Limits of Arbitrage with Optimal Contracts

PETER M. DEMARZO
Stanford University and NBER

DIMITRI VAYANOS
London School of Economics, CEPR and NBER

CRETE Conference

Ios, 10 July 2026

Intermediaries in Capital Markets

- Prices in capital markets are determined to a large extent by financial intermediaries rather than households.
- Role of financial intermediaries is increasingly acknowledged in:
  - Finance. Intermediary-based asset pricing.
  - Macroeconomics.
  - International macroeconomics.
- Key themes:
  - Intermediaries absorb imbalances between demand and supply.
    - In doing so, they bring prices closer to fundamentals.
  - Intermediaries have limited risk-absorption capacity.
    - Their capacity is low when they have low wealth or poor performance.

Example: Term Structure of Interest Rates

- Assume decrease in demand for long-term bonds, with no change to current and future expected short rates.

- Impact on term structure:
  - Intermediaries must accommodate decrease in demand by buying long-term bonds.
  - Prices of long-term bonds must decrease ($\rightarrow$ long rates must increase) to induce intermediaries to buy.
  - Impact is large when intermediaries have low risk-absorption capacity.

  Vayanos-Vila (ECMA 2021).

- Time-varying risk premia.
  - Impact on term structure is through risk premia.
  - Prices move without change in "fundamentals."

- Similar mechanisms have been explored in other markets:
  - Stocks. *Greenwood (JFE 2005), He-Krishnamurthy (RES 2012, AER 2013)*
  - Currencies. *Gabaix-Maggiori (QJE 2014), Itskokhi-Mukhin (JPE 2021)*
  - Derivatives. *Garleanu-Pedersen (RFS 2011)*
  - Mortgage-backed securities. *Gabaix-Krishnamurthy-Vigneron (JF 2007)*

Intermediation Frictions

- For intermediaries to matter, they must face frictions in raising capital from households.
  - In the absence of frictions, intermediary-based models reduce to standard models with households. Representative agent, etc.

- Examples of assumed frictions:
  - Intermediaries can raise no equity capital from households.
    - External finance can only be in the form of debt.

    Grossman-Miller (JF 1988), Geanakoplos (1997), Gromb-Vayanos (JFE 2002), Brunnermeier-Sannikov (AER 2014)
  - Equity capital that households invest with intermediaries is exogenous increasing function of intermediaries' past performance.

    Shleifer-Vishny (JF 1997)
  - Equity capital is determined by static contracting even though relationship is dynamic.

    He-Krishnamurthy (RES 2012, AER 2013)

- Strong assumptions on contracting.
  - Emphasis is on equilibrium asset prices rather than on contracting.

Contracting Model

- Optimal contracting between intermediaries (agents) and households (principals).
  - No restriction to debt.
- Dynamic aspects are important.
  - Poor performance $\rightarrow$ Principals invest less capital with agents.
- Dynamic financial contracting.
  - Agency arises from moral hazard. Agent can "steal."
  - Agent can be rewarded using money or future continuation utility.
  - Used in corporate finance to study firms' capital structure and investment decisions.

DeMarzo-Sannikov (JF 2006), Biais-Mariotti-Platin-Rochet (RES 2007), DeMarzo-Fishman (RFS 2007).

This Paper

- Intermediaries raise capital from households using optimal contracts.
- Use dynamic financial contracting, bringing it into equilibrium asset pricing.
  - Two-way feedback between contracts and equilibrium asset prices.

Main Results

- Effectiveness of arbitrage depends on size of imbalances between demand and supply.
  - Imbalances are absorbed with small price impact when they are small.
  - Imbalances have large price impact and can generate amplification when they are large.
- Longer-term contracts between principals and agents result in more effective arbitrage.
  - Lower price impact.
  - Lower price volatility feeds back into better contracting.
- Possible implications for contracts in asset management.
  - Should longer-term contracts be encouraged?

Baseline Model

Assets

- Three periods $t = 0, 1, 2$.
- Riskless asset, zero net return.
- Risky asset.
  - Dividend D_2 in period 2, distributed in $[0, \infty)$ with continuous density $f_2(D_2)$.
  - Price S_t in period $t = 0, 1$.
  - Supply $\theta \geq 0$ shares.

Agents

- Direct traders.
- Delegated traders.
  - Pairs of principals and agents.
- Noise traders.
  - Supply u_t shares in period $t = 0, 1$.
  - u_0 deterministic.
  - u_1 distributed in $[0, \infty)$ with continuous density $f(u_1)$.
    - Imbalance. What will be its price impact?
  - Noise traders can be folded into direct traders for a welfare analysis.

Direct Traders

- Competitive, continuum with measure one.
- Risk-neutral with utility

$$W_{2d} - \frac{\alpha_{0d}}{2} I_{0d}^2 - \frac{\alpha_{1d}}{2} I_{1d}^2.$$

  over wealth W_{2d} in period 2 and investment $I_{td} \geq 0$ in risky asset in period $t = 0, 1$.
  - Quadratic terms are reduced form for risk aversion.
  - α_{td} is risk-aversion coefficient in period $t = 0, 1$.
- Endowed with asset supply θ in period 0.

Delegated Traders

- Pairs of principals and agents. Competitive, continuum with measure one.

- Agents.
  - Hold risky asset in a fund on behalf of principals.
  - Observe asset prices and dividend.
  - Moral hazard: Can reduce fund value and divert some of the money.
    - Reduce fund value by $X \geq 0 \rightarrow$ Divert ϵX, where $\epsilon \in [0, 1]$.
    - No agency when $\epsilon = 0$.
    - Diversion can occur in periods $t = 1, 2$.

- Principals.
  - Do not observe asset prices and dividend. Observe certified reports of fund value post-diversion.
    - Reports are observed in periods $t = 1, 2$.
  - Risk-neutral with utility
    $$W_2 - \frac{\alpha_0}{2} I_0^2 - \frac{\alpha_1}{2} I_1^2$$
    over wealth W_2 in period 2 and investment $I_t \geq 0$ in period $t = 0, 1$.

Contracts

- Contracts are written in period 0 and can depend on reports of fund value in periods $t = 1, 2$.

- Contracting variables:
  - Investment $I_0 \geq 0$ in period 0.
  - Investment $I_1 \geq 0$ in period 1.
    - Can depend on report $\hat{V}_1$ about fund value V_1 in period 1, or equivalently on report $\hat{u}_1$ about u_1.
  - Cash payment $C_2 \geq 0$ to agent in period 2.
    - Can depend on reports $(\hat{V}_1, \hat{V}_2)$ about fund value (V_1, V_2) in periods 1 and 2, or equivalently on $\hat{u}_1$ and on report $\hat{D}_2$ about D_2.

Equilibrium in Period 1

Optimization by a Direct Trader

- Direct trader's utility in period 1 is

$$\underbrace{\theta S_0 + I_{0d}\frac{S_1 - S_0}{S_0} + I_{1d}\frac{\mathbb{E}(D_2) - S_1}{S_1}}_{\text{expected wealth}} - \frac{\alpha_{0d}}{2} I_{0d}^2 - \frac{\alpha_{1d}}{2} I_{1d}^2.$$

- First-order condition is

$$I_{1d} = \frac{\mathbb{E}(D_2) - S_1}{\alpha_{1d} S_1}.$$

- Combining with market clearing

$$I_{1d} + I_1 = (\theta + u_1) S_1,$$

  yields equilibrium price in period 1.

- Assume S_1 decreasing in u_1. Verify that this holds in equilibrium.

Incentive Constraint in Period 2

- C_2 must rise with $\hat{V}_2$ to deter under-reporting. Constraint is binding, implying

$$C_2(\hat{V}_1, \hat{V}_2) = \epsilon \hat{V}_2 + \underline{C}_2(\hat{V}_1).$$

  Cash payment consists of
  - Performance fee $\epsilon \hat{V}_2$.
  - Base payment $\underline{C}_2(\hat{V}_1)$. Can be interpreted as fee in period 1.

- Since in equilibrium $\hat{V}_1 = V_1$ and $\hat{V}_2 = V_2 = I_1(V_1)\frac{D_2}{S_1}$, agent's utility in period 1 is

$$U_{1a}(V_1) = I_1(V_1)\frac{\mathbb{E}(D_2)}{S_1} + \underline{C}_2(V_1).$$

  - Utility increases in $I_1(V_1)$ (fund's AUM).
  - Can compensate agent with AUM in period 1 ($\rightarrow$ fee in period 2) and fee in period 1.

- Contract choice reduces to choosing $(I_0, I_1(u_1), \underline{C}_2(u_1))$. Take I_0 as given for now.

Optimization by a Principal

- Principal's utility in period 0 is

$$\underbrace{\mathbb{E}\left[I_0\frac{S_1(u_1)-S_0}{S_0}+I_1(u_1)\frac{\mathbb{E}(D_2)-S_1(u_1)}{S_1(u_1)}\right]}_{\text{expected wealth gross of agent's utility}}$$
$$-\frac{\alpha_0}{2}I_0^2-\frac{\alpha_1}{2}\mathbb{E}\left[(I_1(u_1))^2\right]-\underbrace{\mathbb{E}\left[I_1(u_1)\frac{\mathbb{E}(D_2)}{S_1(u_1)}+\underline{C}_2(u_1)\right]}_{\text{agent's utility}}.$$

- Principal maximizes utility subject to agent's incentive constraint in period 1 (IC1)

$$\underbrace{I_1(u_1)\frac{\mathbb{E}(D_2)}{S_1(u_1)}+\underline{C}_2(u_1)}_{\substack{\text{expected utility}\\ \text{for report } u_1}}\geq\underbrace{\epsilon I_0\frac{S_1(u_1)-S_1(\hat{u}_1)}{S_0}}_{\text{diversion proceeds}}+\underbrace{\epsilon I_1(\hat{u}_1)\frac{\mathbb{E}(D_2)}{S_1(u_1)}+\underline{C}_2(\hat{u}_1)}_{\substack{\text{expected utility}\\ \text{for report } \hat{u}_1}},$$

for all u_1 and $\hat{u}_1 \geq u_1$.

Key Trade-Off

- In the absence of (IC1), principal sets $\underline{C}_2(u_1) = 0$ and

$$\eta_1(u_1) = \frac{(1-\epsilon)\mathbb{E}(D_2) - S_1(u_1)}{\alpha_1 S_1(u_1)} \equiv I_1^{*P}(u_1).$$

  - $\eta_1^{*P}(u_1)$ is smaller than first-best (FB) value

$$\frac{\mathbb{E}(D_2) - S_1(u_1)}{\alpha_1 S_1(u_1)} \equiv I_1^{*}(u_1),$$

    which maximizes sum of utilities of principal and agent, because higher $I_1(u_1)$ raises agent's compensation.

- If principal sets $I_1(u_1) = I_1^{*P}(u_1)$ for all u_1, then must raise $\underline{C}_2(u_1)$ for low values of u_1 (high values of S_1) to deter diversion by agent.

Carrots

- Principal economizes on $\underline{C}_2(u_1)$ by lowering $I_1(u_1)$ for low values of u_1.

Sticks

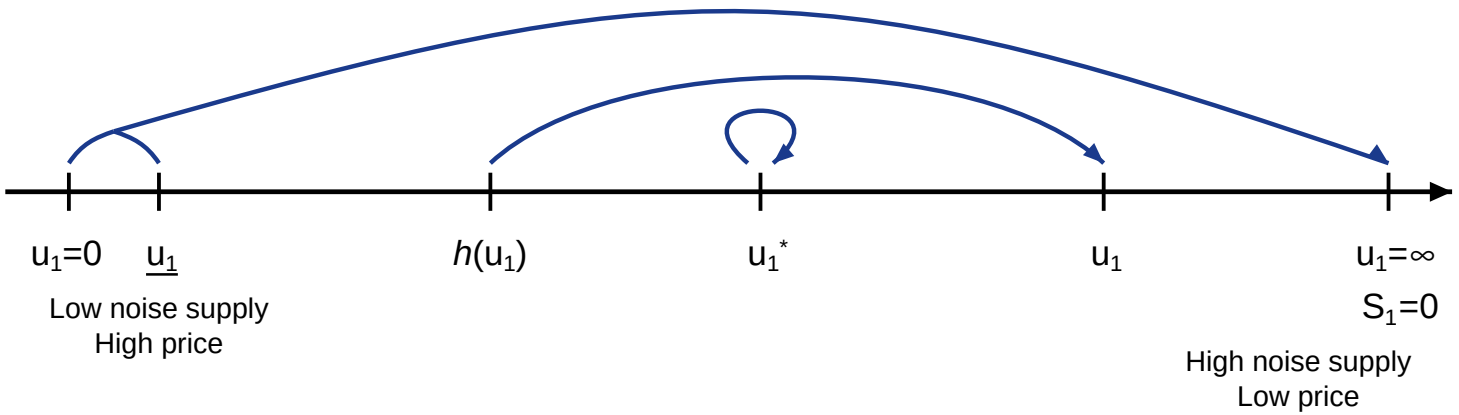
Sorting Property

- (IC1) is two-dimensional: $u_1 \in [0, \infty)$ and $\hat{u}_1 \in [u_1, \infty)$.
  - For which values is (IC1) binding?
  - Is local (IC1) binding, as is common in mechanism design settings?
- Sorting Property: Stronger incentive to divert when agent observes lower u_1.
- Intuition: Agents with lower u_1 have worse investment opportunities $\rightarrow$ Less affected when principal cuts investment.

Relaxed Problem

- Assume (IC1) only between $h(u_1)$ and u_1, where $u_1 \in [u_1^*, \infty)$ and $h(u_1)$ is a decreasing function from $[u_1^*, \infty)$ to $(\underline{u}_1, u_1^*]$.

- Agent observing $h(u_1)$ must prefer
  - To report $h(u_1)$ truthfully.
  - Than to divert and report $u_1 > h(u_1)$.



- (IC1) becomes one-dimensional, but not local.

Solution of Relaxed Problem

- Sticks: For all $u_1 \in [u_1^*, \infty)$,
  - $I_1(u_1) < I_1^{*P}(u_1)$.
  - $\underline{C}_2(u_1) = 0$.

- Carrots: For all $u_1 \in [0, u_1^*)$,
  - $I_1(u_1) > I_1^{*P}(u_1)$.
  - $I_1(u_1) = I_1^*(u_1)$ when $\underline{C}_2(u_1) > 0$.
    - Compensate with cash rather than AUM when $I_1(u_1)$ reaches $I_1^*(u_1)$.

- First-order condition (ODE1) wrt $I_1(u_1)$ for $u_1 \in [u_1^*, \infty)$:
  - When $\underline{C}_2(h(u_1)) = 0$,

$$\underbrace{\left(I_1^{*P}(u_1) - I_1(u_1)\right) f(u_1)}_{\text{benefit of raising } I_1(u_1)} = \underbrace{\left(I_1^{*P}(h(u_1)) - I_1(h(u_1))\right) f(h(u_1))h'(u_1)}_{\text{cost of raising } I_1(h(u_1))}.$$

  - When $\underline{C}_2(h(u_1)) > 0$,

$$\underbrace{\left(I_1^{*P}(u_1) - I_1(u_1)\right) f(u_1)}_{\text{benefit of raising } I_1(u_1)} = \underbrace{-\frac{\epsilon \mathbb{E}(D_2)}{\alpha_1 S_1(h(u_1))} f(h(u_1))h'(u_1)}_{\text{cost of raising } \underline{C}_2(h(u_1))}.$$

Determination of $h(u_1)$

- Solution of relaxed problem has same general form regardless of $h(u_1)$.
- $h(u_1)$ can be chosen so that solution of relaxed problem satisfies (IC1) for all $(u_1, \hat{u}_1)$.
- **Proposition:** Solution of relaxed problem satisfies (IC1) for all $(u_1, \hat{u}_1)$ if reporting u_1 is locally optimal for agent observing $h(u_1)$.
  - Since local optimality requires that $\hat{u}_1 = u_1$ maximizes

$$\epsilon l_0 \frac{S_1(h(u_1)) - S_1(\hat{u}_1)}{S_0} + \epsilon l_1(\hat{u}_1) \frac{\mathbb{E}(D_2)}{S_1(h(u_1))} + \underline{C}_2(\hat{u}_1)$$

and $\underline{C}_2(u_1) = 0$ for $u_1 \in [u_1^*, \infty)$, first-order condition (ODE2) is

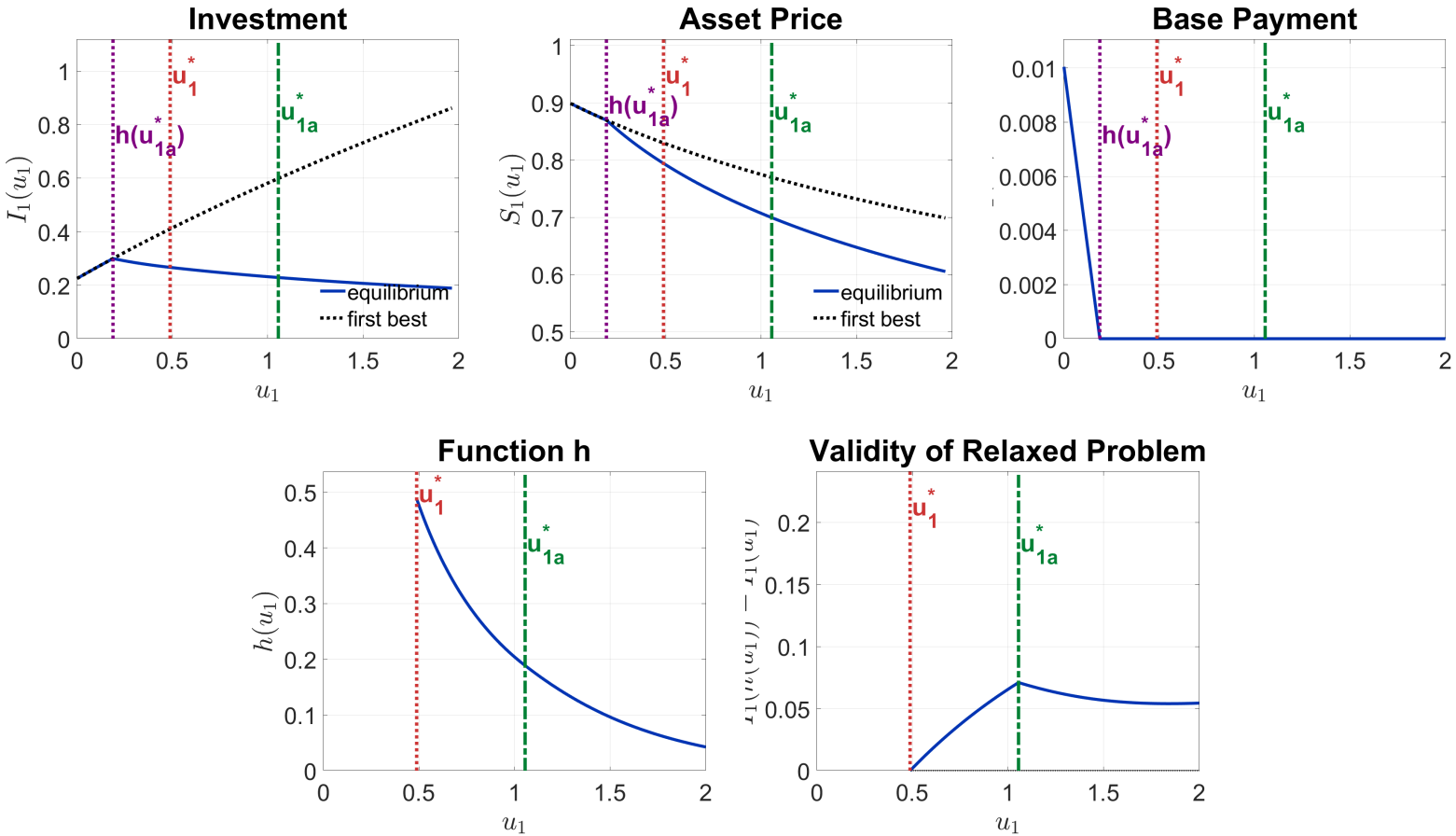
$$-l_0 \frac{S_1'(u_1)}{S_0} + l_1'(u_1) \frac{\mathbb{E}(D_2)}{S_1(h(u_1))} = 0.$$

- Proposition holds if $l_1(h(u_1)) \geq l_1(u_1)$ for all $u_1 \in [u_1^*, \infty)$. Otherwise, solve a modified relaxed problem in which local (IC1) binds for small u_1.

Equilibrium in Period 1

- Equilibrium:
  - $(I_1(u_1), h(u_1))$ solve ODE1 and ODE2 given $S_1(u_1)$.
  - $S_1(u_1)$ is equilibrium price given $I_1(u_1)$.
- Problem reduces to solving ODE system in $(I_1(u_1), h(u_1))$.
- Threshold u_1^* is determined by $h(\infty) = \underline{u}_1$ and by $\underline{u}_1 = 0$ if $I_1(\infty) = 0$.

Equilibrium in Graphs



- Small u_1.
  - Principal's investment $I_1(u_1)$ takes FB value. Increases in u_1.
  - Small price impact.
- Large u_1.
  - Principal's investment significantly below FB value. Decreases in u_1.
  - Large price impact.

Extended Model

Assets

- Add a period 3.
- For now:
  - Risky asset is as in baseline model.
  - New asset can be traded in period 2.
    - Dividend D_3 in period 3, distributed in $[0, \infty)$ with continuous density $f_3(D_3)$.
    - Exogenous deterministic price S_2 in period 2.
- Work in progress:
  - Risky asset pays a dividend D_2 in period 2 and D_3 in period 3.
  - Endogenous price S_2 in period 2 (and in periods 0 and 1).

Contracts

- Contracts are written in period 0 and can depend on reports of fund value in periods $t = 1, 2, 3$.

- Contracting variables:
  - Investment $I_0 \geq 0$ in period 0.
  - Investment $I_1 \geq 0$ in period 1.
    - Can depend on report of fund value in period 1, or equivalently on report $\hat{u}_1$ about u_1.
  - Investment $I_2 \geq 0$ in period 2.
    - Can depend on report of fund value in periods 1 and 2, or equivalently on $\hat{u}_1$ and on report $\hat{D}_2$ about D_2.
  - Payment $C_3 \geq 0$ to agent in period 3.
    - Can depend on reports of fund value in periods 1, 2 and 3, or equivalently on $\hat{u}_1$, $\hat{D}_2$ and on report $\hat{D}_3$ about D_3.

Optimization by a Principal in Period 1

- Agent's utility in period 1 is

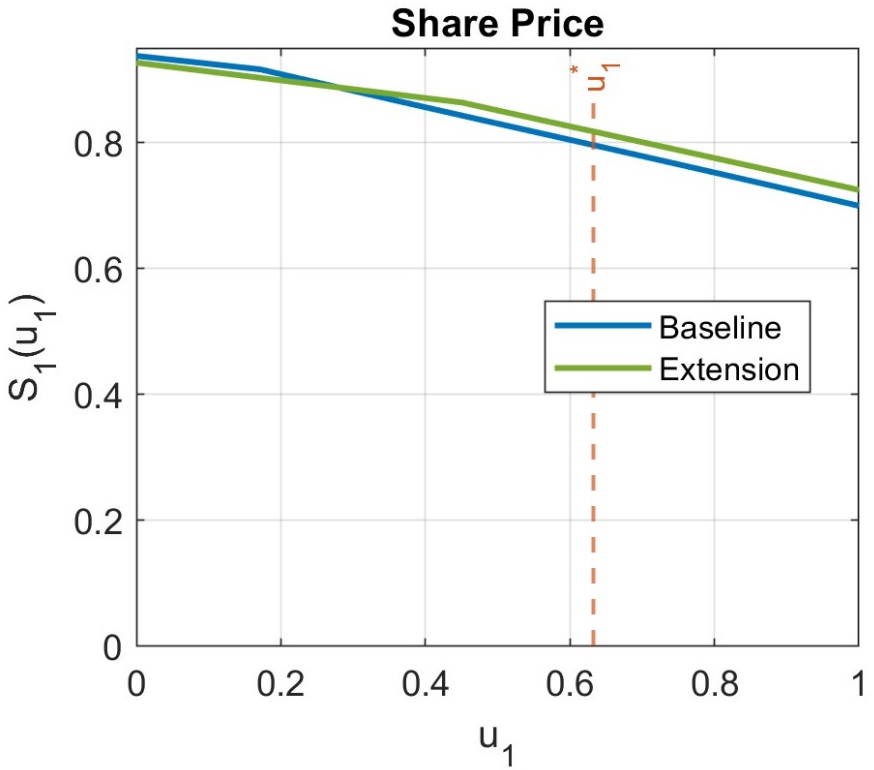
$$U_{1a}(u_1) = \epsilon I_1(u_1)\frac{\mathbb{E}(D_2)}{S_1(u_1)} + \epsilon \underline{I}_2(u_1)\frac{\mathbb{E}(D_3)}{S_2} + \underline{C}_3(u_1),$$

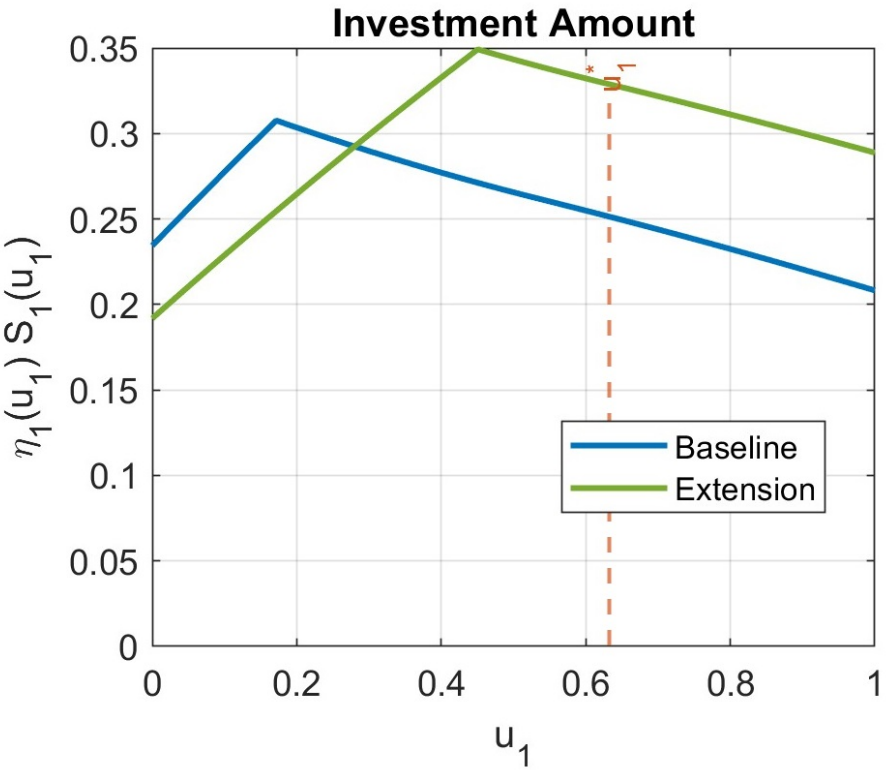
  where $\underline{I}_2(u_1) \equiv I_2(u_1, 0)$ and $\underline{C}_3 \equiv C_3(u_1, 0, 0)$.

- Can compensate agent with AUM in periods 1 and 2 ($\rightarrow$ fee in period 3) and fee in period 1.

- Key trade-off is modified.
  - Sticks: Can reduce not only $I_1(u_1)$ for $u_1 \in [u_1^*, \infty)$, but also $\underline{I}_2(u_1)$.
  - Carrots: Can increase not only $I_1(u_1)$ and $\underline{C}_3(u_1)$ for $u_1 \in [0, u_1^*)$, but also $\underline{I}_2(u_1)$.
  - $\rightarrow$ AUM changes in period 1 are smaller with longer-term contracts.
  - $\rightarrow$ Smaller price impact.

Equilibrium in Graphs





- With long-term contracts, price impact is smaller, especially for large u_1.
- Principal reduces AUM less than with short-term contracts following poor performance in period 1.

Conclusion

- Use dynamic financial contracting to study delegated investment and its asset-pricing implications.
  - Flow-performance relationship is optimal response by principals to agents' moral hazard.
- Main results:
  - Arbitrage is more effective for small shocks than for large shocks.
  - Arbitrage is more effective when contracts between principals and agents are longer-term.