB}$, they can achieve the first-best allocation, as long as the incentive compatibility constraint holds:

$$u(c_1^{FB}) \leq u\left(\frac{(1-\lambda c_1^{FB})}{1-\lambda}R\right)E_\theta[p(\theta)]$$

- Yet, things are not so simple, as one has to think carefully about the mechanic details of how banks serve agents and the resulting equilibria.
- Suppose that banks follow a sequential service constraint:
  - They pay r_1 to agents who demand early withdrawal as long as they have resources.
  - If too many agents come and they run out of resources, they go bankrupt, and remaining agents get no payment.
- Impatient agents demand early withdrawal since they have no choice. Patient agent have to consider the following payoff matrix:

Period	$n < 1/r_1$	$n \geq 1/r_1$
1	r_1	$\begin{cases} r_1 & prob \quad \frac{1}{nr_1} \\ 0 & prob\ 1 - \frac{1}{nr_1} \end{cases}$
2	$\begin{cases} \frac{(1-nr_1)}{1-n} R & prob \quad p(\theta) \\ 0 & prob\ 1 - p(\theta) \end{cases}$	0

Here, n is the proportion of agents (patient and impatient who demand early withdrawal.

Multiple Equilibria

- Assuming that the incentive compatibility condition holds, there are at least two Nash equilibria here:
  - **Good equilibrium:** only impatient agents demand early withdrawal.
    - Clear improvement over autarky. First-best is achieved.
  - **Bad equilibrium:** all agents demand early withdrawal. **Bank Run** occurs.
    - Inferior outcome to autarky. No one gets access to long-term technology and resources are allocated unequally.

Source and Nature of Bank Runs

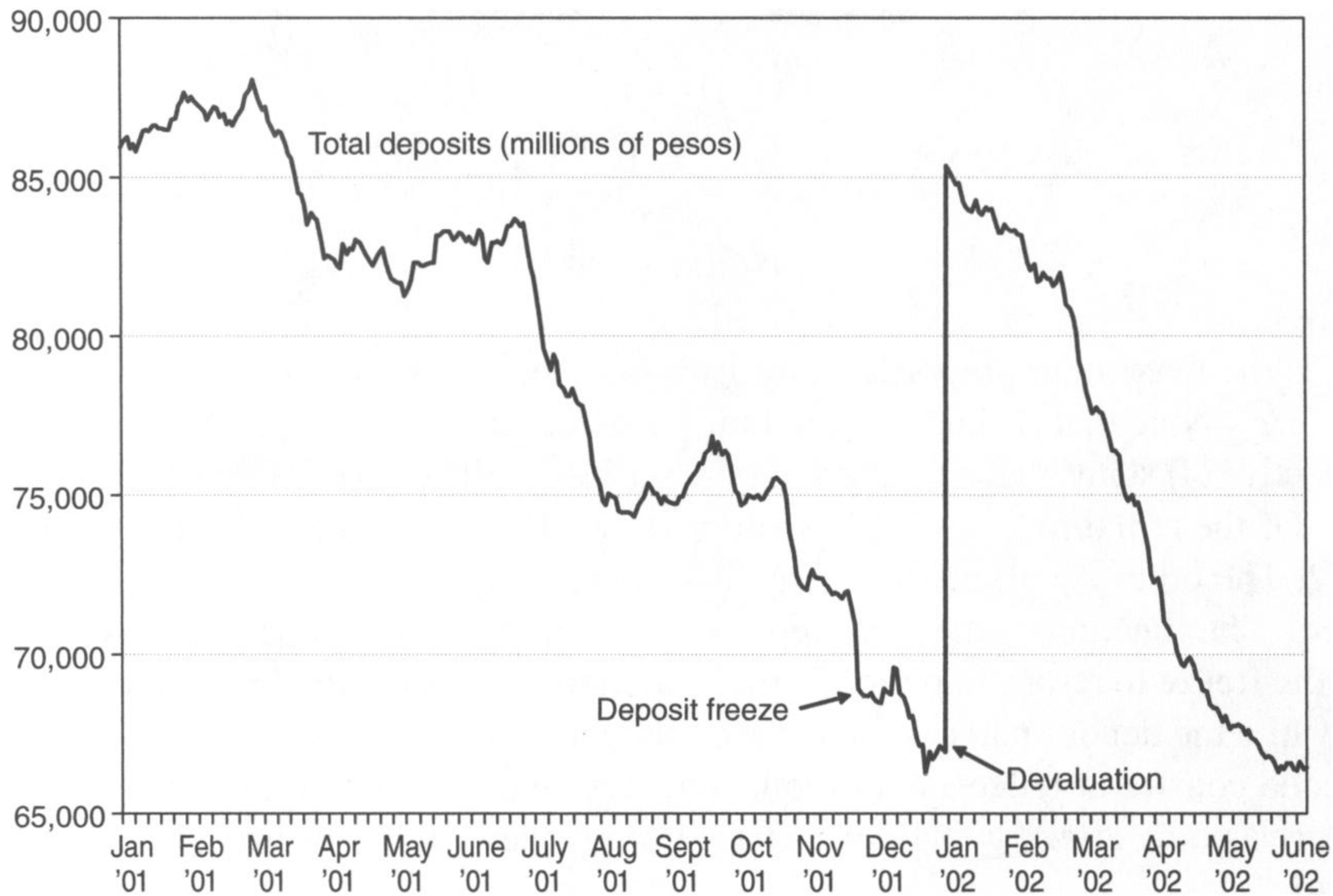
- Bank runs occur because of **strategic complementarities** among agents. They want to do what other agents do.
  - When everyone runs on the bank, this depletes the bank's resources, and makes running the optimal choice.

- As a result, runs are **panic-based**: They occur as a result of the **self-fulfilling beliefs** that other depositors are going to run.

- Moreover, here, they are **unrelated to fundamentals**.
  - Some tend to attribute them to sunspots.

Solutions to Fragility – Suspension of Convertibility:

- Suppose that the bank announces that after λ depositors withdraw in period 1, no one else gets money in this period.
- The good equilibrium becomes the unique equilibrium.
- Patient agents know that no matter what others do, they are guaranteed to get $u\left(\frac{(1-\lambda c_1^{FB})}{1-\lambda}R\right)E_\theta[p(\theta)] > u(c_1^{FB})$.
- Hence, the run is prevented without even triggering suspension.
- **Problem:** What if the number of impatient agents is not known? What about commitment?



Evolution of deposits during the crisis in Argentina; Ennis and Keister (2009)

Solutions to Fragility – Deposit Insurance:

- Suppose that the government provides insurance to the bank in case of excess withdrawals.
  - To maintain the assumption of 'closed' economy, suppose that the government obtains this amount by taxing depositors.
- Again, the good equilibrium becomes the unique equilibrium.
  - Patient agents know that the withdrawal by others is not going to harm their long-term return.
- **Problems:** Deposit insurance might generate moral hazard; deposit insurance can be costly if it is paid from taxes

Problems with Multiplicity

- The model provides no tools to determine when runs will occur.
- This is an obstacle for:
  - **Understanding liquidity provision and runs:**
    - How much liquidity will banks offer when they take into account the possibility of a run and how it is affected by the banking contract?
    - Given that banks may generate a good outcome and a bad outcome, it is not clear if they are even desirable overall.

- **Policy analysis:** which policy tools are desirable to overcome crises?
  - Deposit insurance is perceived as an efficient tool to prevent bank runs, but it might have costs, e.g., moral-hazard.
  - Without knowing how likely bank runs are, it is hard to assess the desirability of deposit insurance.
- **Empirical analysis:** what constitutes sufficient evidence for the relevance (or lack of) of strategic complementarities in fragility?
  - Large body of empirical research associates crises with weak fundamentals. Is this evidence against the panic-based approach?
  - How can we derive empirical implications? See Goldstein (2012).

The Global Games Approach

The Global-Games Approach

- The global-games approach – based on Carlsson and van Damme (1993) – enables us to derive a unique equilibrium in a model with strategic complementarities and thus overcome the problems associated with multiplicity of equilibria (discussed above).
- The approach assumes **lack of common knowledge** obtained by assuming that agents observe slightly noisy signals of the fundamentals of the economy.
- A simple illustration is provided by Morris and Shin (1998).

A Model of Currency Attacks:

Morris and Shin (1998)

- There is a continuum of speculators *[0,1]* and a government.

- The exchange rate without intervention is $f(\theta)$, where $f'(\theta) > 0$, and θ, the fundamental of the economy, is uniformly distributed between *0* and *1*.

- The government maintains the exchange rate at an over-appreciated level (due to reasons outside the model): $e^* > f(\theta), \forall\theta$.

- Speculators may choose to attack the currency.
  - The cost of attack is t (transaction cost).
  - The benefit in case the government abandons is $e^* - f(\theta)$.
    - In this case, speculators make a speculative gain.
- The government's payoff from maintaining is: $v - c(\alpha, \theta)$.
  - v can be thought of as reputation gain.
    - $c(\alpha, \theta)$ is increasing in α (proportion of attackers) and decreasing in θ.

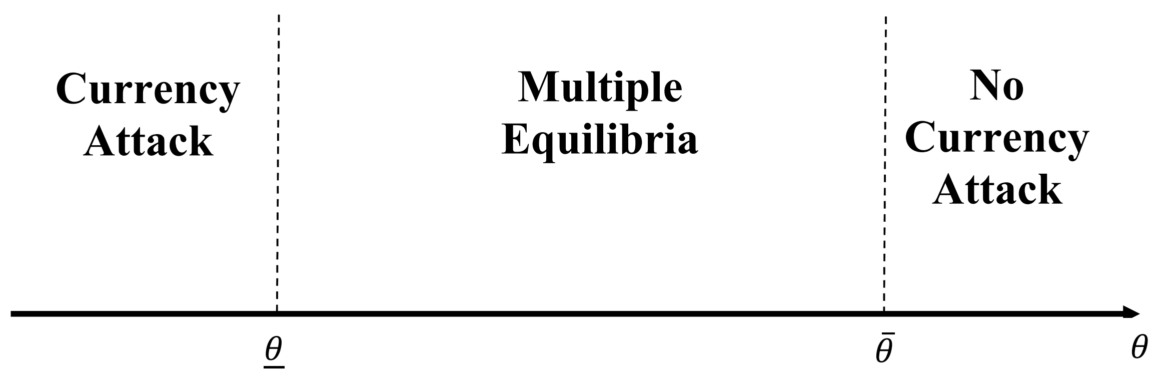
Equilibria under Perfect Information

- Suppose that all speculators (and the government) have perfect information about the fundamental θ.

- Define extreme values of θ, $\underline{\theta}$ and $\overline{\theta}$: $1 > \overline{\theta} > \underline{\theta} > 0$, such that:
  - $c(0, \underline{\theta}) = v$.
  - $e^* - f(\overline{\theta}) = t$.
  - Below $\underline{\theta}$, the government always abandons. Above $\overline{\theta}$, attack never pays off.

- Three ranges of the fundamentals:
  - When $\theta < \underline{\theta}$, unique equilibrium: all speculators attack.
  - When $\theta > \overline{\theta}$, unique equilibrium: no speculator attacks.
  - When $\overline{\theta} > \theta > \underline{\theta}$, multiple equilibria: Either all speculators attack or no speculator attacks (for this, assume $c(1,1) > v$).
- As in Diamond and Dybvig, the problem of multiplicity comes from strategic complementarities: when others attack, the government is more likely to abandon, increasing the incentive to attack.

- Equilibria in the basic model:



Introducing Imperfect Information to Morris and Shin (1998)

- Suppose that speculator i observes $\theta_i = \theta + \varepsilon_i$, where ε_i is uniformly distributed between $-\varepsilon$ and ε. (Government has perfect information.)
- Speculators choose whether to attack or not based on their signals.
- The key aspect is that because they only observe imperfect signals, they must take into account what others will do at other signals.
- This will 'connect' the different fundamentals and determine optimal action at each.

Definitions

- Payoff from attack as function of fundamental and aggregate attack:

$$h(\theta, \alpha(\theta)) = \begin{cases} e^* - f(\theta) - t & if \quad \alpha(\theta) > a(\theta) \\ -t & if \quad \alpha(\theta) \leq a(\theta) \end{cases},$$

where $c(a(\theta), \theta) = v$.

- Payoff as a function of the signal and aggregate attack:

$$V(\theta_i, \alpha(\theta)) = \frac{1}{2\varepsilon} \int_{\theta_i - \varepsilon}^{\theta_i + \varepsilon} h(\theta, \alpha(\theta)) d\theta$$

- Threshold strategy characterized by θ' is a strategy where the speculator attacks at all signals below θ' and does not attack at all signals above θ'.
  - Aggregate attack when speculators follow threshold θ':

$$\alpha(\theta,\theta') = \begin{cases} 0 & if \quad \theta > \theta' + \varepsilon \\ \dfrac{\theta' + \varepsilon - \theta}{2\varepsilon} & if \quad \theta' - \varepsilon \le \theta \le \theta' + \varepsilon \\ 1 & if \quad \theta < \theta' - \varepsilon \end{cases}$$

- We will show that there is a unique threshold equilibrium and no non-threshold equilibria that satisfy the Bayesian-Nash definition.

Existence and Uniqueness of Threshold Equilibrium

- Let us focus on the incentive to attack at the threshold:
  - Function $V\big(\theta', \alpha(\theta, \theta')\big)$ is monotonically decreasing in θ'; positive for low θ' and negative for high θ'.
  - Hence, there is a unique θ^* that satisfies $V\big(\theta^*, \alpha(\theta, \theta^*)\big) = 0$.
  - This is the only candidate for a threshold equilibrium, as in such an equilibrium, at the threshold, speculators ought to be indifferent between attacking and not attacking.

- To show that acting according to threshold θ^* is indeed an equilibrium, we need to show that speculators with lower signals wish to attack and those with higher signals do not wish to attack.
  - This holds because: $V\left(\theta_i, \alpha(\theta, \theta^*)\right) > V\left(\theta^*, \alpha(\theta, \theta^*)\right) = 0$, $\forall \theta_i < \theta^*$, due to the direct effect of fundamentals (lower fundamental, higher profit and higher probability of abandoning) and that of the attack of others (lower fundamental, more people attack and higher probability of abandoning).
  - Similarly, $V\left(\theta_i, \alpha(\theta, \theta^*)\right) < V\left(\theta^*, \alpha(\theta, \theta^*)\right) = 0$, $\forall \theta_i > \theta^*$,

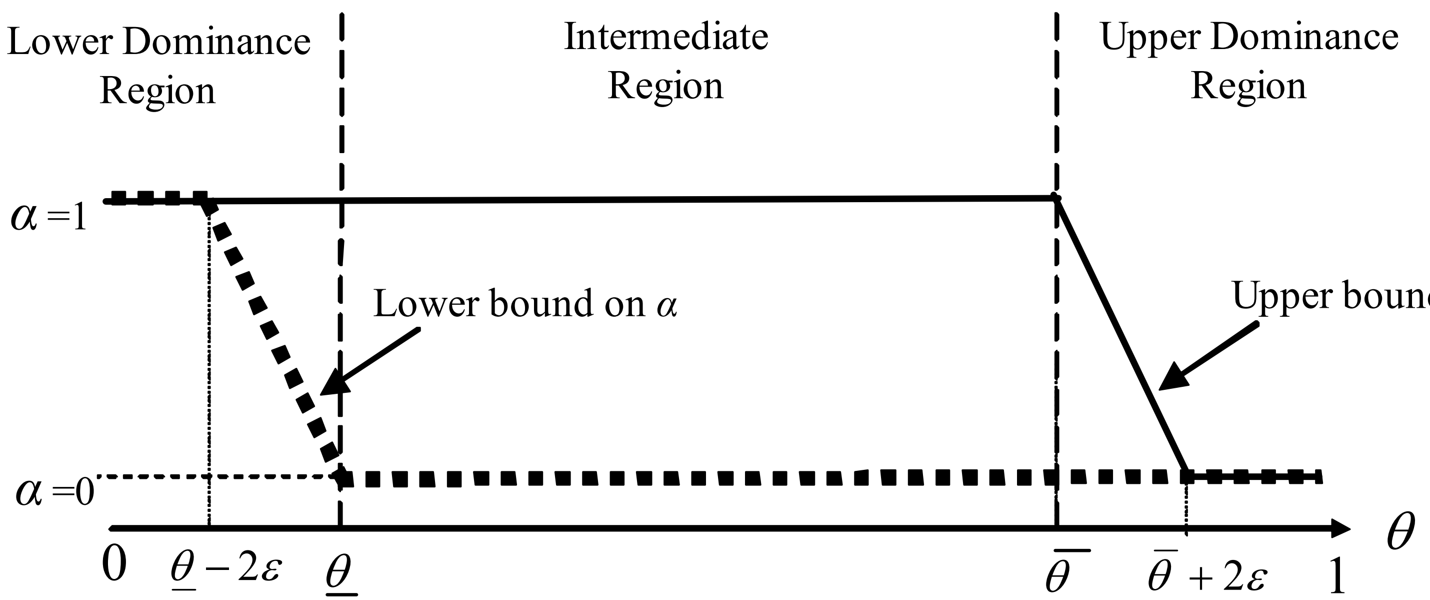
Ruling out Non-Threshold Equilibria

- These are equilibria where agents do not act according to a threshold strategy.
- By contradiction, assume such an equilibrium and suppose that speculators attack at signals above θ^*; denote the highest such signal as θ'^* (we know it is below 1 because of upper dominance region).
- Denote the equilibrium attack as $\alpha'(\theta)$, then due to indifference at a switching point: $V\left(\theta'^*, \alpha'(\theta)\right) = 0$.
- We know that $\alpha'(\theta) \leq \alpha(\theta, \theta'^*)$.

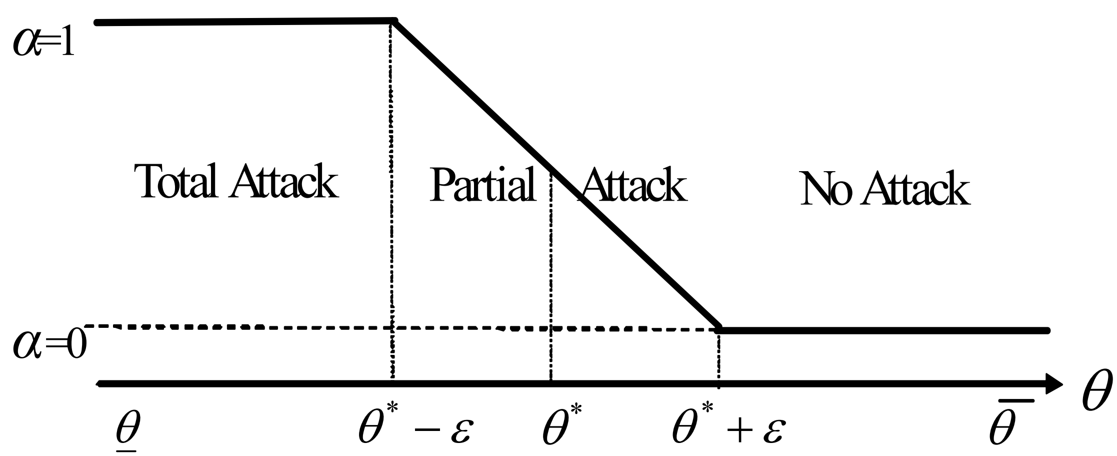
- Then, due to strategic complementarities: $V\left(\theta'^{*}, \alpha(\theta, \theta'^{*})\right) \geq 0$.
- But, this is in contradiction with $V\left(\theta^{*}, \alpha(\theta, \theta^{*})\right) = 0$, since θ'^{*} is above θ^{*} and function $V\left(\theta', \alpha(\theta, \theta')\right)$ is monotonically decreasing in θ'.
- Hence, speculators do not attack at signals above θ^{*}.
- Similarly, one can show that they always attack at signals below θ^{*}.
- This rules out equilibria that are different than a threshold equilibrium, and establishes the threshold equilibrium based on θ^{*} as the unique equilibrium of the game.

Some Intuition

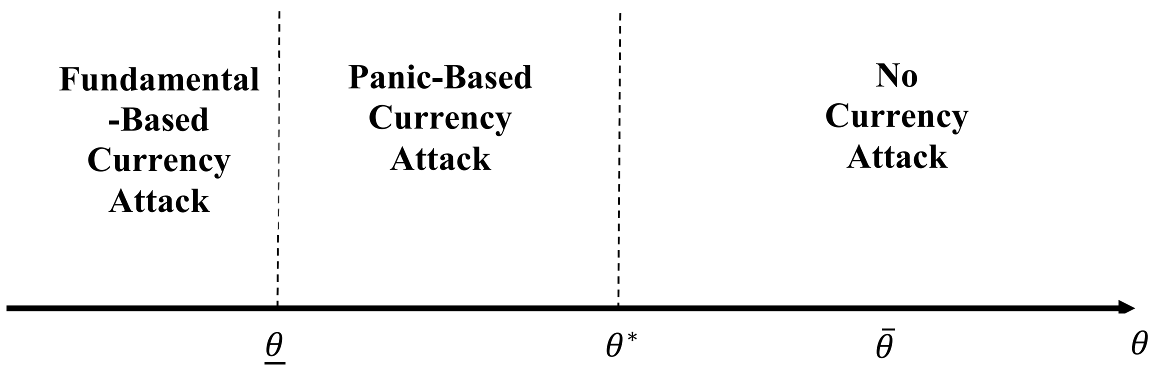
- These are the bounds on the proportion of attack imposed by the dominance regions:



- These bounds can be shifted closer together by iterative elimination of dominated strategies.
- The result is the equilibrium that we found:



- Or, when the noise converges to zero:



Important:

- Although θ uniquely determines α, attacks are still driven by bad expectations, i.e., still panic-based:
  - In the intermediate region speculators attack because they believe others do so.
  - θ acts like a coordination device for agents' beliefs.
- A crucial point: θ is not just a sunspot, but rather a payoff-relevant variable.
  - Agents are obliged to act according to θ.

Why Is This Equilibrium Interesting?

- **First**, reconciles panic-based approach with empirical evidence that fundamentals are linked to crises.
- **Second**, panic-based approach generates empirical implications.
  - Here, the probability of a crisis is pinned down by the value of θ^*, affected by variables t, v, etc. based on: $V\left(\theta^*, \alpha(\theta, \theta^*)\right) = 0$.
- **Third**, once the probability of crises is known, one can use the model for policy implications.
- **Fourth**, captures the notion of strategic risk, which is missing from the perfect-information version.

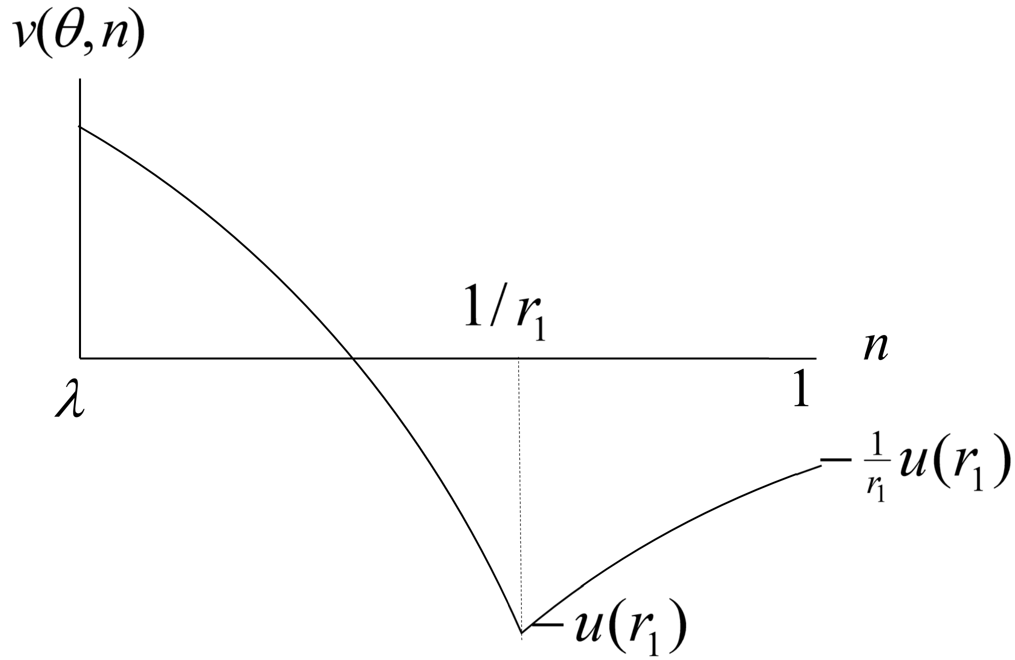
Back to Bank Runs: Goldstein and Pauzner (2005)

- Use global-games approach to address the fundamental issues in the Diamond-Dybvig model.
- But, the Diamond-Dybvig model violates the basic assumptions in the global-games approach. It does not satisfy **global strategic complementarities**.
    - Derive new proof technique that overcomes this problem.
- Once a unique equilibrium is obtained, study how the probability of a bank run is affected by the banking contract, and what is the optimal demand-deposit contract once this is taken into account.

Reminder, Payoff Structure

Period	$n < 1/r_1$	$n \geq 1/r_1$
1	r_1	$\begin{cases} r_1 & prob \quad \frac{1}{nr_1} \\ 0 & prob\ 1 - \frac{1}{nr_1} \end{cases}$
2	$\begin{cases} \frac{(1 - nr_1)}{1 - n} R & prob \quad p(\theta) \\ 0 & prob\ 1 - p(\theta) \end{cases}$	*0*

- Global strategic complementarities do not hold:
  - An agent's incentive to run is highest when $n = 1/r_1$ rather than when $n = 1$.
- Graphically:



- The proof of uniqueness builds on **one-sided strategic complementarities**:
  - v is monotonically decreasing whenever it is positive
- which implies **single crossing**:
  - v crosses zero only once.
- Show uniqueness by:
  - Showing that there exists a unique threshold equilibrium.
  - Showing that every equilibrium must be a threshold equilibrium.

The Demand-Deposit Contract and the Viability of Banks

- We can now characterize the threshold as a function of the rate offered by banks for early withdrawals. At the limit, as ε approaches zero, $\theta^*(r_1)$ is defined by:

$$\int_{n=\lambda}^{1/r_1} u(r_1) + \int_{n=1/r_1}^{1} \frac{1}{nr_1} u(r_1) = \int_{n=\lambda}^{1/r_1} p(\theta^*) u\left(\frac{(1-nr_1)}{1-n} R\right)$$

  - At the threshold, a patient agent is indifferent.
  - His belief at this point is that the proportion of other patient agents who run is uniformly distributed. Effectively, there is no fundamental uncertainty (only strategic uncertainty).

- Analyzing the threshold $\theta^*(r_1)$ with the implicit function theorem, we can see that it is increasing in r_1.
    - The bank becomes more vulnerable to bank runs when it offers more risk sharing.

- Intuition:
    - With a higher r_1 the incentive of agents to withdraw early is higher.
    - Moreover, other agents are more likely to withdraw at period 1, so the agent assesses a higher probability for a bank run.

Finding the optimal r_1

- The bank chooses r_1 to maximize the expected utility of agents:

$$\lim_{\varepsilon \to 0} EU(r_1) = \int_0^{\theta^*(r_1)} \frac{1}{r_1} u(r_1) d\theta$$

$$+ \int_{\theta^*(r_1)}^1 \lambda u(r_1) + (1-\lambda) p(\theta) u\left(\frac{(1-\lambda r_1)}{1-\lambda} R\right) d\theta$$

- Now, the bank has to consider the effect that an increase in r_1 has on risk sharing and on the expected costs of bank runs.

- Main question: Are demand deposit contracts still desirable?

- Result: If $\underline{\theta}(1)$ is not too large, the optimal r_1 must be larger than 1.
- Increasing r_1 slightly above 1 generates one benefit and two costs:
  - **Benefit:** Risk sharing among agents.
    - Benefit is of first-order significance: Gains from risk sharing are maximal at r_1=1.
  - **Cost I:** Increase in the probability of bank runs beyond $\underline{\theta}(1)$.
    - Cost is of second order: Liquidation at $\underline{\theta}(1)$ is almost harmless.

- **Cost II:** Increase in the welfare loss resulting from bank runs below $\underline{\theta}(1)$.
  - Cost is small when $\underline{\theta}(1)$ is not too large.

- Hence, the optimal r_1 generates panic-based bank runs.

- But, the optimal r_1 is lower than c_1^{FB}.
  - Hence, the demand-deposit contract leaves some unexploited benefits of risk sharing in order to reduce fragility.
  - To see this, let us inspect the first order condition for r_1:

$$\lambda \int_{\theta^*(r_1)}^{1} u'(r_1) - p(\theta) R u'\left(\frac{(1-\lambda r_1)}{1-\lambda} R\right) d\theta =$$

$$\frac{\partial \theta^*(r_1)}{\partial r_1}\left(\lambda u(r_1) + (1-\lambda) p\big(\theta^*(r_1)\big) u\left(\frac{(1-\lambda r_1)}{1-\lambda} R\right) - \frac{1}{r_1} u(r_1)\right)$$

$$+ \int_{0}^{\theta^*(r_1)} \frac{u(r_1) - r_1 u'(r_1)}{r_1^2} d\theta$$

- LHS: marginal benefit from risk sharing. RHS: marginal cost of bank runs.
- Since marginal cost of bank runs is positive, and since marginal benefit is decreasing in r_1: The optimal r_1 is lower than c_1^{FB}.

Summarizing the Takeaways

- Likelihood of runs increases in degree of risk sharing
- Banks adjust the demand deposit contract when they take into account its effect on the probability of a run
  - Risk sharing decreases in equilibrium
- In most cases, banks still improve welfare relative to autarky, as some degree of risk sharing is desirable despite the fragility
- Two inefficiencies occur in equilibrium:
  - Level of risk sharing is below optimal
  - Damaging runs still occur

Caveats Concerning Debt Contracts

- Diamond and Dybvig show that demand deposit contracts can generate the first-best risk sharing with the cost of exposing the system to runs
- Jacklin (1987) shows that the benefits of risk sharing can be achieved in a market mechanism without runs
- An important question is why we still see debt contracts or demand deposit contracts that generate fragility
- Several answers have been proposed in the literature, but this is still an active ongoing debate

- Diamond (1997) suggests that some agents are not sophisticated enough to trade in the market and are thus limited to the traditional banking contracts
- Calomiris and Kahn (1991) and Diamond and Rajan (2001) study models where demand deposit contracts play a disciplinary role aligning the incentives of bank managers with the interests of outside claim holders
- Gorton and Pennacchi (1990) show that debt contracts, which are not sensitive to information, protect agents, who have inferior ability to produce information about bank fundamentals

- More recently, this line of argument has been extended to say that a role of banks is to produce safe assets for investors, who demand them for reasons outside the model (Stein (2012)
    - An extreme version of agents liking information-insensitive contracts
- On the other hand, a strong argument has been developed that banks' debt and fragility are inefficient and stem from a moral hazard problem due to implicit and explicit government guarantees (Admati and Hellwig (2013)
    - The policy conclusion out of this is that banks should be required to hold more capital

Extensions: The Effect of a Large Investor:

Corsetti, Dasgupta, Morris, and Shin (2004)

- So far we analyzed situations with many small investors.
- A very relevant question is how things are going to be affected if large investors are present.
- Corsetti, Dasgupta, Morris, and Shin analyze this question motivated by the case of Soros.
  - He is known to have a crucial effect on the attack on the Pound.

- The key intuition can be understood by looking at what happens when instead of a continuum of small investors, there is only one large investor that decides whether to attack/run.

- In the Morris and Shin (1998) model, a large investor would choose to attack if and only if $\theta < \overline{\theta}$.

  - He can force the government to abandon the regime and gain $e^* - f(\theta) - t$, which is positive when $\theta < \overline{\theta}$.

- In the Goldstein and Pauzner (2005) model, a large investor would choose to run if and only if $\theta < \underline{\theta}(1)$.

- He knows that the bank can only pay him 1 in case he demands early withdrawal, which is optimal only when $\theta < \underline{\theta}(1)$.

$\Rightarrow$ In a currency attack model, large investor generates more fragility, while in a bank run model, he generates more stability.

- The unifying theme is that the large investor is able to achieve the best outcome from his point of view.
    - In currency attacks, this means **attack**, whereas in bank runs, this means **no run**.

- What happens when the large investor is present alongside the small investors?
    - The qualitative effect is similar, albeit weaker.
    - Interestingly, the presence of a large investor, affects the behavior of small investors in the same direction.
- Knowing that he is there, they tend to attack more or run less, depending on the context.
- Overall, adding a large investor to the model increases (decreases) the probability of a currency attack (bank run).

Caveats Concerning Global Games Analysis

- **Settings where uniqueness does not hold:**
  - The analysis above did not highlight the role of public information (we will see more below).
  - Overall, uniqueness requires that private signals are sufficiently precise relative to public ones.
  - Angeletos and Werning (2006) analyze how the relative precisions are determined endogenously in the context of trading in a financial market, and the consequences for uniqueness of equilibrium.
  - There are other settings where uniqueness might fail: Angeletos, Hellwig, and Pavan (2006) study the signaling role of the policymaker's

policy and the effect that this has on the informational environment and on the uniqueness of equilibrium.

- But, more recently, Angeletos and Pavan (2013) show that even with multiplicity of equilibria, the general policy analysis and comparative statics analysis go through across equilibria, generating conclusions that could not be obtained in the common-knowledge benchmark.

- **Sensitivity of unique equilibrium to information structure:**
  - In Morris, Frankel, and Pauzner (2003), equilibrium threshold depends on the specification of noise.
  - But, policy analysis and comparative statics analysis will mostly go through.

- **Payoff Structure:**
  - Typical global-games structure is very stylized, forcing global strategic complementarities on the model.
  - Most settings derived from first principles will not have this structure.
  - Bank run model is an example.
    - Micro-founding payoff structure in a bank run game does not yield standard global-games structure.
    - Analysis in Goldstein and Pauzner (2005) deals with this problem.
  - Applications: Dasgupta (2004), Liu (2016), Bouvard and Lee (2016), Daniels, Jager, and Klaassen (2011) (who study a micro-founded model of currency attacks).

Using Global Games for Policy Analysis

Government Guarantees: Allen, Carletti, Goldstein, Leonello (2018)

- In Diamond-Dybvig, deposit insurance eliminates runs and restores full efficiency.
  - It solves depositors' coordination failure without entailing any disbursement for the government.
- However, reality is more complex:
  - Runs also occur because of a deterioration of banks fundamentals and may do so even with deposit insurance.

  - o Design of the guarantee is crucial: should depositors be protected only against illiquidity due to coordination failures or also against bank insolvency?
  - o Guarantees may alleviate crises inefficiencies, but might distort banks' risk taking decisions.
  - o What is the optimal amount of guarantees taking all this into account?
- Notoriously rich and hard to solve model:
  - o Endogenize the probability of a run on banks to see how it is affected by banks' risk choices and government guarantees.

  - Endogenize banks' risk choices to see how they are affected by government guarantees, taking into account investors expected run behavior

- We build on Goldstein and Pauzner (2005), where
  - Depositors' withdrawal decisions are uniquely determined using the global-game methodology.
  - The run probability depends on the banking contract (i.e., amount promised to early withdrawers), and the bank decides on it taking into account its effect on the probability of a run.

- We add a government to this model to study how the government's guarantees policy interacts with the banking contract - our measure of risk- and the probability of a run.
- Some results:
  - Guarantees can increase the probability of crises (via effect on banks' decisions), but still increase welfare.
  - Programs that protect against fundamentals failures may be better than programs protecting only against panics.
  - Distortions in risk taking can go the opposite way of what is typically expected.

Detecting Strategic Complementarities in the Data

Crises: Fundamentals vs. Panic

- For a long time, the theoretical literature provided models of crises that are based either on panic (e.g., Diamond and Dybvig (1983)) or on fundamentals (e.g., Chari and Jagannathan (1988))
- Real-world descriptions of crises often involved a sense of panic:
  - Unexpected events that are not fully explained by fundamentals (Friedman and Schwartz (1963) and Kindleberger (1978))
- Key question is how to test the different mechanisms in the data

- A common approach in the empirical literature was to test whether runs are correlated with fundamentals
  - The idea was that the distinction between the two types of bank runs is that fundamental-based bank runs are correlated with the fundamentals, whereas panic-based bank runs are not
- Following this approach, most empirical studies found a strong link between runs and various types of fundamentals
- Hence, they concluded that they do not find support for the panic-based approach. A brief summary follows

- **Gorton (1988):**
  - Studies the national banking era in the US between 1863 and 1914.
  - Shows that crises were responses of depositors to an increase in perceived risk. Crises occurred whenever key variables that are linked to the probability of recession reached a critical value.
  - The most important variable is the liabilities of failed firms. He also shows an effect of other variables, such as the production of pig iron, which is used as a proxy for consumption.

- **Kaminsky and Reinhart (1999)**
  - Study episodes of banking and currency crises in developing and developed countries between 1970 and 1995.
  - Find that banking crises and currency crises are interrelated and aggravate each other.
    - The twin-crises phenomenon
  - Both are driven by deteriorating fundamentals, as captured by variables like output, terms of trade, and stock prices.

- **Schumacher (2000)**
  - Studies runs on Argentine banks after the 1994 Mexican crisis.
  - Finds that failing banks suffered more withdrawals than surviving banks.
  - These banks were ex-ante 'bad', as measured by variables like capital adequacy, asset quality, liquidity, performance, and size.

- **Martinez-Peria and Schmukler (2001)**
  - Study the behavior of bank deposits and interest rates in Argentina, Chile, and Mexico in the 90's.

- Find that depositors discipline banks, in that they withdraw deposit and/or demand high interest rate when fundamentals deteriorate, as captured by variables like capital adequacy, non-performing loans, and profitability.

- **Calomiris and Mason (2003)**
  - Study bank failures in the US between 1929 and 1931.
  - Show that the duration of survival can be explained by size, asset quality, leverage, and other fundamentals

Using the Global Games Approach: Chen, Goldstein, and Jiang (2010)

- As demonstrated by the theoretical framework, the link between crises and fundamentals does not say much about whether or not coordination failures and strategic complementarities play a role.
  - Even when coordination failures are involved, crises are more likely to occur at low fundamentals.
  - A decrease in fundamentals can trigger the panic.
- Using mutual-fund data, we present an empirical test that relies on **cross-sectional differences in level of complementarities.**

Basic economic force behind bank runs

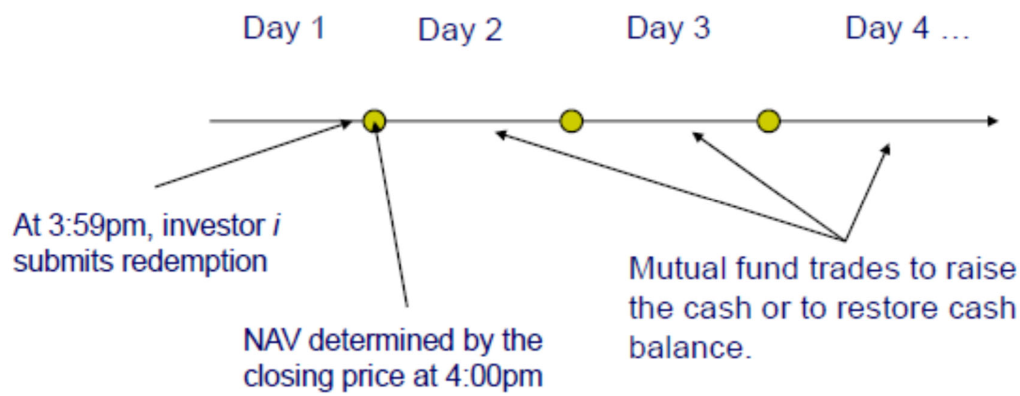
- Strategic complementarities
  - Banks create liquidity by holding illiquid assets and liquid liabilities
  - Depositors are promised a fixed amount if they want to withdraw
  - If many withdraw, the bank will have to liquidate assets at a loss, hurting those who don't withdraw
  - Run arises as a self-fulfilling belief: People run because they think others will do so

What about Non-Bank Institutions?

- Strategic complementarities and run-type behavior are not limited to banks
- Recent example provided by money-market funds: Schmidt, Timmermann, and Wermers (2016)
- One feature that is common to money-market funds and banks is that they have fixed claims, which clearly enhances the first-mover advantage contributing to run dynamics
- New thinking following the crisis involves moving away from the fixed-NAV model to a floating-NAV model as in other mutual funds

Run Dynamics in a Floating-NAV Model

- However, moving to a floating-NAV model does not eliminate the first-mover advantage and the potential for run-like behavior
- In a floating-NAV environment, investors can redeem shares and get the NAV as of the day of redemption
- But, their redemptions will affect fund trading going forward hurting remaining investors in illiquid funds
- This is the source of the first-mover advantage (or strategic complementarities)



- Key feature for empirical analysis:
  - Level of strategic complementarities determined by the illiquidity of the funds' assets
  - Different funds have different levels of illiquidity and thus of strategic complementarities: easy to measure!

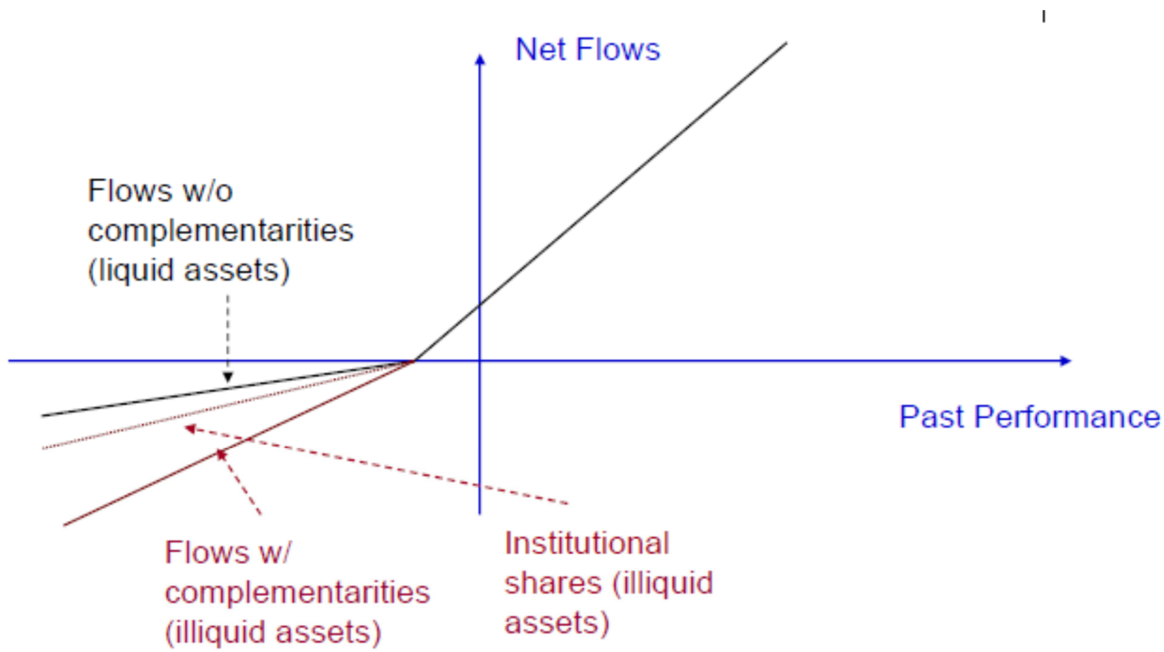
Basic Framework:

- Returns R_1 and R_2. $NAV(t=1) = R_1$.
- Proportion of redeemers: $0 \leq N < 1$.
- Liquidity: need to sell $\$(1+\lambda)$ in order to raise $1.
  - λ measures illiquidity
- Payoff at $t = 2$: $R_1 R_2 [1-(1+\lambda)N]/(1-N)$.
- With inflows $I(R_1)$:

$$\frac{R_1 R_2 [1-(1+\lambda)max\{0,(N-I(R_1))\}]}{1-max\{0,(N-I(R_1))\}}$$

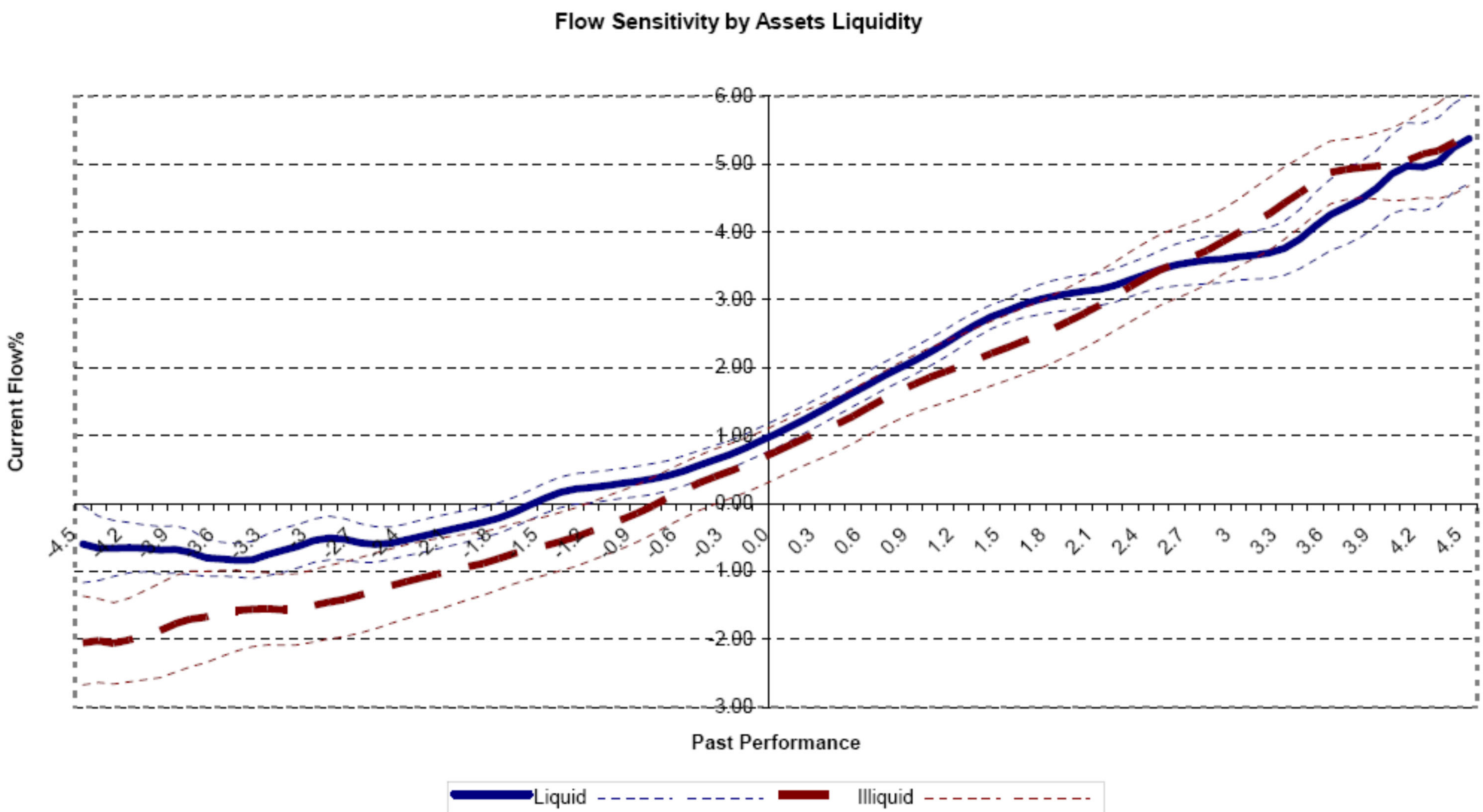
- Two Premises:
  - Complementarities arise when funds experience outflows.
  - Complementarities are stronger when funds hold more illiquid assets.
- Based on a global-game model:
  - **H1:** Conditional on low performance, funds that hold illiquid assets will experience more outflows.
- Sharpen the test (based on Corsetti et al., 2004):
  - **H2:** Pattern weakens when fund is held by large investors.

Summary of Predictions:



Empirical Analysis of Flows in Equity Mutual Funds

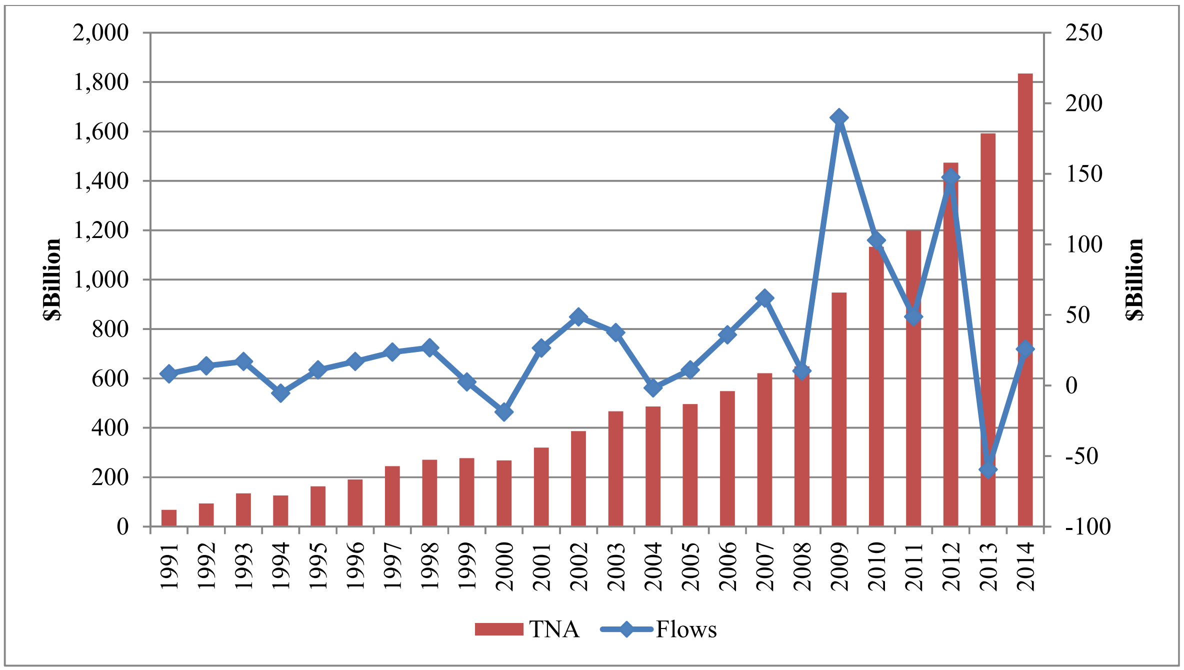
- Chen, Goldstein and Jiang (2010) study flows in 4,393 actively-managed equity funds from 1995-2005
- Find stronger sensitivity of outflows to negative performance in illiquid funds
  - Illiquid funds are: small-cap & mid-cap equity funds (domestic or international), or single-country funds excluding US, UK, Japan and Canada.
    - Or continuous measure of liquidity of portfolio
- Pattern is weaker in funds that are mostly held by institutional investors


Flow Sensitivity by Assets Liquidity
Current Flow%
Past Performance
6.00
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.5
-4.2
-3.9
-3.6
-3.3
-3
-2.7
-2.4
-2.1
-1.8
-1.5
-1.2
-0.9
-0.6
-0.3
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3
3.3
3.6
3.9
4.2
4.5
Liquid
Illiquid

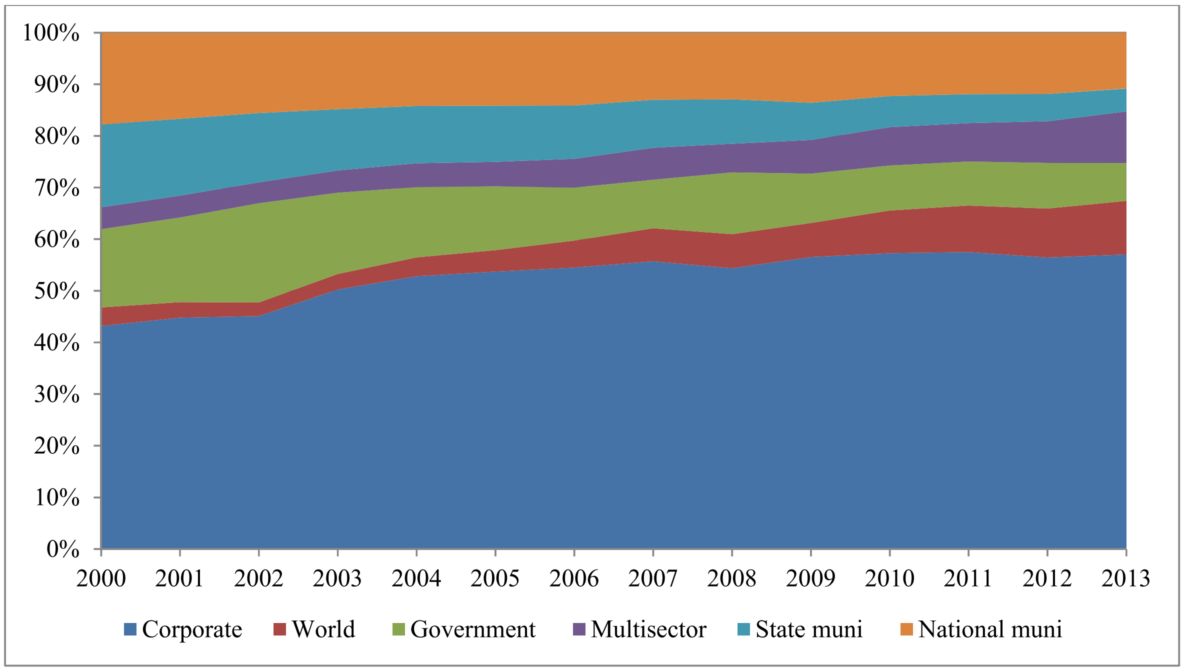
Corporate Bond Funds: Goldstein, Jiang, and Ng (2017)

- Following the crisis, massive inflows into corporate bond funds come largely as a response to changes in investment opportunities and regulation elsewhere in the financial system
- Concerns mentioned about potential fragility mounting in the corporate bond funds sector
    - For example, Feroli, Kashyap, Schoenholtz, and Shin (2014) raise concerns for fragility and outflows in case of tightening of monetary policy
- Do we see signs of fragility in the data?

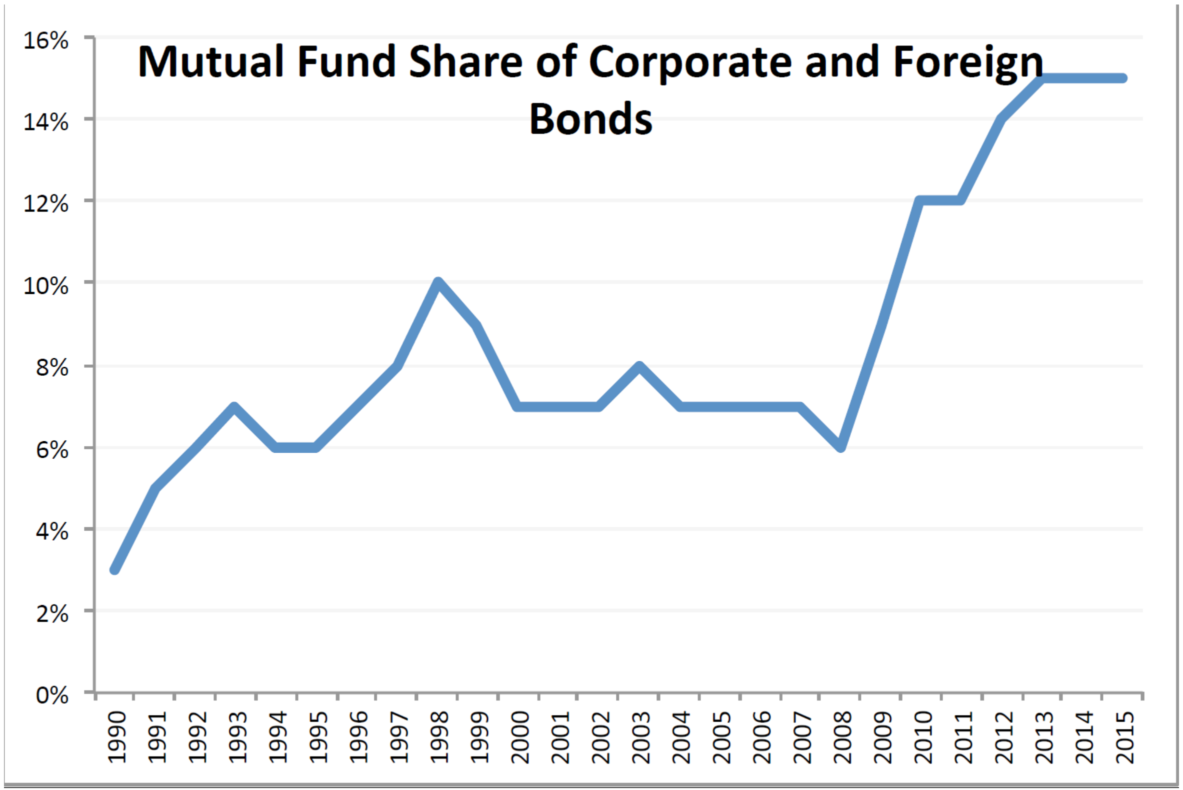
Total Net Assets and Flows of Active Corporate Bond Funds



Distribution of Bond Fund Assets



Mutual-Fund Share of the Market



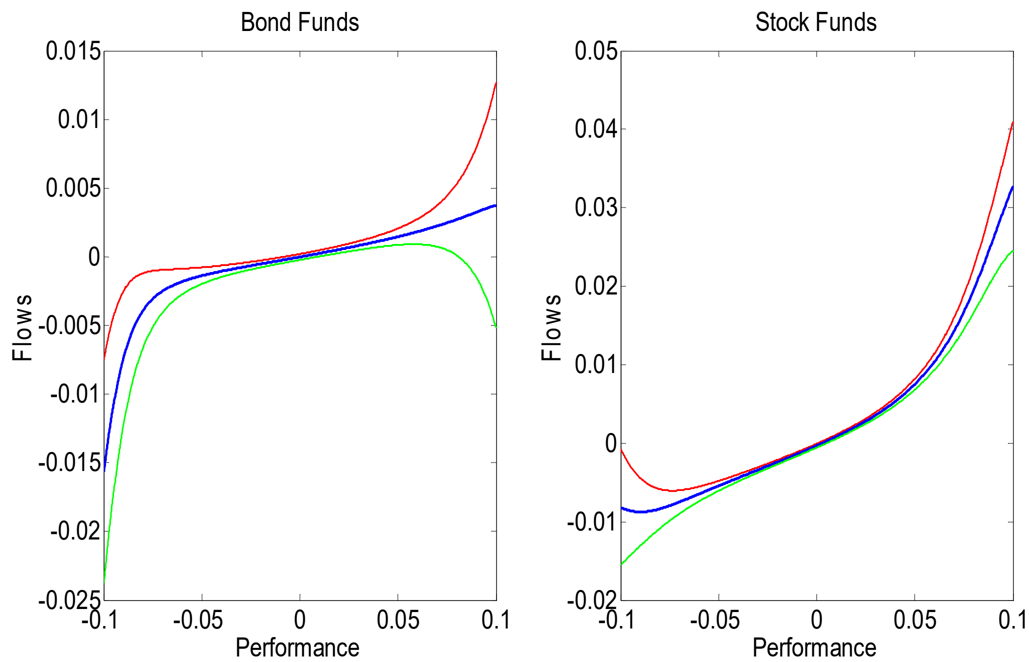
Empirical Analysis of Flows in Corporate Bond Mutual Funds

- Goldstein, Jiang and Ng (2017) study flows in 1,660 actively-managed corporate bond funds from 1992-2014
    - Compare the pattern with that of equity funds
    - Link pattern to illiquidity
    - Motivation based on the fragility argument in Chen, Goldstein, and Jiang (2010)
        - Mutual funds create strategic complementarities for investors especially when the assets are illiquid

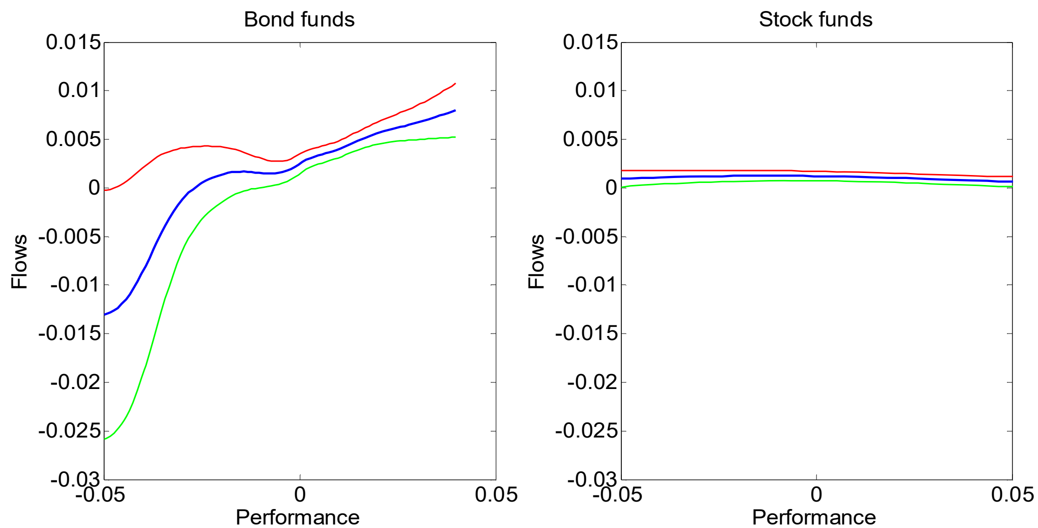
- Illiquidity is much more severe for corporate bonds:
  - Corporate bonds trade infrequently
    - Corporate bonds trade OTC; individual bond issues do not trade on 48% of days in their sample
    - Corporate bonds account for only ~ 2.5 to 3.7% of trading volume in U.S. bonds.
    - U.S. Treasuries account for 69% of volume.
  - More difficult to get up-to-date price for corporate bonds
  - Price impact and other illiquidity costs are high

- Large literature on the flow-to-performance relation in equity funds, finding convex relation (greater sensitivity on upside than on downside)
- We find that corporate bond funds are different:
  - Flow-to-performance relation tends to be concave (greater sensitivity on downside than on upside)
  - Pattern strengthens with illiquidity
    - Funds that hold less cash or periods with greater aggregate illiquidity
- Evidence is consistent with fragility, even in the aggregate

Flow Performance Relation of Corporate Bond Funds vs. Equity Funds



Does redemption sensitivity disappear in aggregation?



Flow-Performance in Corporate Bond vs. Stock Funds

	(1)	(2)
	Corporate Bond Funds	Stock Funds
Alpha	0.238***	0.994***
	(2.71)	(34.23)
Alpha× (Alpha<0)	0.621***	-0.575***
	(4.34)	(-14.70)
Alpha<0	-0.00979***	-0.00723***
	(-18.45)	(-25.06)
Lagged Flow	0.152***	0.118***
	(21.47)	(29.90)
Log(TNA)	0.000728***	0.000459***
	(5.74)	(5.46)
Log(Age)	-0.0157***	-0.0183***
	(-32.08)	(-70.95)
Expense	-0.200***	-0.0522
	(-2.59)	(-0.77)
Rear Load	-0.00280***	-0.134***
	(-3.68)	(-5.51)
Observations	307,242	1,578,506
Adj. R2	0.0646	0.0583

Flow-Performance in Underperforming Funds in Illiquid Times

	(1) VIX	(2) TED	(3) DFL	(4) MOVE
Alpha	-0.131	-0.121	-0.746***	-0.0909
	(-0.77)	(-1.11)	(-3.22)	(-0.73)
Alpha*IlliqPeriod	0.753***	0.749***	1.412***	0.639***
	(3.89)	(5.37)	(5.21)	(4.58)
IlliqPeriod	0.00690***	0.00148**	0.00745***	0.00252***
	(9.81)	(2.44)	(8.11)	(4.19)
Lagged Flow	0.121***	0.123***	0.152***	0.123***
	(15.37)	(15.47)	(14.90)	(15.50)
Log(TNA)	0.000552***	0.000558***	0.000533***	0.000544***
	(3.78)	(3.82)	(2.98)	(3.75)
Log(Age)	-0.0134***	-0.0136***	-0.0124***	-0.0135***
	(-26.78)	(-26.70)	(-17.88)	(-26.70)
Expense	-0.175**	-0.185**	-0.284**	-0.183**
	(-1.98)	(-2.10)	(-2.45)	(-2.08)
Rear Load	-0.00294***	-0.00285***	-0.00611***	-0.00291***
	(-3.40)	(-3.29)	(-5.87)	(-3.36)
Observations	171,006	171,006	100,215	171,006
Adj. R^2	0.0339	0.0330	0.0429	0.0329

Asset Liquidity and Flow-Performance Relation

Alpha<0	Low Cash	Low (Cash + Government Bonds)	Low NSAR Cash	Illiquid Corporate Bond Holdings 1	Illiquid Corporate Bond Holdings 2
Alpha	0.554***	0.567***	0.631***	0.688***	0.662***
	(6.42)	(6.17)	(6.09)	(3.20)	(3.16)
Alpha×IlliqFund	0.814***	0.647***	0.767***	1.305***	1.174***
	(3.21)	(2.74)	(3.82)	(3.02)	(2.82)
IlliqFund	-0.000288	0.00113	0.00211*	0.00472***	0.00435***
	(-0.38)	(1.51)	(1.73)	(2.89)	(2.74)
Lagged Flow	0.131***	0.132***	0.121***	0.180***	0.179***
	(12.50)	(12.52)	(7.15)	(10.67)	(11.11)
Log(TNA)	0.000561***	0.000555***	0.000470*	0.000831***	0.000928***
	(3.18)	(3.15)	(1.80)	(2.58)	(2.86)
Log(Age)	-0.0140***	-0.0140***	-0.0142***	-0.0153***	-0.0157***
	(-20.26)	(-20.22)	(-14.61)	(-12.59)	(-12.95)
Expense	-0.443***	-0.449***	-0.521***	-0.0281	-0.0158
	(-3.99)	(-4.02)	(-3.10)	(-0.14)	(-0.08)
Rear Load	-0.00485***	-0.00482***	-0.00221	-0.00474**	-0.00482**
	(-4.78)	(-4.74)	(-1.45)	(-2.49)	(-2.50)
Observations	108,745	108,745	49,759	25,389	25,370
Adj. R^2	0.0500	0.0498	0.0473	0.0732	0.0750

The Effect of Institutional Investors

Alpha<0	Institutional-Oriented Funds		Retail-Oriented Funds	
	(1)	(2)	(3)	(4)
Alpha	2.056***	2.042***	0.958**	1.024**
	(3.61)	(3.58)	(2.28)	(2.44)
Alpha×LowCash	-0.906	-0.898	1.268***	1.236***
	(-1.17)	(-1.17)	(2.90)	(2.83)
Low Cash	-0.00304	-0.00301	-0.000683	-0.00100
	(-1.57)	(-1.56)	(-0.52)	(-0.76)
Lagged Flow	0.108***	0.108***	0.106***	0.105***
	(5.35)	(5.34)	(4.86)	(4.82)
Log(TNA)	0.000324	0.000391	0.000489	0.000822**
	(0.89)	(1.03)	(1.40)	(2.22)
Log(Age)	-0.0163***	-0.0164***	-0.0132***	-0.0124***
	(-9.70)	(-9.58)	(-10.57)	(-9.96)
Expense	0.0469	-0.0168	-0.543***	-0.331*
	(0.14)	(-0.05)	(-3.02)	(-1.81)
Rear Load	-0.00340	-0.00352	-0.00639***	-0.0062***
	(-1.21)	(-1.25)	(-4.11)	(-3.99)
Inst		-0.00123		0.00766***
		(-0.48)		(4.03)
Observations	19,331	19,331	37,367	37,367
Adj. R^2	0.0398	0.0398	0.0490	0.0500

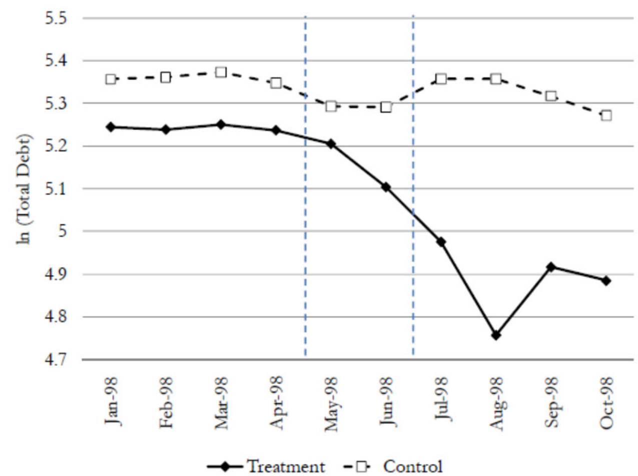
Evidence from a Natural Experiment:
Hertzberg, Liberti, and Paravisini (2011)

- Hertzberg, Liberti, and Paravisini (2010) use a natural experiment, based on the expansion of the Public Credit Registry in Argentina in 1998, to test for strategic complementarities in lending.
    - Prior to 1998, the registry only provided information about borrowers, whose total debt was above $200,000.
    - In 1998, the need for the threshold was eliminated, leading to the disclosure of information about 540,000 borrowers, for which credit assessments were previously only known privately.

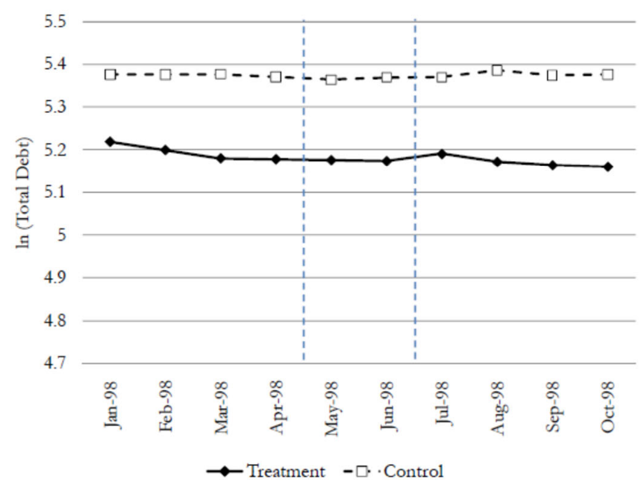
- They identify the presence of complementarities in lending by studying the difference in lenders' behavior following the announcement of the expansion.
- Consider a lender who had negative information about a borrower, for whom the information was not initially disclosed.
- From the point of view of this lender, no new information has arrived between the two periods.
- The only difference is that in the intermediate period, he realizes that the information will become available publicly.
- The authors show that for these borrowers, the amount of credit has decreased following the announcement of expansion.

- This is supposedly because the lenders realized that making this information public will make other lenders reduce credit.
- Moreover, using a differences-in-differences approach, they show that the decrease in debt following the announcement is not observed for:
  - Firms that were slightly above the threshold (for whom the information was always available).
  - Firms who borrow from only one lender (for whom there is no coordination problem).
- Overall, above approaches can be used to assess the relevance of strategic complementarities in other settings and guide policy.

Panel A. Firms with Multiple Lenders before Expansion Announcement



Panel B. Firms with a Single Lender before Expansion Announcement



Credit Market Frictions

Credit Frictions

- Much of the literature on credit and trading frictions focused on problems originating from moral hazard or adverse selection, going back to the seminal paper by Stiglitz and Weiss (1981)
- **Moral hazard:** If borrowers can take an action that affect the quality of the loan, then they need to have enough capital at stake for incentives
- **Adverse selection:** If borrowers know more about the quality of the loan, then markets may break down
- Holmstrom and Tirole (1997) provides a canonical representation of the moral hazard model

Holmstrom and Tirole (1997)

- There is a continuum of firms with access to the same investment technology and different amounts of capital A.
- The distribution of assets across firms is described by the cumulative distribution function $G(A)$.
- The investment required is I, so a firm needs to raise I-A in external resources. The return is either 0 or R, and the probability depends on the type of project that the firm chooses.
- The firm may choose a lower type to enjoy private benefits.

Project	Good	Bad (low private benefit)	Bad (high private benefit)
Private benefit	0	b	B
Probabilit y of success	p_H	p_L	p_L

- The rate of return demanded by investors is denoted as γ, which can either be fixed or coming from a supply function $S(\gamma)$.
- The assumption is that only the good project is viable:

$$p_H R - \gamma I > 0 > p_L R - \gamma I + B.$$

- The incentive of the firm to choose the good project will depend on how much "skin in the game" it has.
- Hence, it would be easier to finance firms with large assets A, since they are more likely to internalize the monetary benefit and choose the good project.

Financial Intermediaries

- In addition to investors who demand a rate of return γ, there are financial intermediaries, who can monitor the firm.
- Monitoring is assumed to prevent the firm from taking a B project, hence reducing the opportunity cost of the firm from B to b.
- Monitoring yields a private cost of c to the financial intermediary.
- Intermediary capital K_m will be important to provide incentives to the intermediary to monitor the firm (the Diamond solution of diversification is not considered here).

Direct Finance

- Consider a contract where the firm invests A, the investor invests I-A, no one gets anything if the project fails, and in case of success the firm gets R_f and the investor gets R_u:

$$R_f + R_u = R$$

- A necessary condition is that the firm has an incentive to choose the good project:

$$p_H R_f \geq p_L R_f + B.$$

- Denoting $\Delta p = p_H - p_L$, we get the incentive compatibility constraint:

$$R_f \geq B/\Delta p$$

- This implies that the maximum amount that can be promised to the investors (the **pledgeable expected income**) is:

$$p_H(R - B/\Delta p)$$

- Due to the participation constraint:

$$\gamma(I - A) \leq p_H(R - B/\Delta p)$$

- This puts a financing constraint on the firm that depends on how much internal capital it has.
- Defining

$$\overline{A}(\gamma) = I - p_H/\gamma(R - B/\Delta p),$$

- We get that only firms with capital at or above $\overline{A}(\gamma)$ can invest using direct finance.
- This is the classic **credit rationing** result going back to Stiglitz and Weiss (1981). The firm cannot get unlimited amounts of capital, for proper incentives to develop, it needs to have "skin in the game".

Indirect Finance

- An intermediary can help relax the financing constraint of the firm by monitoring it and reducing its temptation to take the bad project.
- Now, the intermediary will get a share R_m of the return of the successful project

$$R_f + R_u + R_m = R$$

- The incentive constraint of the firm is now:

$$R_f \geq b/\Delta p$$

- There is also an incentive constraint for the intermediary:

$$R_m \geq c/\Delta p$$

- Then, the pledgeable expected income becomes:

$$p_H(R - (b + c)/\Delta p)$$

- Suppose that the intermediary is making a return of β (which has to exceed γ due to the monitoring cost), and invests I_m: $\beta = p_H R_m / I_m$, because of the incentive constraint it will contribute a least: $I_m(\beta) = p_H c/(\Delta p)\beta$.

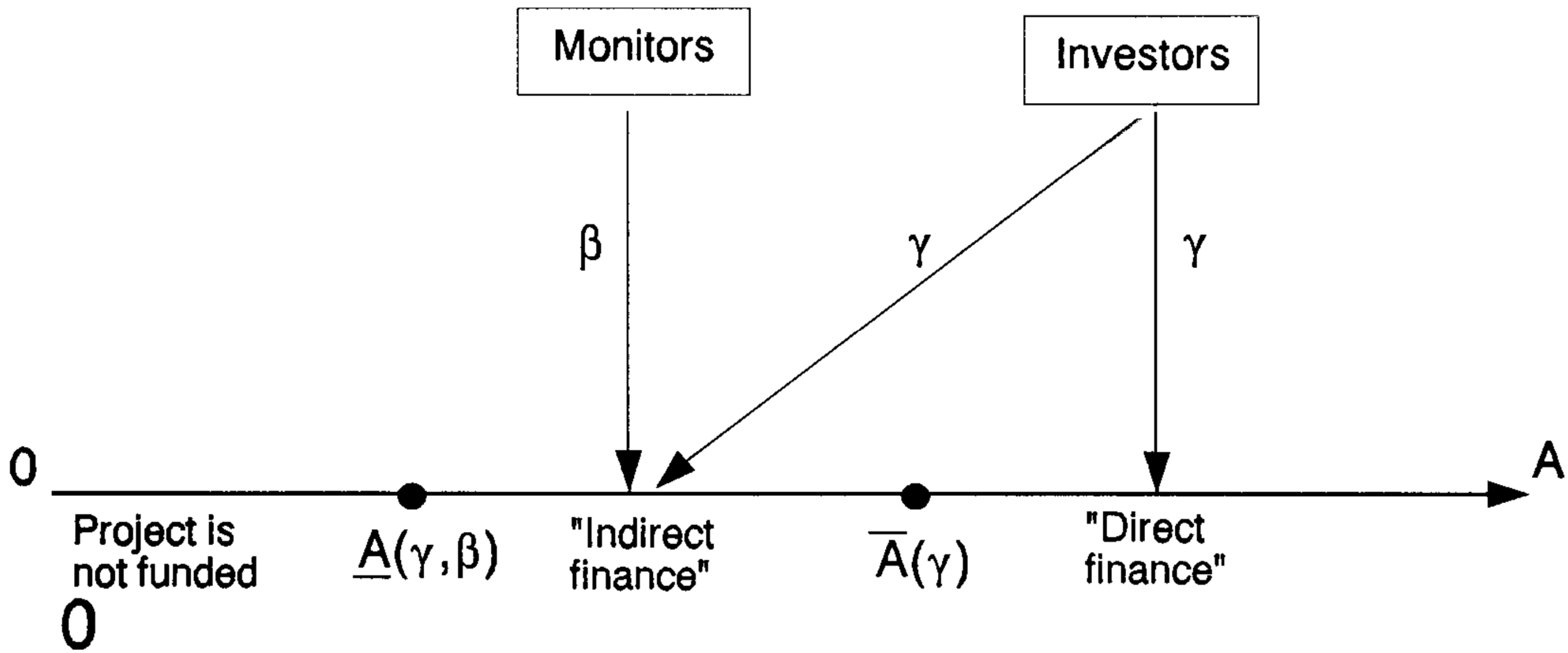
- Now, we can look at the financing constraint imposed by the participation constraint of the investors:

$$\gamma\big(I - A - I_m(\beta)\big) \leq p_H(R - (b + c)/\Delta p)$$

- This can be rewritten as:

$$A \geq \underline{A}(\gamma, \beta) = I - I_m(\beta) - p_H/\gamma(R - (b + c)/\Delta p)$$

- A firm with capital less than $\underline{A}(\gamma, \beta)$ cannot convince investors to supply it with capital even in the presence of intermediation. The firm will not increase reliance on intermediaries as their capital is more expensive.



- There are conditions in the paper guaranteeing that $\underline{A}(\gamma,\beta)$ is below $\overline{A}(\gamma)$.
- The result is that small firms are not financed at all, intermediate firms are financed by intermediaries and investors, and large firms are finance solely by investors.
- In equilibrium, the demand for capital equals the supply.
- The authors analyze the effects of decrease in the supply of capital.
- The main result is that the small firms are hurt most, as the squeeze leads to an increase in $\underline{A}(\gamma,\beta)$.

Relation to Crises

- Consider a negative aggregate shock in the economy, shifting the distribution of capital $G(A)$ to the left
  - This will be amplified via a multiplier effect
  - Entrepreneurs will face stricter financial constraints and will be less able to raise external capital
- Similarly, when the financial sector is hurt, leading to a reduction in K_m, an amplified effect on the economy will also arise
- Related empirical evidence have been provided by Gan (2007 a,b), Chaney, Sraer, and Thesmar (2011) and others

Frictions within the Financial Sector

- While the model above describes frictions in the flow of credit from the financial sector to the real economy, many of the insights apply to the flow of credit between financial institutions
- Rich literature on interbank markets, going back to Bhattacharya and Gale (1987) who analyze the under provision of liquidity in this market due to a free-rider problem
- Recent literature describes the repo market and its breakdown due to moral hazard and adverse selection problems: Martin, Skeie, and von Thadden (2014) and Kuong (2015). This was a key characteristic of the crisis

Link to macroeconomic models

- Financial multipliers of the type described above have been integrated heavily into macroeconomic models to study amplification and persistence over the business cycle
- Bernanke and Gertler (1989): A negative shock to the net worth of a borrower strengthens the agency problem against potential lenders, which reduces lending and investment in equilibrium
- Kiyotaki and Moore (1997): identify an important dynamic feedback mechanism amplifying this effect. The reduction in future investments is reflected in prices today, reducing net worth even further

Link to bank runs

- Credit frictions described here affect the asset side of financial institutions' balance sheets, whereas bank runs described before affect the liability side
- Importantly, the two can interact with each other and amplify each other: as assets deteriorate in value, incentives to run increase, and as runs increase, asset values deteriorate further
- Recently, Gertler and Kiyotaki (2015) combine the traditional macroeconomic model with moral hazard frictions in lending with fragility on the liability side due to potential runs. They analyze the extent to which runs further amplify the effects of shocks on the economy

Link to Currency Crises

- Credit frictions have also been shown to have important interactions with currency problems. Krugman (1999):
  - Firms have a currency mismatch between assets and liabilities (important fact for emerging economies, such as in the 1990s crises)
  - Real depreciation reduces their net worth
  - This implies they can borrow less and invest less
  - This leads to real depreciation, creating a self-fulfilling feedback loop and multiple equilibria
- Key question: why do firms have currency mismatch?

Contagion

Contagion

- One of the most striking features of financial Crises is that they spread across countries/institutions.
- Several leading explanations have been offered:
  - Information.
  - Interbank Connections.
  - Investors' portfolios readjustments.
  - Behavioral explanations.

Interbank Connections: Allen and Gale (2000)

- There are three dates: 0, 1, and 2; one good.
- Investment technology:
  - **Short term**: One unit invested in $t=0$ yields one unit in $t=1$.
  - **Long term**: One unit invested in $t=0$ yields R in $t=2$, or r in $t=1$; $0<r<1<R$.
- There are four different regions: A, B, C, and D.
  - Each region has a continuum [0,1] of agents, who might face liquidity shocks, as in Diamond and Dybvig.

- Utility is given by:

$$U(c_1, c_2) = \begin{cases} u(c_1)\ prob & \omega \\ u(c_2)\ prob\ 1-\omega \end{cases}$$

- The probability of a liquidity shock varies from region to region; there are two equally likely states:

TABLE 1

REGIONAL LIQUIDITY SHOCKS

	A	B	C	D
S_1	ω_H	ω_L	ω_H	ω_L
S_2	ω_L	ω_H	ω_L	ω_H

Optimal Risk Sharing

- Denote $\gamma = (\omega_H + \omega_L)/2$.

- Planner maximizes:

$$\gamma u(c_1) + (1-\gamma)u\left(\frac{1-\gamma c_1}{1-\gamma}R\right)$$

- Hence,

$$u'(c_1) = u'\left(\frac{1-\gamma c_1}{1-\gamma}R\right)R$$

- Achieved by investing c_1 in short asset and $\frac{1-\gamma c_1}{1-\gamma}$ in long asset.
- First-best allocation satisfies incentive-compatibility constraint
  - Thus, first-best can be achieved even if types are not observable.
- The allocation ignores division to regions, and resources move across them to absorb liquidity needs.
- In particular, the planner will shift resources across regions.
  - In state 1, $(\omega_H - \gamma)c_1$ moves from B and D to A and C in *t=1*, and $(\omega_H - \gamma)c_2$ moves from A and C to B and D in *t=2*.

Decentralization

- In each region, consumers deposit their endowments in banks, who offer demand deposit contracts.
- Banks hold deposits in banks of other regions. Suppose the market is incomplete:

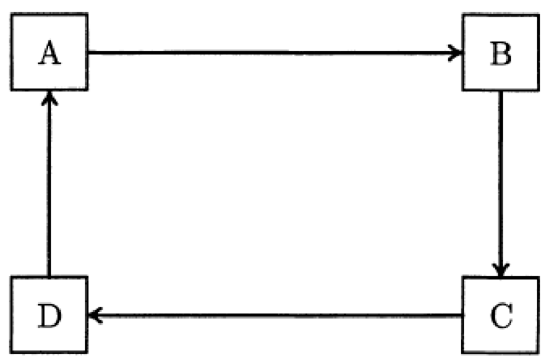


FIG. 2.—Incomplete market structure

- How can banks achieve the first best?
  - They make investments and promise returns as the planner.
  - They hold deposits of $\omega_H - \gamma$ at banks at the adjacent region.
  - In *t=1* banks in regions with high liquidity needs liquidate the deposits at banks in regions with low liquidity needs.
  - In *t=2* banks in regions with low liquidity needs liquidate the deposits at banks in regions with high liquidity needs.
- The fact that banks with low liquidity needs hold deposits in banks with high liquidity needs and vice versa guarantees efficient allocation.

Fragility

- Assume the same allocation as before, but a new state is possible:

TABLE 2

REGIONAL LIQUIDITY SHOCKS WITH PERTURBATION

	A	B	C	D
S_1	ω_H	ω_L	ω_H	ω_L
S_2	ω_L	ω_H	ω_L	ω_H
$\bar{S}$	$\gamma + \epsilon$	γ	γ	γ

- The new state is assigned probability zero; in it, aggregate demand for liquidity requires liquidation of some long-term assets.

- Assume that deposits are liquidated before long-term assets:

$$1 < \frac{c_2}{c_1} < \frac{R}{r}$$

- Banks start liquidating deposits in each other, and banks in region A liquidate some long-term assets.
- If aggregate liquidity shock is large enough, banks in region A must go bankrupt:
    - They liquidate long-term assets to pay early withdrawals, and cannot pay enough to patient investors, who then decide to run.

- If liquidation value is sufficiently low, banks in region D will also go bankrupt.
    - The value of their deposits in region A is low, so they liquidate long-term assets and trigger a run.
- By induction, banks in regions B and C will also go bankrupt.
- Overall, the failure of banks in region A, triggers a failure of region D, which triggers a failure of Region C, which triggers a failure of region B.

Interbank Structures that Reduce Fragility

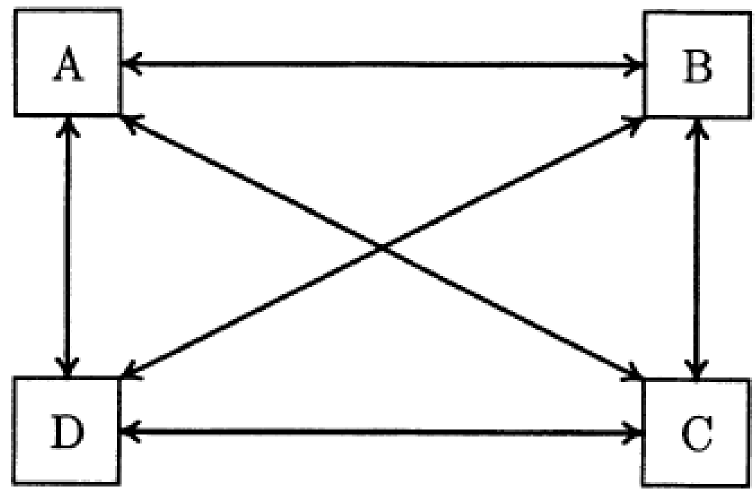


FIG. 1.—Complete market structure

- No bank depends strongly on banks in region A. The damage is spread out evenly, and not big enough to fail other regions.

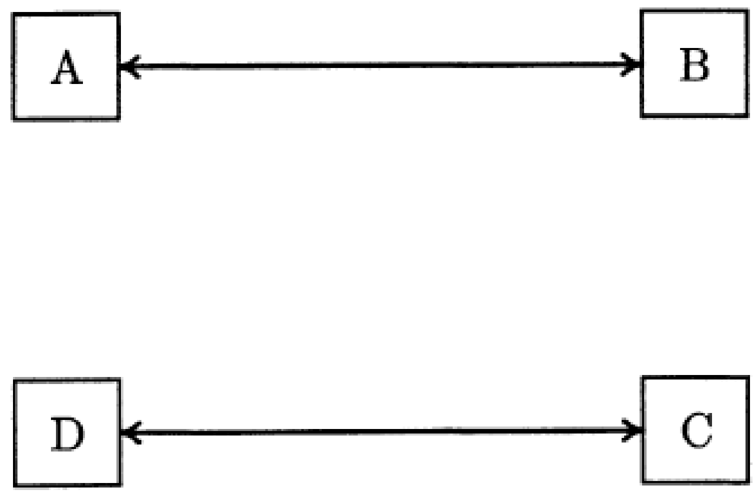


FIG. 3.—Disconnected incomplete market structure

- Failures are limited to regions A and B.
- Overall, the link between market completeness and fragility of the system is non-monotone.

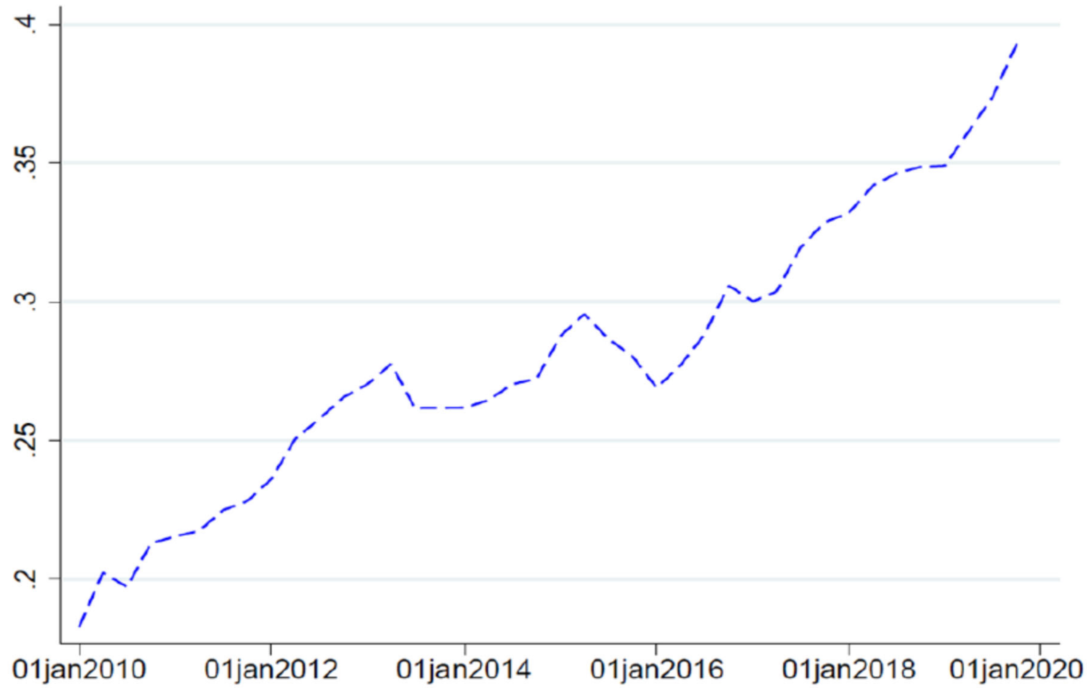
Recent Developments and New Directions

The Fragility of Mutual Funds and Bond Markets

Covid-19 Episode

- In March 2020, following the onset of Covid-19 crisis, there was major turmoil in bond markets.
- Fragilities in corporate-bond mutual funds were on display, as investors rushed to take their money out.
- Quick intervention by the Fed prevented a bigger meltdown.
- This episode and the policy response are studied in Falato, Goldstein, and Hortacsu (2021).

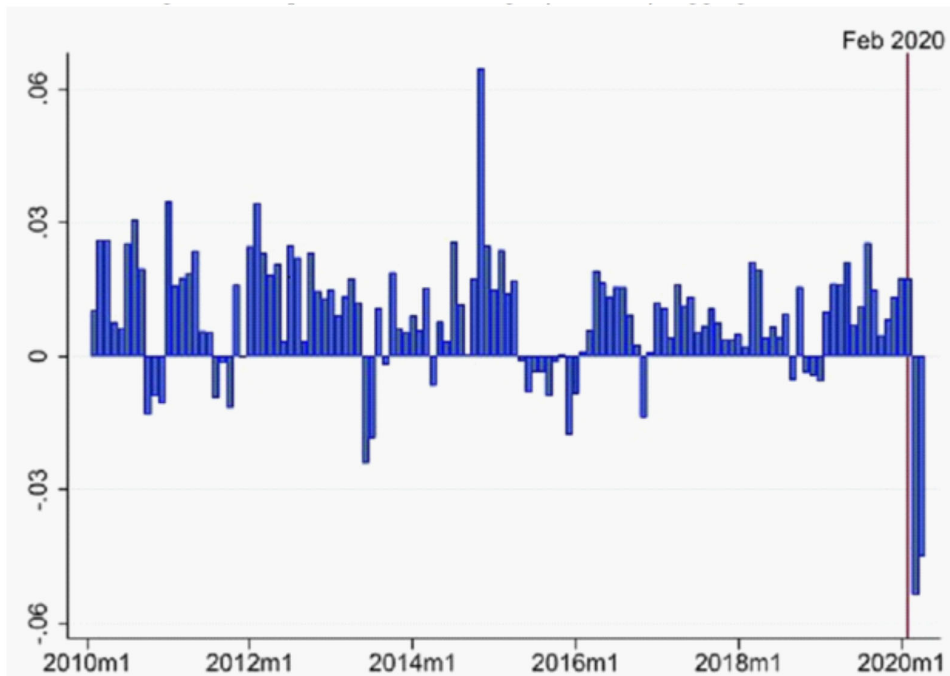
The Growing Importance of Investment Funds in the Corporate Bond Market



Aggregate Net Asset Value of Funds and ETFs divided by Size of Market (from Fed Flow of Funds Z.1)

Outflows in Covid-19 Crisis in Perspective of Recent Decade

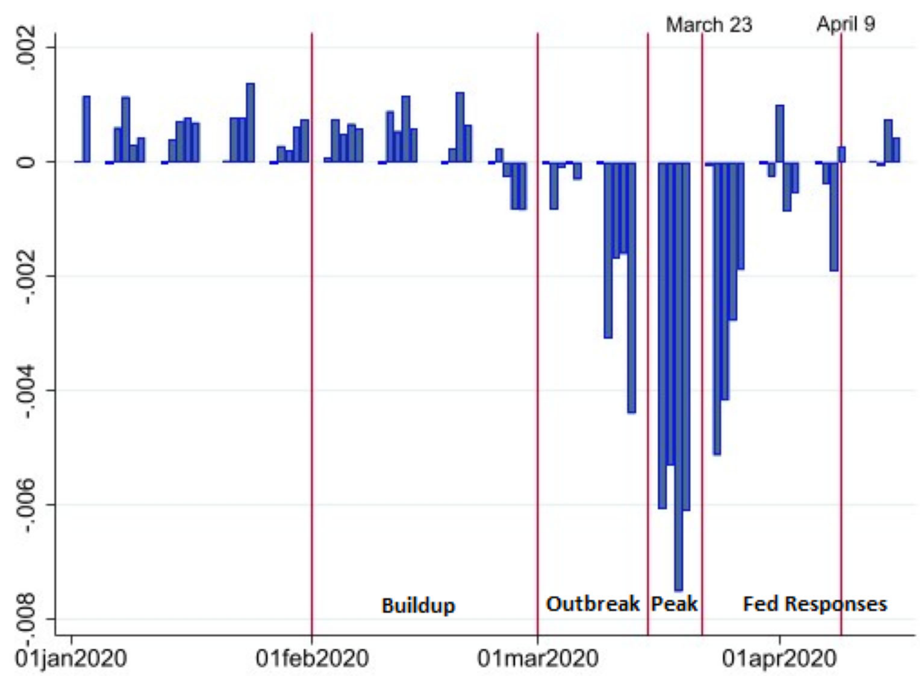
- Mutual funds in corporate bond markets saw massive outflows during the COVID-19 crisis.



Aggregate Net Fund Flows as a Fraction of Lagged Net Asset

Evolution of Flows over the Crisis

- Daily outflows started in the last week of February and accelerated in the second and third weeks of March, peaking at almost 1% of net assets.



Daily Aggregate Net Fund Flows (Fraction of Lagged Net Asset)

Sources of Fragility

- Liquidity mismatch: holding illiquid assets, but providing high level of liquidity to their investors – can lead to "run" type behavior from their investors: Chen, Goldstein, and Jiang (2010); Goldstein, Jiang, and Ng (2017).

- Fire-sale vulnerability: forced asset sales have spillovers on peer funds that can lead to outflows: Falato, Hortacsu, Li, and Shin (2020).

- What happened in Covid-19 crisis? Did these fragilities play a role? Falato, Goldstein, and Hortacsu (2021) show evidence that this was indeed the case.

Other Important Observations from covid-19 episode

- ETFs showed much greater resilience during crisis period than mutual funds. Payoff structure does not create as much liquidity transformation, and so less prone to fragility.
- Equity mutual funds also showed resilience, according to Pastor and Vorsatz (2020). They also provide lower liquidity transformation.
- Stress in corporate bond markets had peculiar features, whereby more liquid securities experienced greater dislocations, as documented by Haddad, Moreira, and Muir (2021). Evidence by Ma, Xiao, and Zeng (2021) ties this to mutual funds following a pecking order.

Policy Lessons Going Forward

- The Federal Reserve interventions were crucial for alleviating the stress. Quick reversal of outflows after two announcements.
- Sustained recovery of flows over the post-crisis period (through August 2020) for funds that held more bonds eligible for purchase by the Fed facilities (as shown in Falato, Goldstein, and Hortacsu, 2021).
- Relying on such interventions in the future might not be sustainable. Moral hazard problem with funds taking excessive risks as they expect outside intervention. If government provides a safety net, then other regulatory measures should be in place to promote resilience, like in banks.

Policy Lessons Going Forward – Cont'd

- Going forward, underlying vulnerabilities should be assessed and potentially addressed:
  - Improving liquidity of underlying corporate bond assets: These are difficult reforms to enact.
  - Requiring funds to hold more liquid securities: Might defeat the purpose of having corporate-bond funds.
  - Reducing liquidity available to investors:
    - Swing pricing is a solution that is directly targeted to the problem.
    - ETF structure acts as natural swing pricing.

Recent Evidence of Bank Fragility

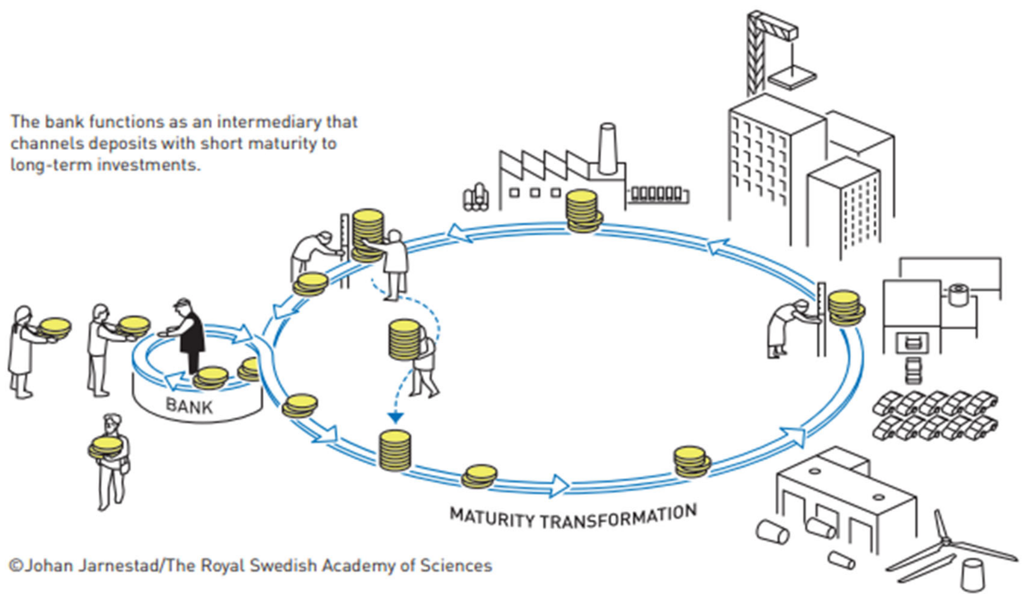
Silicon Valley Bank, Twitter-Fueled Bank Run, 2023

Liquidity Transformation and Fragility

- A lot of it is very familiar: Liquidity transformation is at the core of banks' business model.
    - By providing liquid deposits and investing in illiquid loans, banks create liquidity, but end up with liquidity mismatch on their balance sheets.
    - Liquidity mismatch renders banks vulnerable to panic-based runs (Diamond and Dybvig, 1983). Depositors rush to withdraw deposits expecting that others will do so.

Liquidity Transformation and Fragility - Cont'd

- Illustration from recent Nobel Prize:

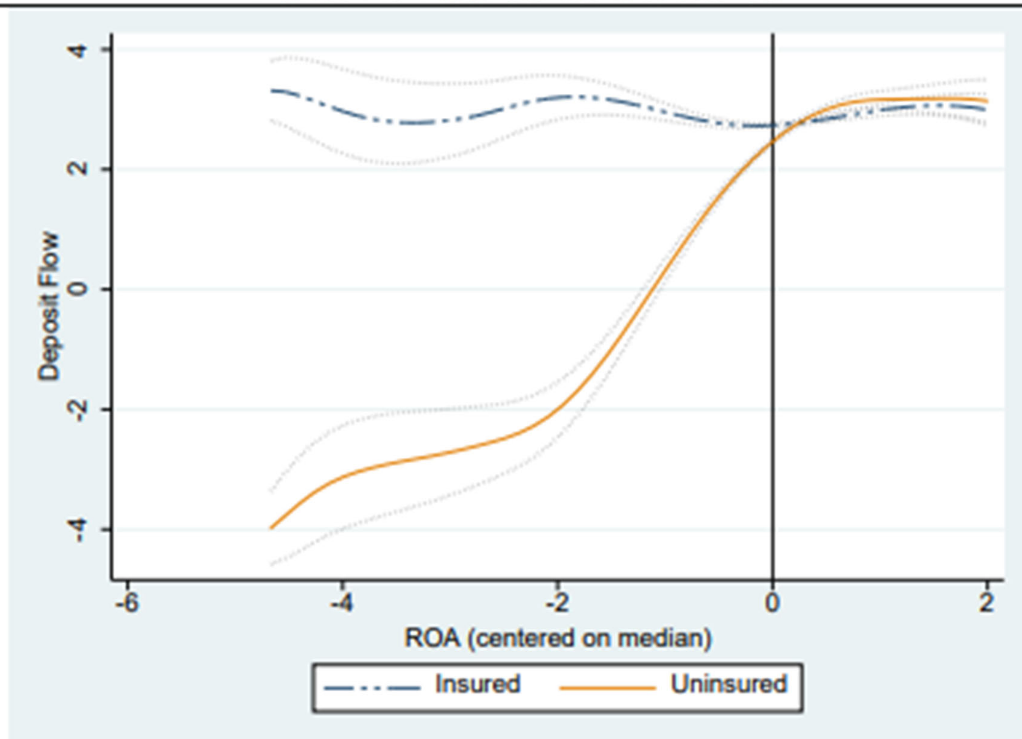


Broad-Base Evidence of Fragility in the Banking Sector

- While the above forces are well known, concerns over fragility of banking sector have decreased over the years with many regulatory measures in place.
- Chen, Goldstein, Huang, and Vashishtha (2022):
    - Using universe of US bank data between 1993 and 2016.
    - Uninsured deposits are flighty and respond negatively to performance decrease.
    - Uninsured deposits respond more strongly when banks perform greater liquidity transformation.

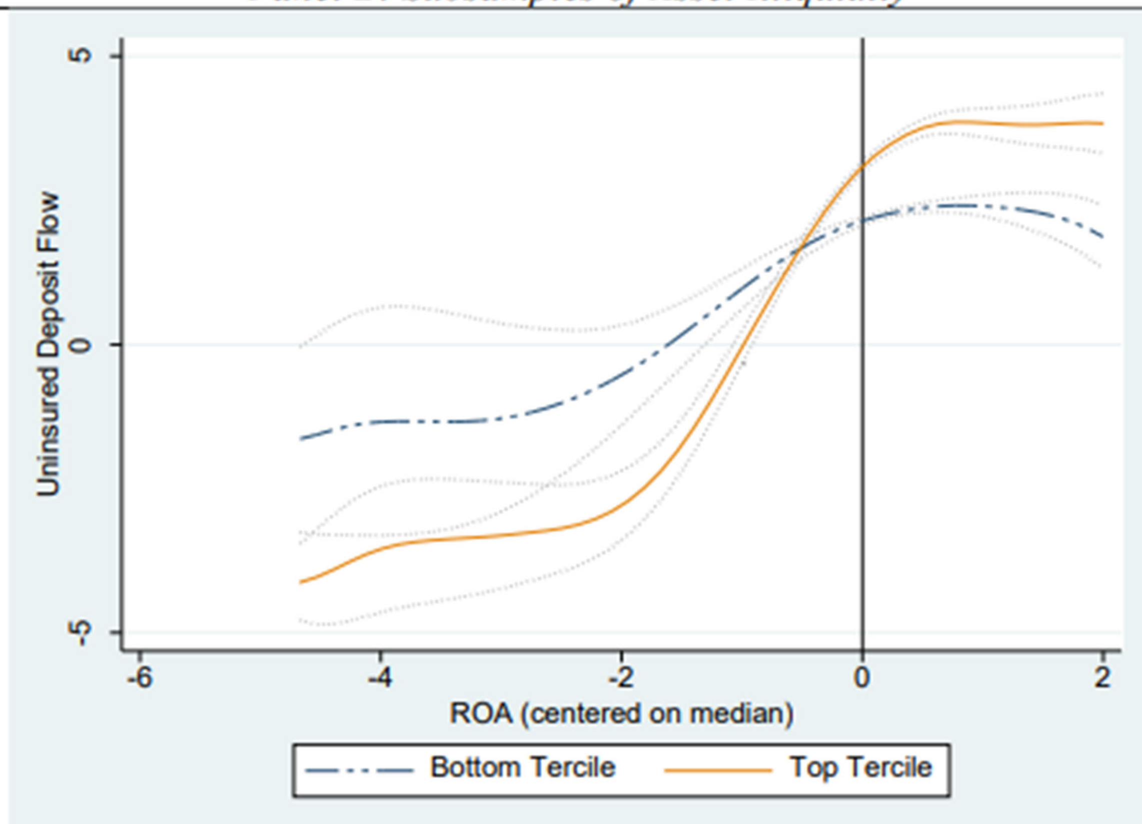
Deposit Flow: Insured vs. Uninsured

Panel A: Insured vs. uninsured

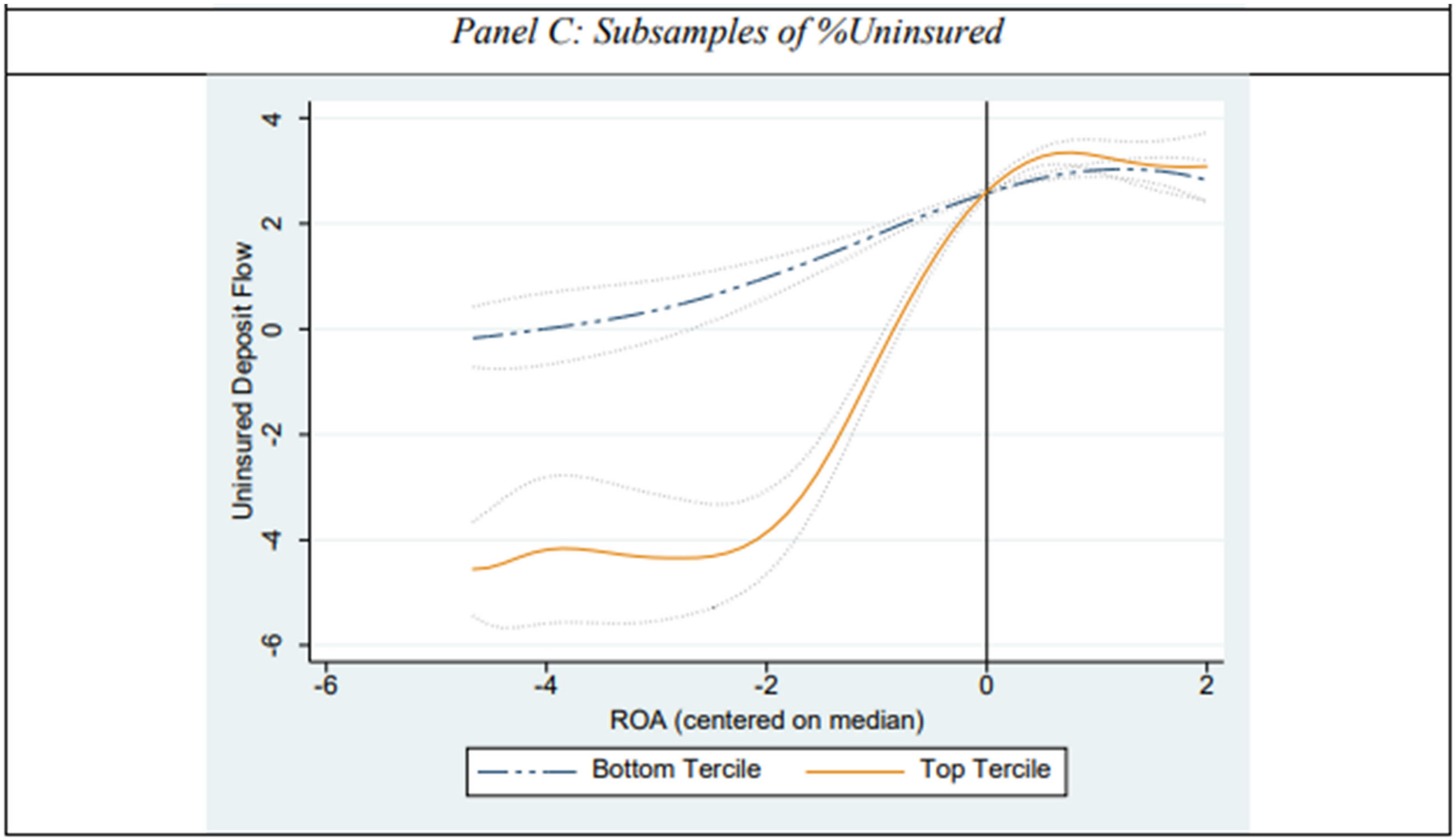


Uninsured Deposit Flow and Asset Illiquidity

Panel B: Subsamples of Asset Illiquidity



Uninsured Deposit Flow and Uninsured Deposit Base



Some Regression Results

- Banks with more illiquid assets and/or uninsured deposits:
  - Exhibit stronger sensitivity of uninsured deposits outflows to bad performance.
  - Exhibit higher outflows conditional on low performance.
- Pattern is reversed for insured deposits:
  - Banks raise insured deposits to substitute for uninsured ones.
  - Yet, this is generally not enough to completely compensate banks for deposit loss.
- Pattern is stronger when performance shock is systematic than when it is idiosyncratic:
  - Complementarities strengthen when aggregate conditions are bad.

Takeaways from Research and Recent Events

- Banks are fragile:
  - Recent focus has been on other institutions.
  - But fundamental and panic risks are still prevalent in banks.
- Deposit insurance involves tradeoffs:
  - It is an important tool.
  - But it cannot be unlimited.
- Other regulatory tools should be strengthened:
  - Increased scrutiny of mid-size banks.
  - More imaginative stress tests.
  - Capital and liquidity regulation.