

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Asymmetric Information, Repeated Trade, and Asset Prices

### Permalink

<https://escholarship.org/uc/item/3w99x5jq>

### Author

McLoughlin, James

### Publication Date

2012

Peer reviewed|Thesis/dissertation

**Asymmetric Information, Repeated Trade, and Asset Prices**

by

James McLoughlin

A dissertation submitted in partial satisfaction of the  
requirements for the degree of  
Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Christine A. Parlour, Chair  
Professor Robert M. Anderson  
Professor William Fuchs  
Professor Terrence Hendershott

Spring 2012

**Asymmetric Information, Repeated Trade, and Asset Prices**

Copyright 2012  
by  
James McLoughlin

## Abstract

Asymmetric Information, Repeated Trade, and Asset Prices

by

James McLoughlin

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Christine A. Parlour, Chair

Financial intermediaries play an important role in the pricing of financial assets. For example, intermediaries may act on behalf of consumers in deciding how their wealth is invested, or they may act as providers of liquidity. This dissertation explores several ways in which intermediaries impact price informativeness, the transaction costs investors incur, and investor welfare.

In the first chapter, I examine how prices reveal information when intermediaries are informed. Using a model of repeated trade between a long-lived, informed, price-discriminating market maker and risk averse traders with endogenous hedging demands, I first show that traders are weakly better off trading with an informed dealer, as they may learn something about an asset's value in the process of transacting. Second, while long-term incentives can induce an informed market maker to honestly reveal information and increase risk-sharing, they also enable the market maker to hide her information and extract more rents, reducing price informativeness. This less desirable outcome dominates with respect to both the parameter space and a selection criterion. Finally, measures of market quality, such as the transient component of price volatility (illiquidity), may not accurately reflect welfare.

The second chapter discusses how relationships affect prices when intermediaries are concerned about adverse selection. When counter-parties trade in OTC markets, such as those for corporate bonds or derivatives, the lack of anonymity implies that future terms of trade can influence prices today. Using a model of repeated trade between an informed trader and uninformed market makers, I show that information asymmetry can affect the markups charged by dealers in two ways. First, for a given market structure (number of market makers), traders with more private information incur lower trading costs because dealers offer better terms to mitigate adverse selection. Second, even when dealers can not compete directly on price quotes, they compete indirectly by improving the informed trader's outside option, though this competition is imperfect. While repeated trade allows two given counter-parties to ameliorate adverse selection, the maximum number of dealers, and hence the total gains achievable, are limited by information frictions. An empirical implication

is that the comparative statics of transaction costs only make sense conditional on market structure.

The third chapter considers the effect intermediaries have as financial advisors, and whether measures of their performance as mutual fund managers accurately reflect the value they add to an economy. Relative to the existing literature, I look at how the presence of mutual funds affects the price of the underlying asset in an economy. Once this pricing effect is accounted for, I show that standard measures of mutual fund performance may not accurately reflect whether fund management is welfare improving.

# Contents

<b>Contents</b>	<b>i</b>
<b>List of Figures</b>	<b>iii</b>
<b>1 Why trade with Goldman?</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Model . . . . .	5
1.3 The benefit of trading with an informed dealer . . . . .	8
1.4 The ambiguous effects of long-term incentives . . . . .	12
1.5 Welfare and price volatility . . . . .	18
1.6 Robustness and Extensions . . . . .	20
1.7 Conclusion . . . . .	24
<b>2 Transaction costs and asymmetric information in non-anonymous markets</b>	<b>26</b>
2.1 Introduction . . . . .	26
2.2 Model . . . . .	30
2.3 More informed traders pay lower transaction costs . . . . .	31
2.4 Competition indirectly reduces transaction costs . . . . .	35
2.5 Conclusion and Extensions . . . . .	38
<b>3 Measuring mutual fund performance with endogenous asset prices</b>	<b>40</b>
3.1 Introduction . . . . .	40
3.2 Model . . . . .	42
3.3 Equilibrium and Performance Measurement . . . . .	44
3.4 Active management does not necessarily add value . . . . .	46
3.5 Robustness . . . . .	48
3.6 Conclusion . . . . .	52
<b>A Chapter 1 Proofs</b>	<b>53</b>
<b>B Chapter 2 Proofs</b>	<b>64</b>
<b>C Chapter 3 Proofs</b>	<b>69</b>

**Bibliography**

# List of Figures

1.1	Extensive form of the static game . . . . .	10
1.2	Static game equilibria as a function of the parameters . . . . .	12
1.3	Dynamic game equilibria as a function of the parameters . . . . .	16
1.4	Pareto frontier of investor vs. dealer welfare . . . . .	18
1.5	Incentive compatibility with partial shares . . . . .	22
1.6	Dynamic game equilibria with partial shares . . . . .	24
3.1	Equilibrium demand, fees, risk premium, and performance measure . . . . .	46
3.2	Welfare effects of mutual funds . . . . .	48
3.3	Risk premia and profit for small versus big fund managers . . . . .	49
3.4	Welfare implications under monopolist fund management . . . . .	51



## Acknowledgments

I'd like to thank my chair Christine Parlour and my committee members Terry Hendershott, Willie Fuchs and Bob Anderson for their time and effort. Christine deserves special thanks for innumerable hours spent helping me organize my thoughts.

I am grateful to Dmitry Livdan, Adam Szeidl, Nicolae Garleanu, Richard Stanton, Marcus Opp, Brett Green, Johan Walden, Martin Lettau, Andres Donangelo, Bradyn Breon-Drish, Sara Holland, Adam Yonce, Barney Hartman-Glaser, Matteo Maggiori, Javed Ahmed, Isaac Hacamo, Thomas Reyes, Sebastien Betermier, June Wong, and Kim Guilfoyle for their help.

Thanks to my cohort comrades Hari Phatak, Boris Albul, and Sean Wilkoff for their friendship. Finally, thanks to my parents, brother Danny, and sister Maureen for their love and support.

# Chapter 1

## Why trade with Goldman?

### 1.1 Introduction

*The first thing you need to know about Goldman Sachs is that it's everywhere. The world's most powerful investment bank is a great vampire squid wrapped around the face of humanity, relentlessly jamming its blood funnel into anything that smells like money.*

–Matt Taibbi, “The Great American Bubble Machine”, Rolling Stone, 2009

*People lost money in it, but the security itself delivered the specific exposure that the client wanted to have.*

–Lloyd Blankfein, C.E.O. Goldman Sachs, 2010

In the financial markets literature, the strategic use of private information is largely viewed through the lens of equity markets. Much of our intuition rests on a paradigm in which competitive, uninformed market makers set prices against which a large number of agents, some of whom have private information, may trade. For example, a dealer quoting prices in a particular firm's stock is unlikely to have inside information about the firm's fundamentals, but may be exposed to adverse selection if the firm's employees or their associates do. However, a large number of asset classes, from traditional assets such as municipal bonds to more exotic asset-backed securities and derivatives, trade in over-the-counter (OTC) markets that do not fit this paradigm: price quotes are private, dealers have market power in setting prices, and it may be more reasonable to assume that they are the smart money.

If a market maker knows more about an asset's value than her clients, how do these clients avoid being taken advantage of when buying or selling the asset? Are they better off transacting with an uninformed dealer? Can longer term incentives, such as those promoted by recent financial regulation, improve client welfare by making the market maker more cooperative? To address these questions I posit a model of repeated trade between an informed market maker and uninformed risk-averse agents interested in hedging their

endowment shocks. The market maker is able to use her information strategically because she has pricing power—for simplicity I assume she is a monopolist—and may be more or less focused on short-term profits depending on how she discounts future cash flows.

I find that while long-term incentives can increase risk sharing gains, they are potentially detrimental to price informativeness. In a static setting, as long as consumers of a market maker's risk sharing services are sophisticated—i.e. they have rational expectations—they are not easily exploited and may benefit by learning from an informed dealer's price quotes. Increasing a market maker's concern for future business can improve welfare by promoting more trade and informative prices, but this is not the only possible outcome. In fact, long-term incentives are more likely to increase the dealer's trading gains by allowing her to conceal information.

There are three distinct aspects of my model. First, the price makers (dealers) are more informed about the risky asset's value than the price takers (traders). This information structure may apply for OTC assets because dealers often have a role in designing the securities that satisfy their clients' hedging needs, giving them a better sense of what these assets are worth. For example, following the financial crisis, politicians, regulators, and the media censured investment banks, particularly Goldman Sachs, for betting against their clients using superior information about the value of mortgage-backed securities.<sup>1</sup> Even when dealers don't have superior information about an asset's fundamental value, they may have private information about inventory imbalances that give them an edge: Cao, Evans, and Lyons (2006) show private information about order flow can predict future prices changes in the foreign exchange market.

Second, imperfect competition allows the market maker to strategically use her informational advantage when setting prices. For more exotic OTC assets, there may only be a few players who can accommodate a given hedging need, and, even in more staid asset classes, there is empirical support for imperfect competition—Green, Hollifield, and Schurhoff (2007) find that a significant component of transaction costs in the municipal bond market is due to dealer market power.

Third, the market maker is a long-lived player who trades off shorter versus longer term profits depending on her incentives. I use the market maker's discount factor as a reduced form proxy for the extent to which she cares about current versus future gains. Traders in OTC markets know the counter-party they are dealing with, and should consider a dealer's incentives when forming beliefs about what she is or isn't willing to do.<sup>2</sup>

The beliefs traders entertain about asset value, given an observed price quote, drive the model's results. An informed dealer, who posts these quotes, has long-term incentives that

---

<sup>1</sup>See Taibbi (2009) and Hobson (2010) for media coverage, the U.S. Senate Permanent Subcommittee's analysis of the financial crisis (2010) for political investigations, and Securities and Exchange Commission (2010) for the SEC's criminal complaint against Goldman Sachs.

<sup>2</sup>The role of incentives in financial markets has also received a lot of recent attention: following the crisis regulators in the U.S. and abroad have recommended a shift from short to long-term incentive-based compensation schemes. For details, see proposed implementations of the Dodd-Frank act's section 956 and the U.K. Financial Services Authority's "Principle 5" in The Financial Services Authority (2009).

change the trader's viable set of beliefs about asset value, and hence the possible trading outcomes, relative to the myopic case. In my model, a myopic risk-averse trader (the client) wants to hedge an endowment shock of the risky asset each period. For ease of exposition, I assume he wants to sell due to a positive shock. A long-lived, risk-neutral, monopolist market maker knows whether the mean asset payoff is high or low, and bids accordingly, potentially signaling its value.

The market maker's preferred bidding strategy depends on the magnitude of hedging demand. First, consider a myopic market maker who only cares about current period profits. When the trader is relatively desperate—for example, if he is highly risk-averse—the market maker can bid low regardless of what she knows about the asset's value and trade occurs; the trader knows there is a chance the asset's value is high, but is still willing to trade at a low bid. In concealing her private information via a pooling bid, the market maker maximizes the trader's uncertainty and the corresponding insurance rents. When hedging demand is less extreme, this pooling strategy is untenable: given his prior beliefs, the trader is unwilling to trade at a lowball bid, and the dealer cannot commit to posting a higher bid because she loses money when the asset value is low. Any high bid therefore reveals the asset value is high. Signaling the asset value via this separating strategy is the only way trade occurs, so the market maker is willing to reveal her information in order to get the client's business. Her profits are lower than those of the pooling strategy: by revealing information, she reduces trader uncertainty and the corresponding insurance rents. She also has to forgo trade in the low-value state because the trader never trusts a low bid.

The myopic case above demonstrates that when traders have rational expectations, they are not easily exploited. If anything, trading with an informed market maker allows them to learn about asset value, and they are therefore weakly better off trading with an informed rather than an uninformed monopolist. In the separating equilibrium, the market maker is engaging in a form of information sale à la Admati and Pfleiderer (1990): she reveals her information, alleviating some of the trader's uncertainty, in exchange for the residual risk sharing gains.

Long-term incentives yield two additional types of equilibria with higher gains from trade. The first is the type of Pareto improvement we usually associate with inter-temporal incentives: low bids are not trustworthy in the static separating equilibrium, but if future risk sharing gains are large enough, a market maker who cares about the long-term can post a credible low bid when the asset value is indeed low. If she deviates by lowballing when asset value is high, traders will no longer trust her, and she will lose out on these additional profits going forward. This separating outcome is Pareto improving because the trader is no worse off—he still learns about asset value and reduces his uncertainty—and the market maker extracts the additional gains associated with trade in the low value state<sup>3</sup>.

Long-term incentives also produce pooling equilibria that are not Pareto improving.

---

<sup>3</sup>This outcome is the easiest to support. Equilibria also exist where the trader receives some or all of the additional risk-sharing gains for a smaller set of parameters—these are also a Pareto improvement. For ease of exposition I focus on repeated game improvements that are most easily supportable.

When the market maker cares exclusively about current profits, high bids unambiguously reveal her information. However, if hiding her information leads to significant future profits, the market maker can credibly post a high bid that is uninformative. She is willing to incur small losses today in order to keep her information hidden in future periods, some of which will involve large gains on the spread between her pooling bid and the high asset value. Relative to the myopic separating outcome, gains from trade are again increased, but not in a Pareto improving way: because the market maker can now conceal her information, she captures additional risk sharing rents *and* recovers the gains associated with sharing her information in the separating strategy.

When we care about price informativeness in addition to total insurance gains, long-term incentives enable a socially undesirable pooling outcome. How likely is it that this equilibrium occurs relative to the preferable separating improvement? An increase in long-term incentives may be desirable if it enables a shift from the static/constrained separating equilibrium to a full trade separating equilibrium, but not the pooling equilibrium. It turns out this is never the case: generically, any parametrization supporting a separating improvement supports the pooling equilibrium, but the converse isn't true.

As usual, inter-temporal incentives support many possible equilibria in the repeated game. I use the notion of *undefeated* equilibria in Mailath, Okuno-Fujiwara, and Postlewaite (1993) to select the market maker's maximum profit outcome.<sup>4</sup> Intuitively, since the market maker moves first, if she posts a bid in line with her preferred pooling strategy, any separating equilibrium requires that the trader maintain "stubborn" beliefs: under his prior, he would trade, but instead makes the most pessimistic inference possible from the market maker's bid. These stubborn beliefs don't survive the *undefeated* criterion, and this strengthens the above result: long-term incentives only serve to make the market maker better off, at the trader's expense.

The media and the empirical microstructure literature often portray price volatility as a bad thing because it implies higher trading costs for investors.<sup>5</sup> Fixing parameters, I analyze how price volatility varies across the model's equilibria. Decomposing price volatility into its desirable and undesirable components—price discovery and transient illiquidity—I show the latter, a common measure of market quality, does not accurately reflect trader welfare. The static, partial-trade separating equilibrium endogenously constrains trade at extreme prices, reducing the price-impact of trades and leading to a lower illiquidity measure than the Pareto-dominant full insurance outcome. Huang and Wang (2010) also show that transient volatility doesn't accurately reflect welfare, but they reach this conclusion through an entirely different channel—information is symmetric in their model and participation costs drive the welfare implications. Finding the same result herein, in a different context, makes the point that illiquidity is not an indicator of welfare more robust.

The lack of perfect competition and the endogenous hedging demand in my model are

---

<sup>4</sup>See Gomes (2000) for an application of this selection criterion in the finance literature.

<sup>5</sup>Recent discussion, for example, has focused on whether algorithmic / high-frequency trading has increased price volatility. See, for example, Chaboud, Chiquoine, Hjalmarsson, and Vega (2009)

similar in spirit to Glosten (1989), which features a price-discriminating monopolist dealer, and the models of Biais, Martimort, and Rochet (2003), Bernhardt and Hughson (1997), and Dennert (1993), which feature oligopolistic dealers. All of these papers assume that the liquidity demand side is more informed, as is the case in most of the microstructure literature.

There is an earlier literature that considers the strategy of an informed market maker, or, equivalently, a “big” inside trader. Grinblatt and Ross (1985) and Gould and Verrecchia (1985) both analyze the linear pricing strategies of an informed dealer, but they assume she can commit to this strategy ex-ante, meaning she sometimes posts a price that implies expected losses. Laffont and Maskin (1990) use a static model like the myopic case herein to address this commitment problem, and show that prices can be informationally inefficient—pooling is a possible outcome. They restrict the insider to uniform pricing strategies, which constrains trade beyond the adverse selection dimension. I assume full price discrimination, which separates risk-sharing gains from learning effects, and is closer in spirit to how trade occurs in over-the-counter markets, where price schedules are not constant in quantity.<sup>6</sup> More recently, some models of limit-order markets, such as Goettler, Parlour, and Rajan (2009), do allow for informed trade originating from either the liquidity demand or supply side.

I discuss how a *dealer’s* long-term incentives affect prices and welfare in a financial markets context. Several papers show ways in which dealers can use a *trader’s* long-term concerns to their benefit. Seppi (1990) and Benveniste, Marcus, and Wilhelm (1992) look at whether monopolist dealers can screen out informed trade via the threat of reduced-form punishments, while Desgranges and Foucault (2005) explicitly model a dynamic game between a dealer and an informed trader. Bernhardt, Dvoracek, Hughson, and Werner (2005) use a reduced form model—prices are exogenous—along with supporting empirical evidence to show that dealers offer clients price improvements in exchange for their future order flow.

Finally, there is a significant corporate finance literature that highlights problems associated with managerial myopia (e.g. Stein (1988)), and suggests remedies that keep management focused on long-term shareholder value (e.g. Edmans (2009)). This paper provides an example of how myopia can improve, rather than harm, consumer welfare.

## 1.2 Model

Consider the following discrete-time infinite horizon game ( $t = 0, 1, 2, \dots$ ) in which a risk-neutral informed dealer quotes prices for an uninformed, strategic, risk-averse trader. In each period  $t$ , the asset produces an end of period payoff  $v_t$ :

$$v_t = v_{t-1} + \eta_t + \epsilon_t, \quad \mathbb{P}(\eta_t = \Delta) = \mathbb{P}(\eta_t = -\Delta) = \frac{1}{2}$$

---

<sup>6</sup>For example, Edwards, Harris, and Piwowar (2007) analyze how corporate bond trading costs vary with order size.

where  $\epsilon_t$  is a zero-mean residual risk term with variance  $\sigma_\epsilon^2$  and  $\eta_t$  is the fundamental component of the asset payoff. The dealer knows  $\eta_t$  before any trade occurs; traders only see its value after payoffs are realized.<sup>7</sup> I refer to the monopolist dealer as  $M$  and the trader as  $T$ .

Before trade occurs, the trader receives an endowment shock  $y_t \in \{-1, 1\}$  of the risky asset, where positive and negative shocks are equally likely. For simplicity, I assume the endowment shock is common knowledge, so that the trader's risk-sharing motive—whether he wants to buy or sell—is known before prices are set, so on a given trade the market maker only quotes one side of the market. The trader's end of period wealth  $w_t$  is

$$w_t = (y_t + q_t)v_t - q_tP(q_t)$$

where  $q_t \in \{-1, 0, 1\}$  is the amount of the risky asset purchased by the trader from the market maker at price  $P(q_t)$ . I assume traders choose portfolios myopically, ignoring intertemporal hedging demand, or equivalently that they are short-lived and the history of the game is publicly observable. When traders are indifferent, I assume trade occurs. Trader utility  $V_t^T$  is a combination of mean-variance preferences over end-of-period wealth and any benefit from learning about the asset payoff:

**Assumption 1.** Trader utility  $V_t^T$  is given by

$$V_t^T = \mathbb{E}w_t - \frac{\gamma}{2}\mathbf{Var}(w_t) + L(\mathbf{Var}(v_t) - \mathbf{Var}(v_t|\text{price}))$$

where  $L(0) = 0$  and  $L(\cdot)$  is strictly increasing.

The learning benefit  $L(\cdot)$  is a reduced form way of capturing any social gains from price efficiency. For example, more informative prices may allow investors to hedge exposures in correlated assets more efficiently, or may encourage more efficient real investment.<sup>8</sup>

The informed market maker  $M$  sets a price schedule  $P(q_t; \eta_t)$  that maximizes time- $t$  profits  $\pi_t = \mathbb{E}q_t(P(q_t; \eta_t) - v_t)$  and the present value of future profits discounted at rate  $\beta \in (0, 1)$ . Abstracting away from the time-value of money, the parameter  $\beta$  captures the extent to which the dealer's employees have an incentive to favor long-term firm value (e.g. via restricted stock/option grants or bonus clawbacks) over short term profits (e.g. those encouraged by annual cash bonuses). Denote her normalized utility at time  $t$  as

$$V_t^M = (1 - \beta)\mathbb{E}\sum_{s=t}^{\infty} \beta^{s-t}\pi_s = (1 - \beta)\mathbb{E}\pi_t + \beta\mathbb{E}V_{t+1}^M$$

Given that  $M$  has full market power in setting a price schedule, restricting trade to whole units of the asset does not sacrifice much generality in a qualitative sense, and greatly simplifies the specification of equilibrium beliefs; I discuss unrestricted quantities as an extension.<sup>9</sup>

<sup>7</sup>This is for simplicity to retain perfect public monitoring of the market maker's actions.

<sup>8</sup>Leland (1992) shows that insider trading and the corresponding information impounded in prices can lead to a social welfare increase via more efficient real investment.

<sup>9</sup>In states of the world where trade is not possible with whole units, trade is possible at a constrained quantity in the more general case. However, the main welfare conclusions of the paper are unchanged.

We can therefore think in terms of bid/ask prices for one unit of the asset:

$$B(\eta_t) = P(-1, \eta_t) \quad A(\eta_t) = P(1, \eta_t)$$

After the asset pays off, all information becomes public, and  $h^t = \{h^{t-1}, \eta_t, P_t, q_t, v_t\}$  denotes the history of realized play through the end of period  $t$ . This ex-post history includes the fundamental innovation  $\eta_t$ , so the game has perfect public monitoring. The trader has rational expectations: his expected utility and optimal portfolio choice are based on beliefs about the market maker's pricing strategy and any information revealed by a given price schedule realization. The history of prior play can affect his beliefs about the information revealed in prices. Let

$$\mu(P_t; h^{t-1})$$

denote the trader's belief that the fundamental is high ( $\eta_t = +\Delta$ ) given this period's price and the history of prior play. This belief will be specified for all possible price realizations, not just those on the equilibrium path.

In general, the pricing strategy  $P_t = P(\eta_t; h^{t-1})$  may involve randomization, in which case prices will partially reveal information about the fundamental. I restrict attention to pure pricing strategies that reveal either no information (pooling) or all information (separating).

**Definition 1.** A tuple  $\{P^*(\cdot, \cdot), \mu^*(\cdot; \cdot), q^*(\cdot)\}$  is a perfect Bayesian equilibrium of the repeated game if, for any time  $t$  and any ex-ante history  $h^{t-1}$ ,

- Taking the market maker's pricing rule  $P(\eta_t; h^{t-1})$  as given, the beliefs  $\mu^*(P_t; h^{t-1})$  about the fundamental realization  $\eta_t$  are formed using Bayes' rule for any price  $P_t$  that occurs with non-zero probability under the pricing rule.
- The portfolio choice  $q^*(P_t) \in \{-1, 0, 1\}$  maximizes the trader's expected utility over time  $t$  wealth given his endowment shock  $y_t$ , the observed price  $P_t$ , and corresponding belief  $\mu^*(P_t; h^{t-1})$ .
- The pricing rule  $P(\eta_t; h^{t-1})$  maximizes the informed market maker's expected utility given the evolution of trader beliefs  $\mu^*(\cdot; \cdot)$ , optimal demand  $q^*(\cdot)$ , and fundamental realization  $\eta_t$ .

Because information is perfectly revealed at the end of each period, I focus on simple trigger strategies that achieve efficient outcomes. As in most signaling games, beliefs are not restricted off of the equilibrium path without refinements, leading to multiple equilibria for some parameterizations, even in the static game. The *undefeated* selection criterion discussed in Mailath, Okuno-Fujiwara, and Postlewaite (1993) applies naturally to this game, and I use it as a refinement in both the static and dynamic cases.



### 1.3 The benefit of trading with an informed dealer

I first consider a myopic market maker and show that traders are weakly better off if she is informed. Because everything in the model is symmetric and endowment shocks  $y_t$  are common knowledge, I restrict attention to the bid-side of the market when describing results.<sup>10</sup> This side is relevant when  $y_t = 1$ : the trader has a positive endowment shock he wants to hedge by selling the asset, and the market maker need only supply a price on the bid-side of the market. To emphasize that this is the bid side of the quote, I will denote the price  $P_t$  offered for one unit as  $B_t$ . When discussing empirical implications, I reflect these results to the ask-side of the market. I refer to the states  $\eta = +\Delta$  and  $\eta = -\Delta$  as the high and low asset value states throughout, and to the market maker as the high or low type in these cases. Proofs are in the appendix.

When  $\beta = 0$ , the market maker does not care about future consequences and equilibria correspond to static game outcomes. As a benchmark, consider the symmetric information case at time  $t$ —when  $M$  does not know whether asset value is high or low—given prior asset value realization  $v_{t-1}$ :

**Lemma 1.** *In the static symmetric information game where neither party know  $\eta_t$ , full risk sharing occurs, the market maker captures all gains from trade, and prices are uninformative.*

The trader wants to sell to hedge her endowment shock. Maximum gains from trade are achieved via full risk sharing, and the market maker can capture all of these gains by posting a bid  $B^* = v_{t-1} - \frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2)$  that equals  $T$ 's no-trade certainty equivalent. Prices are obviously uninformative because neither party knows the fundamental value's realization.

When the market maker is informed about the fundamental component  $\eta_t$ , she is better off, ex-ante, if she posts the uninformative (pooling) bid of Lemma 1. However, she can not always commit to this bid, and a separating strategy that reveals her information is the only viable alternative. Specifically, she is not willing to bid above the asset's expected value  $\mathbb{E}v_t = v_{t-1} + \eta_t$ , and this constraint may bind in the low-value state. The trader's risk aversion parameter characterizes when these two outcomes—pooling or separating—are possible:

**Proposition 1.** *Risk aversion thresholds  $\gamma_S$  and  $\gamma_P$  characterize outcomes in the static game:*

1. *A separating equilibrium  $\mathcal{S}^0$  with partial trade—trade only occurs when asset value is high—exists for  $\gamma < \gamma_S$ .*
2. *A pooling equilibrium  $\mathcal{P}^0$  featuring full trade exists for  $\gamma \geq \gamma_P$ .*
3.  *$\gamma_P < \gamma_S$ , so both equilibria exist for  $\gamma \in [\gamma_P, \gamma_S)$*

---

<sup>10</sup>With unknown endowments, signaling and beliefs are more complicated because both sides of the quote are relevant strategic choices for the market maker.

4. *No mixed-strategy equilibria exist.*

When  $T$  is highly risk-averse, and correspondingly more desperate to trade,  $M$  can post a low bid regardless of  $\eta_t$  and capture all gains via full insurance. But when  $T$  is relatively risk tolerant, he is no longer willing to trade at such low bids unless they credibly reveal the asset value is low. The market maker generally prefers pooling bids that hide her information because they increase uncertainty and hence insurance rents she extracts. When this is not possible, a separating pricing strategy is her next best option.

Any high bid  $B_H > v_{t-1} - \Delta$  at which trade is feasible credibly signals that the asset value is high to the trader. Maximum gains are extracted when  $M$  offers the trader's outside option:

$$B_H^* = v_{t-1} + \Delta - \frac{\gamma}{2}\sigma_\epsilon^2$$

The intuition is the same as in the symmetric case, except everything is conditional on  $\eta_t = \Delta$ ; the expected payoff is higher, and the trader's uncertainty is reduced, lowering the risk premium  $M$  can charge.  $M$  will never post this bid when the asset value is low because she loses money on the trade. In the low-value state, any low bid  $B_L \leq v_{t-1} - \Delta$  needs to be incentive compatible:  $M$  must not be tempted to post a low bid when the asset value is actually high. This is only possible if  $B_L^*$ , the equilibrium bid offered when  $\eta_t = -\Delta$ , results in no trade. Hence risk-sharing is only possible in the high value state.

Figure 1.1 shows the main structure of the static game on the equilibrium path.<sup>11</sup> To support a separating equilibrium,  $T$  correctly believes the asset value is high when  $B_H^*$  is posted ( $\mu_H = 1$ ). He also has to believe that any low bid  $B_L$  at which trade could occur is not necessarily coming from the low-type—otherwise the high type will deviate to exploit this. Finally, bids that are low enough to prohibit trade, including the equilibrium bid  $B_L^* < v_{t-1} - \Delta - \frac{\gamma}{2}\sigma_\epsilon^2$ , are credible from the low type ( $\mu_L = 0$ ). Under appropriate beliefs, this partial trade equilibrium is possible as long as  $\gamma$  is below  $\gamma_S$ : when risk aversion is too extreme, the high bid  $B_H^*$  is below the low asset value, so the low type will want to mimic this bid, breaking the consistency of  $T$ 's beliefs.

The pooling equilibrium  $\mathcal{P}^0$  is comparable to the symmetric information case in Lemma 1: when the symmetric information bid is below the low asset value  $v_{t-1} - \Delta$ , the informed market maker can credibly bid this amount in all states of the world and capture the full ex-ante risk sharing gains. The trader knows he may hit this bid in a high value state, but his hedging demand is large enough that he still prefers to trade.

Once the symmetric information bid is above the low asset value because hedging demand is lower, which occurs when  $\gamma < \gamma_P$ , it is no longer rational for the informed market maker to maintain this bid in the low state. In that case, the market maker can not credibly bid low for the asset, so she foregoes gains in the low value state of the world while retaining them in the high value state via the separating strategy  $\mathcal{S}^0$ .

Note that any equilibrium has to involve trade in at least one state: under any beliefs the trader holds, a high-type market maker can, at a minimum, post a high bid that signals

<sup>11</sup>A full specification of off-equilibrium beliefs supporting Proposition 1 is given in the appendix.

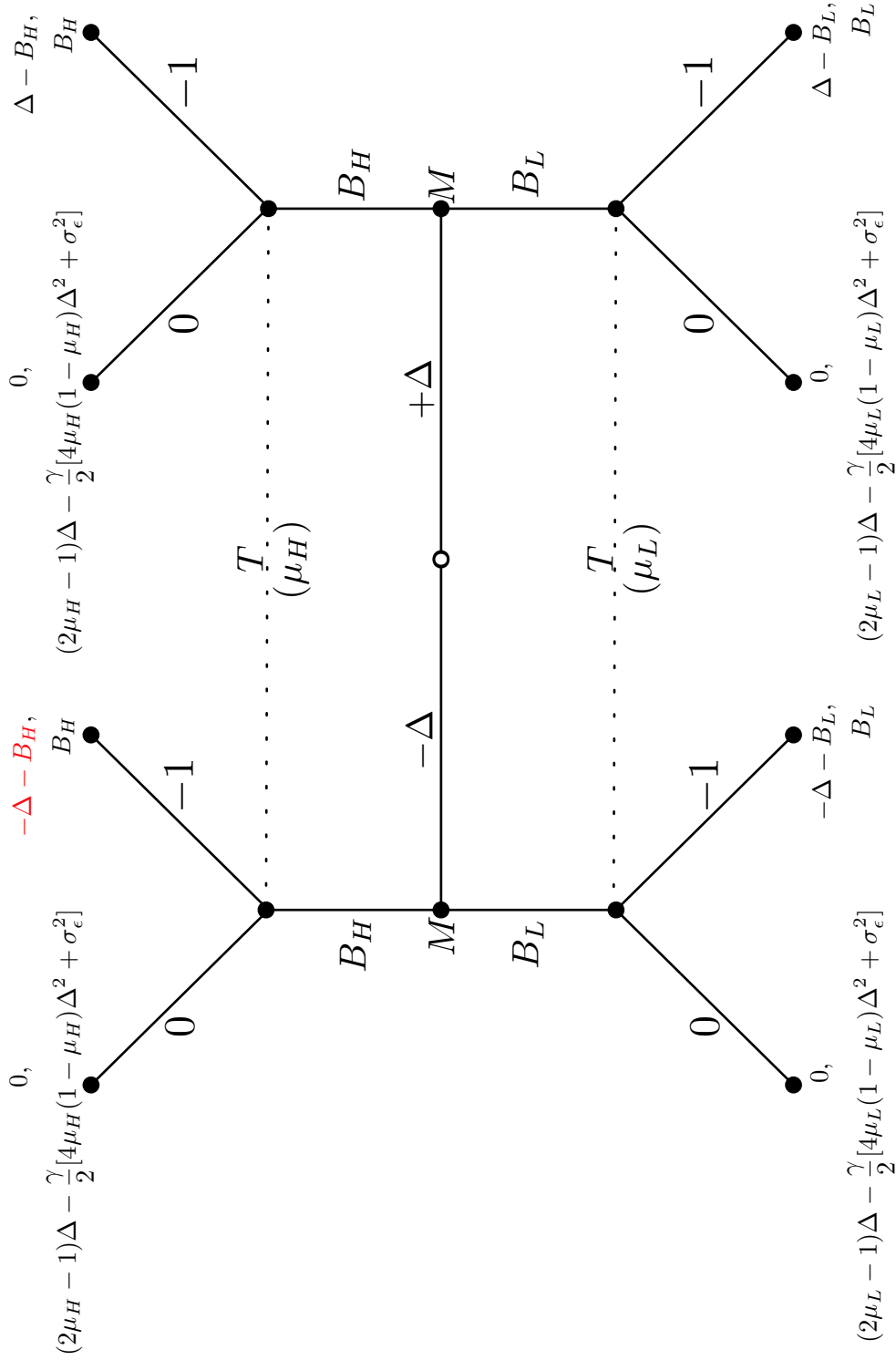


Figure 1.1: The static game: Nature chooses fundamentals; market maker  $M$  chooses either a high or low bid, where  $B_H > -\Delta$  and  $B_L \leq -\Delta$ ; trader  $T$  forms beliefs  $\mu_H$  and  $\mu_L$  about the probability the fundamental is high for each bid and chooses a quantity to trade, 0 or  $-1$ . **Red font** indicates a certain loss.

asset value and captures gains. Similarly, mixed strategy outcomes, which partially reveal information, are not possible. In these cases,  $M$  would occasionally signal the high asset value, but in any such equilibrium she always prefers posting a low-bid to exploit the trader's more generous beliefs.

The market maker's ex-ante profits are higher in the pooling equilibrium  $\mathcal{P}^0$  because uninformative prices increase uncertainty and allow her to extract more risk-sharing gains. There is not a Pareto dominating outcome, so I use the *undefeated* selection criterion:

**Lemma 2.** *When  $\gamma \in [\gamma_P, \gamma_S)$ , so that both equilibria  $\mathcal{P}^0$  and  $\mathcal{S}^0$  are possible outcomes,  $\mathcal{P}^0$  is the only undefeated equilibrium.*

When both equilibria are possible, separation only occurs if  $T$ 's beliefs are highly pessimistic. He needs to stubbornly believe that any low bid is posted by the high type. This equilibrium is *defeated* because there is another equilibrium (pooling) in which that price is posted by both types, and, further, it is the equilibrium they prefer to play. This selection criterion can also be supported with a single set of "robust" beliefs across all parameterizations:

**Corollary 1.** *In the static game, the belief specification*

$$\mu^*(B) = \begin{cases} 1, & \forall B_t > v_{t-1} - \Delta \\ \frac{1}{2}, & \forall B_t \in [\min_{\mu} \underline{u}(\mu), v_{t-1} - \Delta] \\ 0, & \text{otherwise} \end{cases}$$

*supports prices and quantities corresponding to the separating equilibrium  $\mathcal{S}^0$  for  $\gamma < \gamma_P$  and the pooling equilibrium  $\mathcal{P}^0$  for  $\gamma \geq \gamma_P$ .*

The separating threshold  $\gamma_S$  is increasing in  $\Delta$  because higher information asymmetry allows the high type dealer to signal without tempting the low type to mimic this bid. It is decreasing in  $\sigma_\epsilon^2$  because higher residual risk implies a lower bid  $B_H^*$ , which is only above  $v_{t-1} - \Delta$  if risk-aversion is lower. The static pooling threshold  $\gamma_P$  is monotonic in residual risk, but non-monotonic in fundamental risk:

**Remark 1.** The comparative statics of  $\gamma_P$  are

$$\frac{\partial \gamma_P}{\partial \sigma_\epsilon^2} < 0 \quad \frac{\partial \gamma_P}{\partial \Delta} \begin{cases} > 0 \text{ for } \Delta < \sigma_\epsilon \\ < 0 \text{ for } \Delta > \sigma_\epsilon \end{cases}$$

As  $\sigma_\epsilon^2$  increases, traders face more risk and are anxious to hedge, so the pooling threshold is lower. For the standard deviation of fundamentals  $\Delta$ , there are two opposing effects, illustrated in Figure 1.2. When  $\Delta$  is very small, the adverse selection problem is minimal and pooling bids are credible for most risk aversion parameters. As  $\Delta$  increases towards  $\sigma_\epsilon$ , the adverse selection effect dominates, pushing up the pooling threshold because only

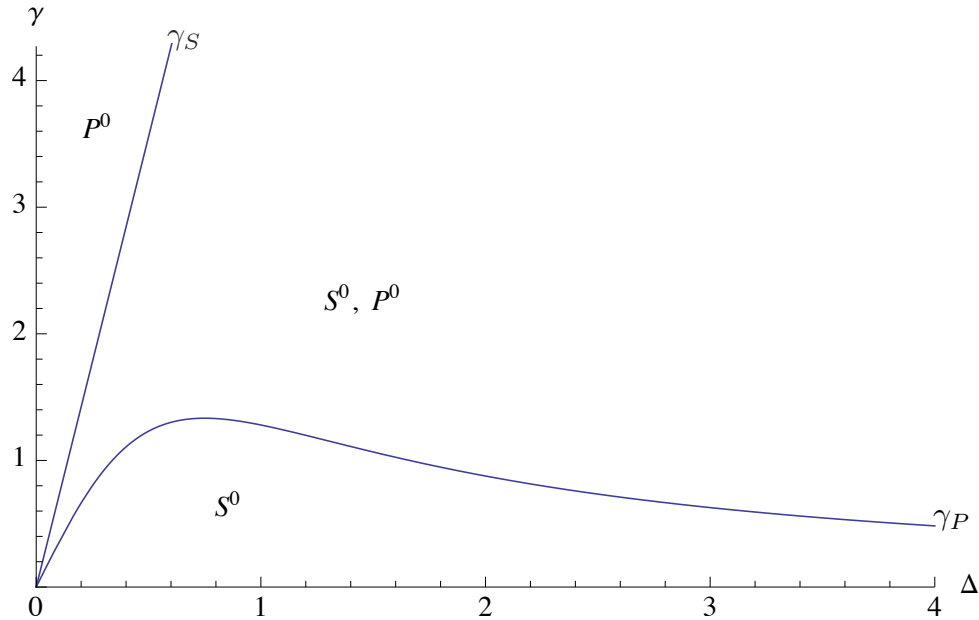


Figure 1.2: The pooling ( $\gamma_P$ ) and separating ( $\gamma_S$ ) thresholds as functions of the fundamental standard deviation  $\Delta$ . Residual risk  $\sigma_\epsilon = 0.75$ . The pooling equilibrium  $\mathcal{P}^0$  occurs above  $\gamma_P$  while the separating equilibrium  $\mathcal{S}^0$  occurs below  $\gamma_S$ . In the region where both are feasible,  $\mathcal{P}^0$  defeats  $\mathcal{S}^0$ .

desperate traders with high risk aversion are willing to trade at the uninformative pooling bid. Finally, for high values of  $\Delta$ , the trader's large fundamental risk exposure and the consequent hedging demand override the adverse selection effect, lowering the pooling threshold.

When the pooling equilibrium is a possible outcome, traders are neither better or worse off transacting with an informed monopolist—she extracts the same insurance rents as an uninformed market maker (Lemma 1). For moderate risk aversion  $\gamma < \gamma_P$ , traders are actually better off facing an informed monopolist, because the adverse selection friction works in their favor. The market maker prefers something to nothing, and is willing to reveal her information via aggressive bids in the high-value state to achieve trade. She still extracts any residual risk-sharing gains, but traders learn something in the process and reduce their uncertainty.

## 1.4 The ambiguous effects of long-term incentives

When the market maker is myopic, the preceding results show that consumers (traders) are better off in the separating equilibrium that occurs for moderate risk aversion—even though  $M$  extracts all residual insurance gains, the information in prices reduces trader uncertainty. When hedging demand is high,  $M$  is able to maintain uninformative prices and extract the

maximum possible rents. Do stronger inter-temporal incentives, represented by higher values of the discount rate  $\beta$ , improve trader welfare?

When hedging demand is high ( $\gamma \geq \gamma_P$ ), the pooling outcome  $\mathcal{P}^0$  is an equilibrium of the static game and therefore a sub-game perfect equilibrium of the repeated game. It is also the undefeated outcome of the repeated game because there is no credible threat that gives incentives any bite. I impose the following assumption to focus discussion on the additional gains achievable in a repeated context.

**Assumption 2.** Hedging demand is moderate, i.e.  $\gamma < \gamma_P$ , where  $\gamma_P$  is defined in Proposition 1.

The only static equilibrium under this restriction is separating and constrains the quantity traded in the low-value state to zero. This is also sub-game perfect in the repeated game and can be used to support more efficient equilibria.

Both equilibria I discuss have the same flavor: the market maker is able to credibly commit to a pricing strategy that achieves additional gains; traders believe this is the strategy being played until evidence arises to the contrary and make trades according to these beliefs; any deviation triggers reversion to the static separating equilibrium where these additional gains are lost. The static separating equilibrium  $\mathcal{S}^0$  is the harshest playable trigger and supports efficient equilibria in the repeated game.

There are two ways to increase gains from trade in the repeated context: an additional pooling equilibrium involving small losses for  $M$  on bids in the low-value state, and an additional separating equilibrium in which she can credibly signal the asset value in the low state without constraining the quantity traded.<sup>12</sup> Given a choice, the market maker prefers the pooling improvement while the trader prefers the separating improvement.

## Incentives may encourage more informative prices

The first dynamic result shows that a Pareto improvement is possible for some parameterizations because  $M$  can credibly signal the asset value is low. Traders believe high and low bids correspond to the true value of the asset. If  $M$  takes advantage of this by bidding low when the asset value is high,  $T$  no longer believes these low bids, and play reverts to the static equilibrium. Under Assumption 2, the static separating equilibrium  $\mathcal{S}^0$  is the market maker's outside option should she choose this deviation, implying her ex-ante continuation value  $V_{\mathcal{S}^0}^M$  is

$$V_{\mathcal{S}^0}^M = \frac{1}{2} \cdot 0 + \frac{1}{2} \frac{\gamma}{2} \sigma_\epsilon^2 = \frac{\gamma}{4} \sigma_\epsilon^2$$

---

<sup>12</sup>There are also a set of additional mixed strategy, partially revealing equilibria in between, but they are ignored for simplicity. They are also inferior to the pooling outcome from  $M$ 's perspective.

In a full-trade separating equilibrium, she earns the residual spread in both states of the world for a continuation value of  $\frac{\gamma}{2}\sigma_\epsilon^2$ , so her incentive constraint in the low state is

$$(1 - \beta)\frac{\gamma}{2}\sigma_\epsilon^2 + \beta\frac{\gamma}{2}\sigma_\epsilon^2 \geq (1 - \beta)(2\Delta + \frac{\gamma}{2}\sigma_\epsilon^2) + \beta V_{S^0}^M$$

The first term on the right-hand side is  $M$ 's most profitable deviation: by posting a low bid when the asset value is high, she makes the spread between high and low asset values ( $2\Delta$ ) plus the residual risk sharing gains. The following proposition summarizes the results of this heuristic argument.

**Proposition 2.** *If the market maker has significant long-term incentives ( $\beta > 2/3$ ), a full-trade separating equilibrium  $\mathcal{S}^\infty$  of the repeated game exists when trader risk aversion  $\gamma$  exceeds the threshold  $\gamma_S^\infty$ .  $\gamma_S^\infty$  is decreasing in the market maker's long-term incentives  $\beta$ , increasing in information asymmetry  $\Delta$ , and decreasing in residual risk  $\sigma_\epsilon^2$ .*

The market maker is able to credibly signal low-asset value through her bid because she would forgo significant future gains from trade by deceiving the trader. However, this is only supportable when the trader's risk-aversion, and therefore hedging demand, are relatively high. When risk aversion is low, the future gains are not high enough to keep the market maker honest. Similarly, there is a lower bound on how much  $M$  cares about future gains, parameterized by  $\beta$ .

The comparative statics of the supporting risk-aversion threshold  $\gamma_S^\infty$  are straightforward. It is decreasing in residual risk  $\sigma_\epsilon^2$  because it implies higher hedging demand for a given risk aversion level, and hence more gains that are lost by deviating; it is increasing in fundamental risk  $\Delta$  because this magnifies the benefit from deviating in the present period if the asset value is high; and it is decreasing in  $\beta$  because, as usual, more weight on future gains supports cooperation over a larger range of parameters.

The trader is no better off here, but he still has the same informational benefit—regardless of whether or not trade occurs in each state, both the repeated and static separating equilibria fully reveal the asset's value to him. The market maker achieves additional insurance rents in the low-value state, and for an outside observer, prices are more informative than the static separating case, where prices are unobservable in the low state.

## Incentives also enable information to be concealed

The static pooling equilibrium  $\mathcal{P}^0$  breaks down once the symmetric uninformed bid, which captures all ex-ante gains from trade, is above the low asset value and no longer credible for a low-type market maker. In that case, such a bid leads to an expected loss for the market maker when the asset value is low. In a repeated context, the market maker can credibly commit to incurring such losses in exchange for large gains in the high value state, again modulo a restriction on risk aversion so that future gains are large enough to support this strategy.

The market maker is tempted to abstain from bidding in the low value state, forgoing the expected losses in the present period at the expense of future profits. Her outside option is again the static separating continuation value  $V_{S^0}^M$ . By staying on the equilibrium path, she always earns the full gains  $\frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2)$ , so her IC constraint in the low state is

$$(1 - \beta)\left[\frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2) - \Delta\right] + \beta\frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2) \geq (1 - \beta) \cdot 0 + \beta V_{S^0}^M$$

where the first term is her loss on the pooling bid—it's negative under Assumption 2. The solution to this constraint characterizes the repeated pooling outcome:

**Proposition 3.** *A pooling equilibrium  $\mathcal{P}^\infty$  of the repeated game exists when trader risk aversion  $\gamma$  exceeds the threshold  $\gamma_P^\infty$ , where  $\gamma_P^\infty \leq \gamma_P$ .  $\gamma_P^\infty$  is decreasing in the market maker's long-term incentives  $\beta$ , and its other comparative statics are analogous to the static threshold  $\gamma_P$ .*

In this case, the market maker's long term incentives allow her to extract full gains for more parameterizations because the pooling threshold is lower than in the static case.  $M$ 's continuation utility allows her to credibly post a high bid in the low value state. Maximal insurance gains are achieved, but prices are uninformative and the trader is left with no share of the surplus. In this sense, pooling is an undesirable outcome, yet it is more easily supported when the market maker has long term incentives.

In sum, long-term incentives can have two effects: an informed market maker may be willing to honestly reveal asset value by signaling through prices, but she is also willing to incur losses in order to maintain uninformative prices. Is the uninformative pooling equilibrium  $\mathcal{P}^\infty$  a legitimate concern? The following proposition shows that, in terms of the parameter space, it is the more likely result of long-term incentives:

**Lemma 3.** *The pooling threshold supported by repeated play,  $\gamma_P^\infty$ , is always lower than the repeated separating threshold  $\gamma_S^\infty$ .  $\gamma_S^\infty$  is only below  $\gamma_P$  when the market maker has sufficient long-term incentives:*

$$\gamma_S^\infty < \gamma_P \iff \beta > \underline{\beta}$$

where  $\underline{\beta} \in (\frac{4}{5}, 1)$ .  $\underline{\beta}$  is increasing in information asymmetry  $\Delta$  and decreasing in residual risk  $\sigma_\epsilon^2$

The full-trade separating equilibrium  $\mathcal{S}^\infty$  is harder to support because  $M$ 's short-term gains from deviating in the low value state are significantly larger than the losses she incurs in the pooling equilibrium  $\mathcal{P}^\infty$ . The separating improvement is only possible for moderate risk aversion ( $\gamma < \gamma_P$ ), so inter-temporal incentives ( $\beta$ ) must be sufficiently high to support  $\mathcal{S}^\infty$ . When information asymmetry is high,  $M$ 's profitable deviation is more attractive, so she needs to care more about long-term gains in order to resist it. When residual risk is high, future gains from staying on the equilibrium path are higher, and cooperation can be supported for lower long-term incentives.



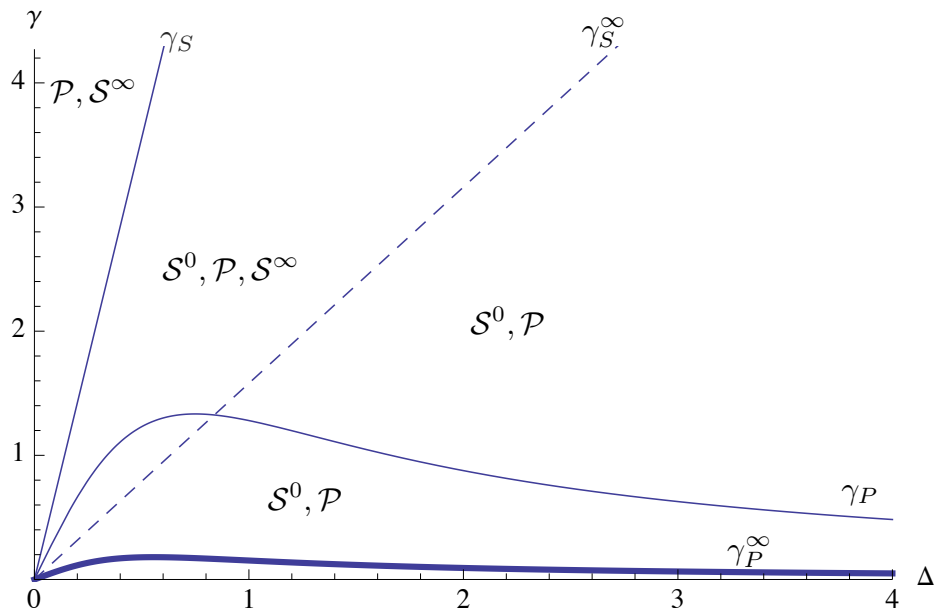


Figure 1.3: Static thresholds  $\gamma_P$  and  $\gamma_S$  (thin lines),  $\gamma_S^\infty$  (dashed line), and  $\gamma_P^\infty$  (thick line) as a function of the fundamental standard deviation  $\Delta$ , for residual risk  $\sigma_\epsilon = 0.75$  and discount factor  $\beta = 0.9$ .  $\mathcal{P}$  denotes a region where either of the pooling outcomes,  $\mathcal{P}^0$  or  $\mathcal{P}^\infty$ , are possible.

Figure 1.3 shows the regions of the parameter space in which each equilibrium outcome is possible. It is the repeated-game analog of Figure 1.2. When the market maker has long term incentives ( $\beta = 0.9$  in the figure), the repeated separating equilibrium  $\mathcal{S}^\infty$  featuring maximal gains, a fair surplus split, and fully informative prices can be supported for some moderate risk aversion levels  $\gamma \in [\gamma_S^\infty, \gamma_P)$ . However, whenever  $\mathcal{S}^\infty$  is possible, so is  $\mathcal{P}^\infty$ . Whether long-term incentives are desirable depends on how the equilibria discussed thus far compare to each other in a welfare sense.

### Welfare Summary

The pooling equilibrium occurs for high risk aversion regardless of long-term incentives using the *undefeated* criterion. For more moderate risk aversion (Assumption 2), the partial-trade separating equilibrium  $\mathcal{S}^0$  occurs absent long-term incentives. This equilibrium does not achieve full insurance, but does provide the trader with some gains via informative prices.

Do long-term incentives improve things in this parameter region? If risk aversion and incentives are high enough to support  $\mathcal{S}^\infty$ , and beliefs can be influenced such that this is the equilibrium played, then incentives can achieve a Pareto improvement—traders still learn

from prices, and more insurance occurs.<sup>13</sup> However, the feasibility of  $\mathcal{S}^\infty$  implies the pooling equilibrium  $\mathcal{P}^\infty$  is also possible (Lemma 3). The converse is not true—in this sense, pooling improvements dominate separating improvements with respect to the underlying parameter space.

Applying the *undefeated* selection criterion in the repeated game—that traders interpret dealer actions in the context of all possible equilibria they might be playing—long-term incentives hurt the trader because they only enable  $\mathcal{P}^\infty$ . A move from  $\mathcal{S}^0$  to  $\mathcal{P}^\infty$  does generate additional gains from trade, but it allocates these gains to the market maker while reducing price informativeness.

**Proposition 4.** *Total risk sharing gains are highest in the pooling equilibria  $(\mathcal{P}^0, \mathcal{P}^\infty)$ , while the trader surplus is highest in the separating equilibria  $\mathcal{S}^0$  and  $\mathcal{S}^\infty$ .*

Equilibrium	Dealer Surplus	Trader Surplus
$\mathcal{P}^0, \mathcal{P}^\infty$	$\frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2)$	0
$\mathcal{S}^0$	$\frac{\gamma}{4}\sigma_\epsilon^2$	$L(\Delta^2)$
$\mathcal{S}^\infty$	$\frac{\gamma}{2}\sigma_\epsilon^2$	$L(\Delta^2)$

There is a fundamental tradeoff between price efficiency and the amount of risk sharing that takes place. Because they both have the same welfare implications, let  $\mathcal{P}$  denote either of the pooling equilibria. Thus far, I've argued that these pooling equilibria are socially undesirable because they are bad for the consumer and result in uninformative prices. When we care about the trader's take or price informativeness, these equilibria are inferior to the full-trade separating equilibrium ( $\mathcal{S}^\infty \succ \mathcal{P}$ ). The repeated separating equilibrium is a Pareto improvement over the static separating outcome, so  $\mathcal{S}^\infty \succ \mathcal{S}^0$ .

I've used a monopolist market maker as a proxy for imperfect competition, but it is reasonable to argue that some of the additional gains generated in the pooling equilibrium would accrue to consumers under a more competitive market structure. To that end, consider a weighted social objective

$$V = V^T + \alpha V^M$$

where  $V^T$  and  $V^M$  are the ex-ante utilities of the trader and dealer in a given equilibrium, and  $\alpha \in (0, 1)$  denotes the level of competition in the dealer market. When  $\alpha = 0$ ,  $\mathcal{S}^0 \succ \mathcal{P}$ . But even when  $\alpha \in (0, 1)$ , there are parameterizations which result in the same social preference—as long as the market is not perfectly competitive, additional risk sharing gains enabled by the pooling equilibrium are potentially offset by the loss in price informativeness  $L(\Delta^2)$ . This occurs when information asymmetry ( $\Delta$ ), and hence the gain from informative prices, is large relative to residual risk  $\sigma_\epsilon$ . As an example, consider the simple Pareto frontier in Figure 1.4 (where  $L(x) = \frac{\gamma}{2}x$ ): in this case, if the introduction of long-term incentives induces a shift from  $\mathcal{S}^0$  to  $\mathcal{P}^\infty$ , there is a small increase in total social gains at the expense of a large loss in trader welfare.

<sup>13</sup>Depending on the beliefs, traders can also receive some of the additional insurance gains, but these outcomes are harder to support.

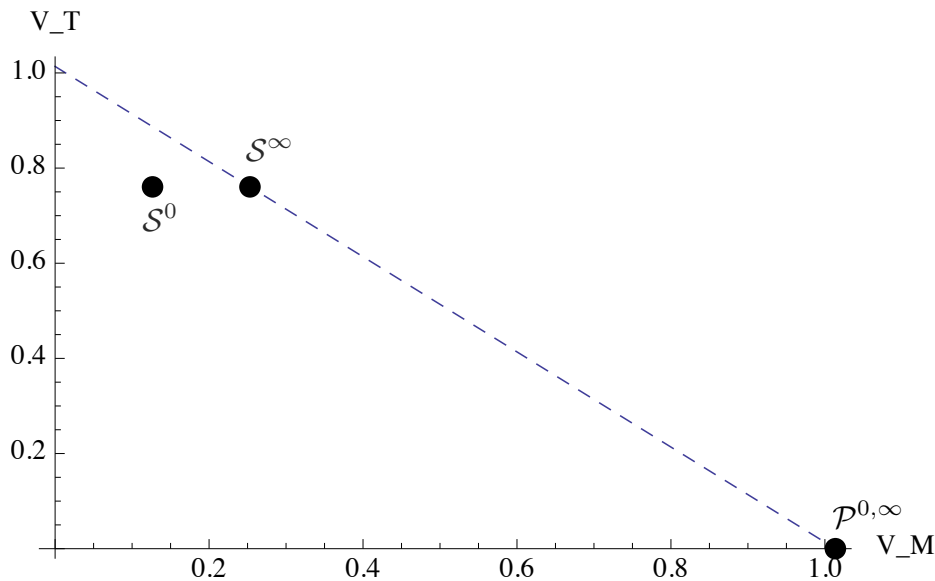


Figure 1.4: The Pareto frontier: ex-ante trader utility  $V^T$  (y-axis) vs. market maker utility  $V^M$  for model parameters  $\sigma_\epsilon = 0.75$ ,  $\gamma = 0.9$ ,  $\Delta = 1.3$  and learning benefit  $L(x) = \frac{\gamma}{2}x$

For a social planner instituting longer-term incentives, even if the selection criterion is relaxed, it's ambiguous which outcome will occur when both full-trade equilibria are possible. If the planner is not sure of the underlying parameters, it's more likely that incentives will support  $\mathcal{P}^\infty$  rather than  $\mathcal{S}^\infty$  because the former occurs for a strict superset of the parameter space. Hence it's possible that the introduction of these incentives only serves to make a less desirable pooling outcome possible—this is the case, for example, in the region of Figure 1.3 where  $\mathcal{S}^0$  and  $\mathcal{P}^\infty$  are the only viable outcomes.

The upshot, then, is that whether longer-term incentives make sense depends on a social tradeoff between risk sharing and price informativeness. When we care significantly about the latter, longer-term incentives can have negative welfare consequences, at least if competition is imperfect. In this case, it only makes sense to introduce them if some other policy will also influence beliefs and actions towards the Pareto improving outcome  $\mathcal{S}^\infty$ , or if some other transfer will restore the consumer's lost gains.

## 1.5 Welfare and price volatility

Given the welfare conclusions of the previous section, what can we expect to see in the data? Specifically, are standard measures of market quality aligned with welfare? For the sake of argument, fix the model parameters and suppose social preferences over equilibria are given by the example in Figure 1.4 ( $\mathcal{S}^\infty \succ \mathcal{S}^0 \succ \mathcal{P}$ )—the conclusions of this section are unchanged

if we reverse the last two.

Now consider price volatility as a measure of market quality: recent media attention has focused on the presumably negative effects of commodity speculation and high-frequency trading on volatility—Chaboud, Chiquoine, Hjalmarsson, and Vega (2009) test the latter claim. Volatility is not necessarily a bad thing if it is associated with price discovery, and I will therefore focus on the transient “illiquidity” component of price volatility. In Roll’s (1984) classic model of the bid-ask bounce, this measure corresponds to fixed trading costs or markups.

I have described the results so far from the bid side of the market, which occurs when the endowment shock is  $y_t = 1$  and the trader wants to sell. This is without loss of generality because I’ve assumed the market maker knows the desired trade direction—she only has to quote one side of the market. This is often the case in OTC markets, as dealers are not necessarily required to quote both sides. The results on the ask side of the market are therefore symmetric, and that implies a data generating process for observed prices in each of the equilibria discussed so far:

$$\begin{aligned}\mathcal{P} : P_t &= v_{t-1} - y_t \frac{\gamma}{2} (\Delta^2 + \sigma_\epsilon^2) \\ \mathcal{S}^\infty : P_t &= v_{t-1} + \eta_t - y_t \frac{\gamma}{2} \sigma_\epsilon^2 \\ \mathcal{S}^0 : P_{\tau_n} &= v_{\tau_{n-1}} + \eta_{\tau_n} - y_{\tau_n} \frac{\gamma}{2} \sigma_\epsilon^2\end{aligned}$$

where  $\{\tau_n\}_{n=1}^\infty$  is the sequence of realized trade times for  $\mathcal{S}^0$  because trade doesn’t occur in every state. These price processes can be used to decompose the variance of price changes in each equilibrium into permanent and transient components:

**Proposition 5.** *For each equilibrium, let  $\Sigma$  denote the total variance of price changes and  $\lambda$  the transient component. Then,*

$$\begin{aligned}\Sigma_{\mathcal{P}} &= (\Delta^2 + \sigma_\epsilon^2) + \underbrace{\frac{\gamma^2}{2} (\Delta^2 + \sigma_\epsilon^2)^2}_{\lambda_{\mathcal{P}}} \\ \Sigma_{\mathcal{S}^\infty} &= (\Delta^2 + \sigma_\epsilon^2) + \underbrace{\frac{\gamma^2}{2} \sigma_\epsilon^4}_{\lambda_{\mathcal{S}^\infty}} \\ \Sigma_{\mathcal{S}^0} &= 2(\Delta^2 + \sigma_\epsilon^2) + \underbrace{\frac{\gamma^2}{2} \sigma_\epsilon^4 - \Delta\gamma\sigma_\epsilon^2}_{\lambda_{\mathcal{S}^0}}\end{aligned}$$

and  $\lambda_{\mathcal{P}} > \lambda_{\mathcal{S}^\infty} > \lambda_{\mathcal{S}^0}$

The details are in the appendix, but, as in Roll (1984), the transient component  $\lambda$  corresponds to the spread between price and fundamental value, which reverts towards fundamentals. The ranking of  $\lambda$  across equilibria is not consistent with the posited social preferences.

Note that comparing illiquidity across the two full-trade equilibria,  $\mathcal{P}$  and  $\mathcal{S}^\infty$  matches welfare: the pooling equilibria  $\mathcal{P}$  is more illiquid because uninformative prices induce higher trader uncertainty, allowing the market maker to charge greater spreads. The confounding effect comes from the fact that  $\lambda_{\mathcal{S}^\infty} > \lambda_{\mathcal{S}^0}$ : the full-trade separating equilibrium is Pareto dominant, but displays higher illiquidity. The reason is that trade in  $\mathcal{S}^0$  is constrained to preclude extreme price impact: sells at the bid only happen when the asset value is high, while buys at the ask only happen when it is low. Adverse selection endogenously constrains trade and dampens the transient component of price volatility.

Huang and Wang (2010) find an analogous result—that lower transient volatility does not necessarily represent higher welfare—via an entirely different channel. In their model, prices are competitive, information is symmetric, and the friction is participation costs that restrict the supply and demand of liquidity. Here, prices are not competitive and adverse selection is the cause of inefficiency. The common theme in both cases is that an endogenous constraint on trade simultaneously reduces both welfare and the transient component of price volatility. The illiquidity measure  $\lambda$  defined above also corresponds directly to the measure of corporate bond illiquidity proposed by Bao, Pan, and Wang (2009).<sup>14</sup>

In sum, decomposition of volatility into good and bad components does not accurately reflect underlying welfare<sup>15</sup>. This suggests that such measures of market quality are not necessarily a good intermediate policy objective, at least in markets characterized by the features of this model, and they may not be good indicators of whether a change in incentives has had a positive or negative effect on welfare.

## 1.6 Robustness and Extensions

### Divisible shares

Thus far, I've assumed that trading the asset is an all or nothing proposition: the trader is unable to sell a fractional share of the asset.<sup>16</sup> This greatly simplifies the specification of equilibrium strategies and beliefs. In this section, I show that the main welfare implication of the paper—that inter-temporal incentives reduce trader welfare—is robust to this assumption.

Pooling outcomes  $\mathcal{P}^0$  and  $\mathcal{P}^\infty$  are the same when they are possible:  $M$  posts uninformative prices at which full insurance occurs. Similarly, the full-trade separating improvement  $\mathcal{S}^\infty$ , possible for high  $\beta$ , has full-trade in both states at the same prices. What changes is the static separating outcome  $\mathcal{S}^0$ : when the asset value is low,  $M$  can make an incentive compatible bid that results in trade. This is achieved by constraining the quantity traded

<sup>14</sup>Their measure  $\gamma$  is the negative of the first-order auto-covariance of price changes, which equals  $\lambda/2$

<sup>15</sup>The measure used here is strictly price based. A more nuanced decomposition of these components using both prices and order flow as in Hasbrouck (1993) does not substantially change the results: for a large portion of the parameter space, the ranking is unchanged.

<sup>16</sup>He will never be able to sell more than his share in equilibrium, as there are no gains beyond full insurance.

in the low value state such that a high type market maker prefers to trade the full quantity at her high bid rather than deviate by bidding low—she makes a larger spread in doing so, but the reduced quantity negates this gain.

**Proposition 6.** *The risk aversion threshold  $\hat{\gamma}_P$  characterizes outcomes in the static game with arbitrary quantities:*

1. *A full insurance pooling equilibrium  $\mathcal{P}^0$  exists for  $\gamma \geq \hat{\gamma}_P$ .*
2. *A separating equilibrium  $\mathcal{S}^0$  with partial trade always exists. Full residual insurance occurs in the high value ( $\eta_t = \Delta$ ) state, while trade in the low value state is constrained to a quantity  $Q_L^* \in (0, 1)$ .*

When quantities are unconstrained,  $M$  can fully price discriminate, so it is without loss of generality to consider her action space as offering a price-quantity pair  $(B, Q)$ , where  $Q$  is the quantity she is willing to buy. Trader  $T$  then chooses an optimal sale quantity  $q \in \{0, Q\}$ —let  $q$  be positive when he is selling. Beliefs  $\mu(B, Q; h_{t-1})$  that the asset value is high are formed over both price and quantity dimensions, and the equilibrium concept is identical to Definition 4, other than the change from an optimal pricing strategy  $B^*(\eta_t; h_{t-1})$  to a price-quantity strategy  $(B^*, Q^*)(\eta_t; h_{t-1})$ .

Any high bid  $B_H > v_{t-1} - \Delta$  credibly signals that the asset value is high ( $M$  loses money if she posts this in the low state), and maximum gains are extracted when  $M$  offers the trader's outside option:

$$(B_H^*, Q_H^*) = (v_{t-1} + \Delta - \frac{\gamma}{2}\sigma_\epsilon^2, 1)$$

In the low-value state, any bid  $B_L \leq v_{t-1} - \Delta$  needs to be incentive compatible:  $M$  must not be tempted to use this bid when the asset value is actually high. This is possible if  $Q_L^*$ , the equilibrium quantity offered when  $\eta_t = -\Delta$ , is constrained.

First ignore incentive compatibility: given any quantity  $q \in [0, 1]$ , the low bid price  $B_L(q)$  needs to meet the trader's participation constraint ( $IR_q$ ) by offering his autarky value, implying:

$$B_L(q) \geq v_{t-1} - \Delta - (2 - q)\frac{\gamma}{2}\sigma_\epsilon^2 \quad (IR_q)$$

This price will be incentive compatible for small  $q$ , as low bid profits  $\pi_L(q)$  are lower than the full insurance profits  $\pi_H^* = \frac{\gamma}{2}\sigma_\epsilon^2$  at  $(B_H^*, Q_H^*)$ . However, as  $q$  increases, the market maker is tempted to deviate when the asset value is high and capture the much larger spread provided by  $B_L$ , requiring the incentive constraint  $\pi_H^* \geq \pi_L(q)$  when asset value is high. This constraint implies a lower bound on  $B_L$  for a given quantity

$$B_L(q) \geq v_{t-1} + \Delta - \frac{\gamma}{2q}\sigma_\epsilon^2 \quad (IC_q)$$

Figure 1.5 shows the relationship between these constraints: there is a  $q^* \in (0, 1)$  such that for quantities higher than  $q^*$ , the  $IC$  constraint binds. Trader  $T$  believes that any

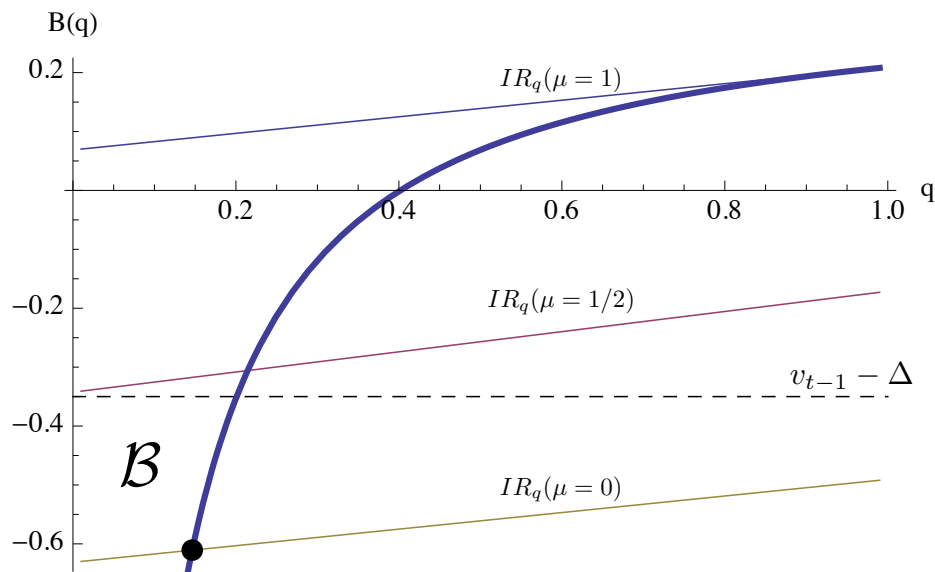


Figure 1.5: The static separating bid with partial shares. The thin slanted lines are the participation constraints  $IR_q(\mu)$  for different trader beliefs  $\mu$ . The horizontal dashed line denotes the low asset value  $v_{t-1} - \Delta$ —bids below this are profitable in the low state. The dark, curved line is the  $IC_q$  constraint.  $\mathcal{B}$  denotes the region of incentive compatible, rational low-state bids. The intersection of  $IR_q(\mu = 0)$  and  $IC_q$  is the optimal bid-quantity pair in the low state. Parameters:  $\gamma = 0.5, \sigma_\epsilon = 0.75, \Delta = 0.35$ .

bid-quantity pair above the  $IC$  constraint credibly signals the asset value is low (region  $\mathcal{B}$ ). As  $q$  increases from 0,  $M$  only has to ensure participation, and profits are increasing. Once the incentive constraint binds at  $q^*$ , profits decrease. Hence she will select this quantity in the separating equilibrium, and bid such that the trader participates.

The modification of  $\mathcal{S}^0$  when partial shares can be bought/sold does not qualitatively change the results regarding long-term incentives. The additional constrained trade in the low value state improves  $M$ 's outside option, but this options supports both the separating ( $\gamma_S^\infty$ ) and pooling ( $\gamma_P^\infty$ ) equilibria. It is still the case that the pooling equilibrium  $\mathcal{P}^\infty$  dominates  $\mathcal{S}^\infty$  because  $\gamma_S^\infty > \gamma_P^\infty$ . The welfare conclusions that follow are unchanged.

### Asymmetric fundamental shocks

To derive the main results of the paper, I use a model with symmetric, equally likely shocks to the privately known component of asset value  $\eta_t$ . I now relax this assumption and show the results of the paper are largely unchanged. Suppose asset value follows a process with asymmetric shocks

$$v_t = v_{t-1} + \eta_t + \epsilon_t, \quad \mathbb{P}(\eta_t = \Delta_H) = p, \mathbb{P}(\eta_t = \Delta_L) = 1 - p$$

where  $\Delta_H > \Delta_L$  and  $p \in (0, 1)$ . Under these dynamics, it is no longer without loss of generality to consider one side of the market exclusively. There is still symmetry in the sense that ask-side results when  $\mathbb{P}(\eta_t = \Delta_H) = p$  correspond to bid-side results where  $\mathbb{P}(\eta_t = \Delta_H) = 1 - p$ . This implies different pooling thresholds for each side of the market:

**Proposition 7.** *Risk aversion thresholds  $\gamma_S, \gamma_{P,A}, \gamma_{P,B}$  characterize outcomes in the static game with asymmetric fundamental shocks:*

1. A separating equilibrium  $\mathcal{S}^0$  with partial insurance—sells only occur when asset value is high, buys when asset value is low—exists for  $\gamma < \gamma_S$ . This threshold is independent of whether the trader is buying or selling.
2. A pooling equilibrium  $\mathcal{P}_B^0$  featuring full insurance on the bid-side of the market ( $y_t = 1$ ) exists for  $\gamma \geq \gamma_{P,B}$ .
3. A pooling equilibrium  $\mathcal{P}_A^0$  featuring full insurance on the ask-side of the market ( $y_t = -1$ ) exists for  $\gamma \geq \gamma_{P,A}$ .
4. A pooling equilibrium  $\mathcal{P}^0$  featuring full insurance on the both sides of the market exists for  $\gamma \geq \max\{\gamma_{P,A}, \gamma_{P,B}\}$ .
5.  $\gamma_S$  is greater than both pooling thresholds, so pooling and separating are possible for  $\gamma \in [\gamma_{P,A}, \gamma_S)$  on the ask-side and  $\gamma \in [\gamma_{P,B}, \gamma_S)$  on the bid-side.
6.  $\gamma_{P,A} < \gamma_{P,B} \iff p > \frac{1}{2}$ .
7. No mixed-strategy equilibria exist.

The effects of asymmetric fundamental shocks in the myopic case are illustrated in Figure 1.6. The results largely correspond to the symmetric case (Figure 1.2), but the spread  $\Delta_H - \Delta_L$  determines the volatility of fundamentals and hence the degree of information asymmetry. The previous comparative statics, where this spread replaces the parameter  $\Delta$ , are unchanged. The bid threshold  $\gamma_{P,B}$  is increasing in  $p$ , while the ask threshold  $\gamma_{P,A}$  is decreasing in  $p$ —as the trader is more sure the asset value is high, it is harder to support a low pooling bid.

The major difference is the asymmetry between when pooling is supported on each side of the market. In the figure, the region between the two pooling thresholds supports pooling at the ask, but not at the bid (for  $p < 1/2$ , the results are reversed). This region is largest when  $p$  takes on extreme values. The ask-side pooling outcome  $\mathcal{P}_A^0$  defeats  $\mathcal{S}^0$  because  $M$  prefers pooling where possible, so it will be the equilibrium in that region under my selection criteria, and the market maker has an improved outside option in the repeated game. I have not yet analyzed how this affects the feasibility of improvements in the repeated game. However, when  $\gamma < \min\{\gamma_{P,B}, \gamma_{P,A}\}$ , separating prices are still the only possible outcome on both sides of the market, so  $M$ 's outside option is the same as in the symmetric case. In this region, the results are qualitatively the same as the symmetric information case:



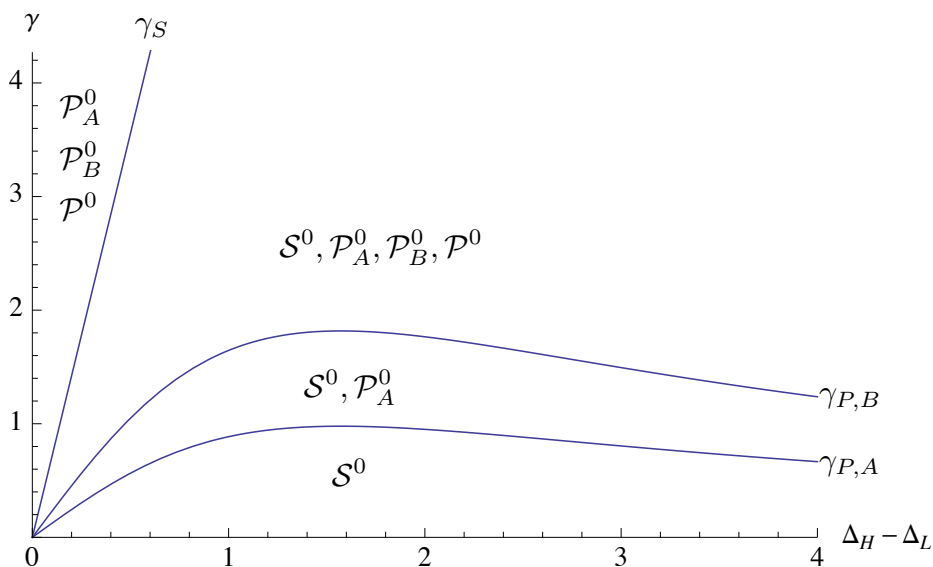


Figure 1.6: The pooling ( $\gamma_{P,A}, \gamma_{P,B}$ ) and separating ( $\gamma_S$ ) thresholds as functions of the private information spread  $\Delta_H - \Delta_L$ . Residual risk  $\sigma_\epsilon = 0.75$  and high-state probability  $p = 0.65$ . Pooling on both sides of the market is possible for  $\gamma > \gamma_{P,B}$ , but only on the ask-side for  $\gamma \in [\gamma_{P,A}, \gamma_{P,B})$ .  $\mathcal{P}^0$  is undefeated among all equilibria, while  $\mathcal{P}_A^0$  and  $\mathcal{P}_B^0$  defeat  $\mathcal{S}^0$ .

**Proposition 8.** *In the repeated game, for  $\gamma < \min\{\gamma_{P,B}, \gamma_{P,A}\}$ , there exists a risk aversion threshold  $\gamma_S^\infty$  supporting a separating improvements on both the bid and ask sides of the market. Thresholds  $\gamma_{P,B}^\infty$  and  $\gamma_{P,A}^\infty$  support pooling improvements on the bid and ask sides, respectively, where*

$$\gamma_{P,B}^\infty < \gamma_S^\infty \quad \gamma_{P,A}^\infty < \gamma_S^\infty$$

This confirms that, if the static separating equilibrium  $\mathcal{S}^0$  is the market maker’s outside option, whenever long-term incentives enable a separating improvement, they also enable a pooling improvement. While the separating outcome is a Pareto improvement, the pooling outcome reduces trader welfare. Hence the qualitative conclusions about the effects of long-term incentives on trader welfare remain unchanged.

## 1.7 Conclusion

Using a stylized model of repeated trade, I’ve shown that traders interested in hedging a risk are weakly better off trading with an informed dealer. If they’ve got rational expectations, the dealer can’t use the information to her advantage, and may actually share what she knows to facilitate trade. Surprisingly, when the dealer has long-term incentives—she values future

profits highly—traders may be worse off because the dealer is willing to incur short-term losses to conceal her private information.

Using simple assumptions on the dynamics of asset value, I decompose price volatility to show that liquidity measures do not necessarily reflect trader welfare (or price efficiency). Enriching the fundamental dynamics will allow for more precise statements about which types of assets are more likely to benefit or suffer from long term incentives in this context (for example those with bond-like versus stock like payoffs), as well as sharper empirical implications. For example, fundamentals could more closely track the dividends of a consumption tree, where the market maker knows whether consumption growth will be high or low. I also made a strong perfect monitoring assumption; relaxing it will not necessarily change the welfare results, but it will imply other interesting price dynamics: when traders experience significant losses, temporary periods of lower market liquidity may follow if they infer the market maker has deviated from a given pricing strategy.

## Chapter 2

# Transaction costs and asymmetric information in non-anonymous markets

### 2.1 Introduction

When counter-parties trade in OTC markets, such as those for corporate bonds or derivatives, the lack of anonymity implies that the consideration of future transactions can influence prices today. In the presence of adverse selection, informed traders may be willing to forgo information rents today in exchange for future liquidity rents, resulting in gains from trade that would be lost in a static or anonymous setting. This tradeoff, as well as the degree of competition in a given market, jointly determine the transaction costs—a result of dealer markups—that traders incur.

Using a model of repeated trade between an informed trader and uninformed market makers, I show that repeated interaction affects transaction costs through two channels. First, for a given market structure (number of dealers) more informed traders pay lower costs. Second, even if dealers can not compete directly on price, a trader's costs are still lowered by the presence of competition because it improves their outside option. This indirect effect of competition is imperfect: a Bertrand outcome does not obtain, so costs will vary with the number of dealers. The main implication of the model, then, is that measurement and comparison of transaction costs in markets characterized by non-anonymous trade should proceed with caution, at least to the extent that information frictions apply, by conditioning on market structure.

The model used is stylized in the sense that it analyzes the effect of a specific friction – information asymmetry – on a specific component of trading costs – the markups charged by dealers due to imperfect competition. The model does not include inventory effects or fixed order processing costs, so the spread charged by market makers consists of an adverse selection component plus any markups they are able to charge. As in most OTC markets, the

notion of a binding bid-ask spread does not apply, but these two components can be identified by the econometrician from the history of price changes: adverse selection costs correspond to the “permanent” component of price changes, markups to the “transitory” component, which tends to revert following a trade. In this sense, the primitive parameters of the model induce an endogenous bid-ask bounce. Finally the information friction is assumed to be extreme enough such that trade is only viable in a repeated setting. When adverse selection is not extreme, a competitive, anonymous market as in Glosten and Milgrom (1985) is viable and transaction costs (the transitory component) are zero – this will serve as a benchmark.

Both of the model’s effects are driven by the incentive compatibility constraints of an informed trader facing a market maker’s price schedule. In each period, traders come to the market driven by either an information or liquidity motive. In the informed state, the trader knows the asset’s end-of-period common value, while in the liquidity state she has a private value that differs from current fundamentals. This private value is not extreme – the trader is somewhat elastic to price – and the dealer’s break-even price, given he doesn’t know the trading motive, is above this valuation. In a one-shot anonymous setting, the market breaks down. With the potential for repeated interaction, the dealer can incentivize the trader to cooperate by revealing her trading motive, and gains from trade are realized.

First consider the case of a monopolist dealer. He is able to solicit the trader’s motive by offering a share of future gains from trade in the liquidity state. There are two dimensions of information asymmetry: the likelihood of informed versus liquidity trade, and the magnitude of common value shocks about which the trader has private information. When information asymmetry is high in either dimension, all else equal, the trader values future liquidity gains less, so her share must be increased to maintain incentive compatibility. The dealer wants to impose markups to the point where incentive compatibility on the part of the trader binds, so prices are determined by this constraint. These markups are a function of the primitive parameters, not realized fundamentals, so they show up as the transitory component of price changes. Measurements of this component via Roll’s 1984 covariance estimate or trade indicator regressions along the lines of Madhavan, Richardson, and Roomans (1997) and Huang and Stoll (1997) will produce the same conclusion: for a fixed market structure (e.g. one dealer), transaction costs are decreasing in the trader’s informational advantage.

The second result of the paper concerns the entry of additional competition on the liquidity supply side. In the monopolist model described above, the dealer ensures cooperation by: a) promising a share of future gains from trade; b) credibly threatening to terminate the relationship if a trader deviates by lying about her motive, destroying any future gains. If there are competing dealers present in the market, the trader’s outside option is significantly improved – if she deviates and burns an existing relationship, she can trade with the remaining dealers. While the OTC nature of the market does not allow dealers to compete directly on price within a given trading period, competition increases the trader’s outside option, and therefore the surplus share dealers must offer to maintain cooperation. Hence the level of competition in a given market can be a confounding factor in determining the causes of transaction costs, and it should be controlled for when attributing costs to other factors.

The effect of competition on transaction costs suggests a simple explanation for some of the observed cost variation in, for example, corporate bond markets. Bao, Pan, and Wang (2009) show that bonds with lower ratings have higher transaction costs, and suggest that this illiquidity might be a priced risk factor. However, to the extent that my model reflects the OTC nature of the corporate bond market, it may just be that there is less competition in lower rated bonds: according to MarketAxess, an electronic bond trading platform, there are almost twice as many dealers in the U.S. high-grade corporate bond market as there are in the high-yield market.<sup>1</sup>

While I mainly considered the effect of exogenous changes in dealer competition, each additional dealer who enters reduces the liquidity supply side's surplus share, placing an endogenous upper bound on the feasible level of competition in a given market. The fact that competition is endogenously limited due to incentive effects, even with free-entry, is also unique to the repeated/non-anonymous nature of the market examined in this paper. In many microstructure models, Bertrand competition is invoked to justify setting prices such that dealers break even in the presence of two or more competitors. In models that allow informed traders to split their orders among competing dealers, this result does not hold for a given level of competition, but free entry still implies that an infinite number of dealers will enter and profits will go to zero. In the present case, even with free entry, there is a limit on the per-capita number of dealers—the number of dealers per trader. Dealers will enter until profits are zero, but the limited number of dealers prevents maximum social gains from being achieved—some traders are not able to find a market maker to transact with.

To motivate the model empirically, I appeal to stylized facts in the markets for corporate and municipal bonds because they trade in an OTC/non-anonymous fashion and there is a wealth of existing research on their transaction costs.<sup>2</sup> Both Bao, Pan, and Wang (2009) and Edwards, Harris, and Piwowar (2007) document that bonds with lower ratings or higher yields are more costly to trade.<sup>3</sup> If markets for those bonds are less competitive, this is in line with the model's predictions. Second, in the municipal bond market, Green, Hollifield, and Schurhoff (2007) conclude that a significant portion of transaction costs are due to dealer markups (not other "marginal" cost components) via a structural estimation, and it's natural to assume the same applies for corporate bonds. Finally, to address concerns about whether the model's main friction, which is informational, is relevant in the corporate bond market, I appeal to Hotchkiss and Ronen (2002), who show that bonds incorporate informational events into prices as quickly as equity markets.

---

<sup>1</sup>See <http://www.marketaxess.com>: they advertise 67 dealers in the high-grade credit market compared to 37 dealers in the high-yield market as of April 2012.

<sup>2</sup>Recent studies of corporate bond transaction costs include Edwards, Harris, and Piwowar (2007), Bao, Pan, and Wang (2009), Bessembinder, Maxwell, and Venkataraman (2006), and Goldstein, Hotchkiss, and Sirri (2007). In the municipal bond markets, see Green, Hollifield, and Schurhoff (2007) and Green, Li, and Schurhoff (2010)

<sup>3</sup>In Bao, Pan, and Wang (2009) costs are measured via transitory price changes, whereas Edwards, Harris, and Piwowar (2007) look at the total average spread

## Literature Review

Theoretical motivation for relationship/reputation effects in financial markets begins with Seppi (1990) and Benveniste, Marcus, and Wilhelm (1992), who look at whether monopolist dealers can screen out informed trade via the threat of reduced form punishments. Desgranges and Foucault (2005) are the first to explicitly model the dynamic game between a dealer and an informed trader – they construct a stationary equilibrium in which the dealer offers price improvements as a function of past profits. In all of the above papers the dealer is a monopolist and the focus is on whether and how cooperation can be maintained – my focus is more on the pricing effects of relationship trade and imperfect competition. Empirical support for reputation effects is given by Battalio, Ellul, and Jennings (2007), who use specialist relocations on the NYSE as a natural experiment to show that brokers who move with the specialist trade on better terms.

The second result in this paper adds competing dealers and looks at their effect on prices. Static adverse selection models often assume dealers earn zero profits to pin down prices, but there is a significant literature exploring imperfect competition: Glosten (1989) examines how a monopolist sets his price schedule against a trader who has mixed information and liquidity motives, and notes that monopolistic pricing is socially preferable when adverse selection is significant. Dennert (1993) and Bernhardt and Hughson (1997) consider stylized models in which a finite number of dealers compete for order flow and traders are able to split orders amongst dealers. The result is that with any finite number of market makers, expected profits are positive (Bertrand competition doesn't hold) and transfer schedules are convex (agents are charged more for the second unit than the first). This result is generalized by Biais, Martimort, and Rochet (2003), who also show welfare implications of these schedules: oligopolistic competition increases trading volume (and hence efficiency) over the monopolist case, but falls short of ex-ante social efficiency (the allocation an informed social planner would choose). In these models, profits go to zero as the number of dealers increases to infinity, which will occur with free entry. The second result of this paper shows that adverse selection and relationship effects combine to limit the viable level of competition for a given parameterization, and hence market structure provides a link between prices and parameters.

Though the focus of this paper is on an implication of adverse selection in a relationship market, Carlin, Lobo, and Viswanathan (2007) show that cooperation among traders in a market can be sustained for non-informational reasons – they abstain from “preying” on each other in a repeated setting. Bernhardt, Dvoracek, Hughson, and Werner (2005) show that traders can obtain price improvements from a dealer's posted quotes with the promise of future order flow. Both of these models take prices as exogenous, while I solve for prices as a function of the model's parameters.

There is a significant literature concerning over-the-counter trade. The search models of Duffie, Garleanu, and Pedersen (2005) and Duffie, Garleanu, and Pedersen (2007) are driven by hedging motives, while more recent work including Duffie, Giroux, and Manso (2010) and Golosov, Lorenzoni, and Tsyvinski (2009) study information driven trade. In my model, repeated interaction enables the adverse selection problem to be negotiated around, whereas

trade in the above models is entirely anonymous. The common theme I share with these models is that outside options determine prices: in OTC search models, the likelihood of meeting another counter-party tomorrow determines an agent's valuation today, whereas in my model, the number of dealers in a market determines a trader's market power and hence the terms of trade.

## 2.2 Model

I first discuss the effect information asymmetry has in a two party setting where there is no competition.

Consider a discrete-time infinite horizon game ( $t = 0, 1, 2, \dots$ ) in which a risk-neutral, uninformed dealer ( $m$ ) provides price quotes for a strategic trader ( $i$ ) with unknown trading motives. Both players are long-lived. With probability  $\alpha$ , the trader has superior information about an innovation in asset value, and is otherwise trading due to a private hedging motive. The fundamental common value of the asset  $v_t$  evolves as

$$\begin{aligned} v_t &= v_{t-1} + \eta_t y_t \Delta \\ \mathbb{P}(\eta_t = 1) &= \alpha \quad \mathbb{P}(\eta_t = 0) = 1 - \alpha \\ \mathbb{P}(y_t = 1) &= \mathbb{P}(y_t = -1) = \frac{1}{2} \end{aligned}$$

$\eta_t$  indicates whether trade is information driven or not, and, if so,  $y_t$  indicates whether the common value innovation is positive or negative. The common value innovations are symmetric with magnitude  $\Delta$ . These two shocks are assumed to be independent.

When trade is liquidity driven ( $\eta_t = 0$ ),  $y_t$  indicates the direction of the trader's private valuation, which has magnitude  $\delta$ , so  $i$ 's valuation in period  $t$  is

$$v_t + (1 - \eta_t)y_t\delta$$

Market maker  $m$  sets a price schedule  $P(q_t, \hat{\eta}_t)$  contingent on the trader's demand for the asset  $q_t$  and a report  $\hat{\eta}_t$  about the true value of  $\eta_t$ : trader  $i$  can indicate whether he is trading for informational or liquidity reasons.

After trade occurs, the end of period asset value  $v_t$  is revealed, so  $m$  knows whether  $i$  was truthful in his report of trading motive, and can respond with punishments if necessary.<sup>4</sup> Trader  $i$ 's realized period  $t$  profit  $\pi_t^i$  is

$$\pi_t^i = q_t[v_t + (1 - \eta_t)y_t\delta - P(q_t, \hat{\eta}_t)]$$

where  $q_t \in \{-1, 0, 1\}$  is the amount of the asset purchased by  $i$  (trade only occurs in whole units of the asset). Market maker  $m$ 's period  $t$  profit  $\pi_t^m$  is

$$\pi_t^m = q_t[P(q_t, \hat{\eta}_t) - v_t]$$

---

<sup>4</sup>This is a game with perfect public monitoring.

All agents maximize the present value of expected payoffs using discount rate  $\beta$ , which I interpret as the probability of continuation play: the trading relationship may end for some exogenous reason with probability  $(1 - \beta)$ . Their ex-ante utilities prior to trading in period  $t$  are

$$V_t^i = (1 - \beta) \mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \pi_t^i$$

$$V_t^m = (1 - \beta) \mathbb{E}_{t-1} \sum_{s=t}^{\infty} \beta^{s-t} \pi_t^m$$

Each player optimizes conditional on their information sets: trader  $i$  knows the value of the shocks  $\eta_t$  and  $y_t$  (and uses expectation operator  $\mathbb{E}_t$ ) while dealer  $m$  does not (so she uses  $\mathbb{E}_{t-1}$ ). After the period  $t$  asset value is revealed,  $h^t = \{h^{t-1}, P_t(\cdot, \cdot), q_t, \hat{\eta}_t, \eta_t, y_t\}$  denotes the updated history of play through time  $t$ . An equilibrium of the game is defined as follows:

**Definition 2.** A tuple  $\{P(\cdot, \cdot; h^{t-1}), q(h^{t-1}, \hat{\eta}(h^{t-1}))\}$  is an equilibrium of the repeated game if, for any time  $t$  and ex-ante history  $h^{t-1}$ :

1. Taking the market maker's pricing strategy  $P(q, \hat{\eta}; h^{t-1})$  as given, trader  $i$ 's trading motive report  $\hat{\eta}^*(h^{t-1})$  and demand  $q^*(h^{t-1})$  maximize his expected utility  $V_t^i$ .
2. Market maker  $m$ 's pricing rule  $P(q, \hat{\eta}; h^{t-1})$  maximizes her utility  $V_t^m$  given the traders reporting/demand strategy.

For simplicity, I restrict attention to pure-strategy, stationary equilibria in which  $m$  follows simple trigger strategies.<sup>5</sup> While more complex dynamics strategies are potentially interesting in their own right, I focus on the implications for *average* transaction costs in this paper, so this restriction is largely without loss of generality.

## 2.3 More informed traders pay lower transaction costs

Because of the symmetry in the model, the market maker is effectively posting a bid-ask spread around the prior period's asset value:

$$P(1, \hat{\eta}; h^{t-1}) = A(\hat{\eta}; h^{t-1}) = v_{t-1} + s(\hat{\eta})$$

$$P(-1, \hat{\eta}; h^{t-1}) = B(\hat{\eta}; h^{t-1}) = v_{t-1} - s(\hat{\eta})$$

where the symmetric spread  $s(\hat{\eta})$  is conditional on trader  $i$ 's reported motive  $\hat{\eta}$ . For expositional purposes, it is simpler to consider  $m$ 's choice of spreads  $s(\hat{\eta})$ , and trader  $i$ 's choice of whether to trade or not ( $q \in \{0, 1\}$ ) and whether to accurately report his motive—whether the trade is a buy or sell doesn't matter in describing the equilibrium.

---

<sup>5</sup> $m$  might be able to extract more surplus by offering quotes that either front-load or backload the surplus shared with trader  $i$ .



## For moderate hedging demand, trade breaks down in a static setting

Consider the stage game at time  $t$  taking the fundamental realization  $v_{t-1}$  at the end of the prior period as given and assuming no continuation play beyond the current period. The potential ex-ante gains from trade in the stage-game are  $(1 - \alpha)\delta$ , and the following proposition describes its equilibrium:

**Proposition 9.** *In the static game, there are effectively three-equilibrium outcomes depending on the parameter region:*

1. *If  $\delta < \alpha\Delta$ :  $m$ 's optimal spreads are  $s^*(0) = s^*(1) \geq \Delta$ , trader  $i$  does not trade ( $\hat{q}^*(\eta = 0) = \hat{q}^*(\eta = 1) = 0$ ), both players receive payoffs of 0 and no gains from trade are realized.<sup>6</sup>*
2. *If  $\delta \in (\alpha\Delta, \Delta)$ :  $m$ 's optimal spreads are  $s^*(0) = s^*(1) = \delta$ ; trader  $i$  always trades ( $\hat{q}^*(\eta = 0) = \hat{q}^*(\eta = 1) = 1$ ); trader  $i$ 's ex-ante payoff is  $V_t^i = \alpha(\Delta - \delta)$ ; market maker  $m$ 's ex-ante payoff is  $V_t^m = \delta - \alpha\Delta$ ; ex-ante gains from trade  $V_t^i + V_t^m = (1 - \alpha)\delta$  are achieved.*
3. *If  $\delta > \Delta$ :  $m$ 's optimal spreads are  $s^*(0) = s^*(1) = \delta$ ; trader  $i$  only trades in the liquidity state ( $\hat{q}^*(\eta = 0) = 1, \hat{q}^*(\eta = 1) = 0$ ); trader  $i$ 's payoff is 0; market maker  $m$ 's ex-ante payoff is  $\mathbb{E}V_t^m = (1 - \alpha)\delta$ ; ex-ante gains from trade  $V_t^i + V_t^m = (1 - \alpha)\delta$  are achieved.*

In the second and third regions of the parameter space above, all possible gains from trade are achieved and there is no extra social benefit to repeated interaction. In the second region, the gains are split: the trader profits on information trades, but the dealer extracts gains on liquidity trades. In the third region, the dealer is able to screen out information trade by setting a wide spread, while still capturing all liquidity rents because the trader is relatively desperate to trade in those states. In a perfectly competitive market a la Glosten and Milgrom (1985) trade will occur at zero-profit bid-ask prices

$$A_t = v_{t-1} + \alpha\Delta, B_t = v_{t-1} - \alpha\Delta$$

in regions (2) and (3). To focus on the benefits of repeated interaction, I'll assume throughout the rest of the paper that gains from trade are only achieved from repeated interaction:

**Assumption 3.** The private valuation discount  $\delta$  satisfies

$$\delta < \alpha\Delta$$

---

<sup>6</sup>There is a trivial equilibrium in which trade occurs at  $s^* = \Delta$  and payoffs/gain are still zero for both players. I'll ignore it because the presence of any other trading costs would eliminate it.

Assumption 3 captures the elasticity of liquidity demand to price: under this assumption, any anonymous / one-shot market breaks down. In a repeated OTC context, gains from trade can still be achieved if traders credibly signal whether trade is liquidity based or not. When  $\delta > \alpha\Delta$ , liquidity demand is less elastic to price, and trade is viable in an anonymous market.

## Relationships induce honesty and gains from trade

For any stationary symmetric pricing strategy, let  $s(\hat{\eta})$  denote the spread charged when  $i$  claims a trading motive of  $\hat{\eta}$ . We can define the spreads  $s_L = s(0)$  charged by the market maker on liquidity trade and  $s_I = \Delta - s(1)$  conceded by the market maker on informed trades. For any equilibrium involving sustained cooperation at stationary spreads  $(s_I, s_L)$ , because things are i.i.d.,

$$\begin{aligned} V_i^1 &= (1 - \alpha)(\delta - s_L) + \alpha s_I \\ V_m^1 &= (1 - \alpha)s_L - \alpha s_I \end{aligned}$$

where  $V_i^1$  and  $V_m^1$  are the ex-ante utilities of the trader,  $i$ , and market maker  $m$ . The one super-script denotes the number of market makers, for use when we consider multiple market makers later.  $m$ 's optimization problem ( $P$ ) can then be stated as

$$\max_{s_L, s_I} (1 - \alpha)s_L - \alpha s_I \quad (P)$$

such that:

$$\begin{aligned} (1 - \beta)s_I + \beta[\alpha s_I + (1 - \alpha)(\delta - s_L)] &\geq (1 - \beta)(\Delta - s_L) + \beta \cdot 0 & (IC_1) \\ (1 - \beta)(\delta - s_L) + \beta[\alpha s_I + (1 - \alpha)(\delta - s_L)] &\geq 0 & (IC_2) \\ \alpha s_I + (1 - \alpha)(\delta - s_L) &\geq 0 & (IR) \end{aligned}$$

where  $IC_1$  ensures that trader  $i$  doesn't disguise an informed trade as a liquidity trade (only applicable for  $s_L < \Delta - s_I$ ),  $IC_2$  ensures that trader  $i$  executes her liquidity trade through the dealer rather than abstaining (applicable when  $s_L > \delta$ ), and the  $IR$  constraint ensures that traders will want to continue on with the relationship under a given pricing strategy. Note that for both  $IC$  constraints, the market maker uses the credible threat of the static game Nash equilibrium, where trader  $i$ 's continuation value is 0.

The solution to this problem formulation is given by the following proposition:

**Proposition 10.** *The unique stationary strategy equilibrium of the repeated game with one market maker, is characterized by the parameter  $\delta$ . Define  $\underline{\delta}$  as*

$$\underline{\delta} = \frac{\alpha\Delta(1 - \beta)}{1 - \alpha^2\beta}$$

For  $\delta \in (\underline{\delta}, \alpha\Delta)$ ,

- Market maker  $m$  sets spreads contingent on the trader's report

$$s_L^* = \frac{\delta + \alpha\beta\Delta}{1 + \alpha\beta} \quad s_I^* = \frac{(1 - \alpha\beta)(\Delta - \delta)}{1 + \alpha\beta}$$

after every possible history in which the trader truthfully reports her motive. Following any history involving a false report,  $m$  plays the stage-game strategy in all future periods.

- Trader  $i$  truthfully reports her motive when spreads are quoted as above ( $\hat{\eta} = \eta, q = 1$ ), and plays the stage-game equilibrium for all other spreads.

For all  $\delta \in (0, \underline{\delta})$ , incentive compatibility can not be profitably maintained by  $m$ , and both players play the stage-game equilibrium in every period (no trade occurs).

The appendix derives the equilibrium spreads in detail, but, roughly, both *IC* constraints will bind at any optimum, and the *IR* constraint will be satisfied. The lower bound on the liquidity gain  $\underline{\delta}$  required to support trade is increasing in the information asymmetry parameters ( $\alpha, \Delta$ ) because bigger gains are required to reward traders for cooperating when adverse selection is more severe.

## Implied Price Dynamics and Cost Measurement

Given the equilibrium spreads above, I now examine how prices and corresponding measures of transaction costs will vary with the underlying parameters. First note that the evolution of fundamentals and prices is given

$$v_t = v_{t-1} + \eta_t y_t \Delta \tag{2.1}$$

$$P_t = v_{t-1} + (\Delta - s_I^*)\eta_t y_t + s_L^*(1 - \eta_t)y_t \tag{2.2}$$

We can consider two measures of costs:

1. The negative auto-covariance of price changes  $\gamma = -\mathbf{Cov}(\Delta P_t, \Delta P_{t-1})$  as in Bao, Pan, and Wang (2009).
2. A regression of prices changes  $\Delta P_t$  on buy sell indicators  $x_t$  in a typical reduced form decomposition of the spread into permanent (price-impact) and transitory (fixed) costs. These specifications typically model an underlying efficient price  $f_t$

$$f_t = f_{t-1} + w_t$$

$$w_t = \lambda x_t + u_t$$

where  $\lambda$  measures the permanent impact of a trade on fundamental value due to informational effects and  $x_t \in \{1, -1\}$  indicates whether a trade is a buy or a sell, where  $x_t$  and  $x_{t-1}$  are i.i.d. The observed transaction price process  $p_t$  is assumed to be

$$p_t = f_t + cx_t$$

where  $c$  represents the transitory cost a trader incurs. In the present setting, this cost reflects the markup dealers are able to charge, though in models that assume perfect competition, it is usually interpreted as a fixed order processing cost. Price changes are then given by

$$p_t - p_{t-1} = (\lambda + c)x_t - cx_{t-1} + u_t$$

where  $u_t$  is a residual term capturing public information innovations—in the model, these are zero, so costs can be estimated exactly from price and trade observations. This specification matches Huang and Stoll (1997) (minus inventory effects) and Madhavan, Richardson, and Roomans (1997).

I'll focus on the second measure for several reasons: 1) the covariance measure  $\gamma$  is closely related to  $c$ ; 2) the estimates from the second specification have a simpler form, facilitating easier comparative statics.

**Proposition 11.** *In a monopolistic market, where the true price dynamics are given by equations (1) and (2), the estimated transitory cost  $\hat{c}$  and price impact  $\hat{\lambda}$  are given by*

$$\hat{c} = (1 - \alpha)s_L^* - \alpha s_I^* \quad \hat{\lambda} = \alpha\Delta$$

*with comparative statics*

$$\frac{\partial \hat{c}}{\partial \alpha} < 0 \quad \frac{\partial \hat{c}}{\partial \Delta} < 0$$

The intuition for the cost estimate is that, on average, traders pay the liquidity spread  $s_L^*$  with probability  $(1 - \alpha)$ , and recoup the informed spread concession  $s_I^*$  with probability  $\alpha$ . The comparative static demonstrates the first main result of the paper: When trader's are more informed – either by being more frequently informed ( $\alpha$ ), or having better information (about larger price moves  $\Delta$ ) – they pay lower costs. Higher  $\alpha$  implies the trader cares less about future gains from liquidity trade, and needs to be compensated with lower costs today in order to ensure cooperation. Higher  $\Delta$  makes disguising an informed trade as a liquidity trade more tempting today, and requires a similar concession from the market maker.

In a market characterized by anonymous trading, where there are no relationship effects, transaction costs will be unrelated to the information parameters  $\alpha$  and  $\Delta$ , and hence to permanent price impact estimates ( $\hat{\lambda}$ ). Here, the comparative statics of  $\lambda$  and  $c$  have opposite signs with respect to each information asymmetry parameter. Controlling for the effects of other parameters, a negative relationship between cost and price impact estimates in a given market is indicative of potential relationship effects.

## 2.4 Competition indirectly reduces transaction costs

This section demonstrates the second main result of the paper: that the comparative statics of transaction costs only make sense conditional on market structure—the viable number of dealers in a given market—and that the concentration of dealers has an endogenous limit. To

address this issue in a tractable way, consider a matching extension to the one-dealer/one-trader model presented above.

There is a unit mass of traders  $i \in [0, 1]$ , and a mass of dealers with exogenous measure  $\mu \leq 1$  indexed by  $m$ . Because this is an over the counter context, dealers do not simultaneously compete via price quotes: traders can only request a quote from the dealer they are currently matched with.

Maintaining a relationship with traders is costly, so dealers are limited in the number of relationships they can have at one time—I'll assume they can not be in more than one trading relationship.<sup>7</sup> Each period, existing relationships end with exogenous probability  $(1 - \beta)$ , and unmatched players of both types die with probability  $(1 - \beta)$ . In both cases, players are replaced by new players of the same type.<sup>8</sup>

Dealers in this setup still have the potential to generate trade and extract some surplus, but their threat is less powerful than the monopoly case. They can terminate a relationship and send traders back to the unmatched pool, but the severity of this threat depends on how easy it is to form another relationship, i.e. on the mass of dealers  $\mu$ . Gains from trade via repeated trade are only achievable if  $\mu < 1$ ; otherwise the dealers have no credible threat because traders are sure to be matched in the next period with a new dealer. The unmatched pool of players will consist entirely of traders.

In this setting, a dealer-trader pair only observe the history  $h^t$  of their prior transactions, and act upon those. Once a relationship is terminated, any new matches begin with a fresh (null) history. The equilibrium is similar, with the addition that things need to be in a steady state:

**Definition 3.** A tuple  $\{P(\cdot, \cdot; h^{t-1}), q(h^{t-1}, \hat{\eta}_t(h^{t-1}))\}$  is an equilibrium of the repeated game if, for any time  $t$  and ex-ante history  $h^{t-1}$  of a given dealer-trader pair:

1. Taking the market maker's pricing strategy  $P(q, \hat{\eta}; h^{t-1})$  as given, trader  $i$ 's trading motive report  $\hat{\eta}^*(h^{t-1})$  and demand  $q^*(h^{t-1})$  maximize his expected utility  $V_t^i$ .
2. Market maker  $m$ 's pricing rule  $P(q, \hat{\eta}; h^{t-1})$  maximizes her utility  $V_t^m$  given the traders reporting/demand strategy.
3. Flows of traders / market makers into the unmatched pool are such that the fraction of traders in a relationship is constant.

The resulting outcome is similar to the monopolist case, but the level of competition in the market  $\mu$  alters spreads and the range of  $\delta$  for which gains from trade are feasible.

---

<sup>7</sup>Allowing them to be in some finite number  $N$  at a given time will not qualitatively change the results.

<sup>8</sup>Allowing one player in a relationship to die while the other continues complicates the algebra without changing much qualitatively. See section 5.2 of Mailath and Samuelson (2006) for a discussion.

**Proposition 12.** *There exists a stationary equilibrium of the repeated matching game characterized by the parameter  $\delta$ . Define  $\underline{\delta}'$  as*

$$\underline{\delta}' = \frac{\alpha\Delta(1-\beta)}{1-\beta\mu-\alpha^2\beta(1-\mu)}$$

For  $\delta \in (\underline{\delta}', \alpha\Delta)$ , and for each dealer-trader pairing  $(m, i)$

- Market maker  $m$  sets spreads contingent on trader  $i$ 's report

$$s_L^* = \frac{\delta + \alpha\beta - \beta\mu(\delta + \alpha\Delta)}{1 + \beta[\alpha - (1 + \alpha)\mu]} \quad s_I^* = \frac{[1 - \alpha\beta - \beta\mu(1 - \alpha)](\Delta - \delta)}{1 + \beta[\alpha - (1 + \alpha)\mu]}$$

after every possible history in which the trader truthfully reports her motive. Following any history involving a false report,  $m$  plays the stage-game strategy in all future periods.

- Trader  $i$  truthfully reports her motive when spreads are quoted as above ( $\hat{\eta} = \eta, q = 1$ ), and plays the stage-game equilibrium for all other spreads.

For all  $\delta \in (0, \underline{\delta}')$ , incentive compatibility can not be profitably maintained by  $m$ , and both players play the stage-game equilibrium in every period (no trade occurs).

Note that for  $\mu = 0$ , the equilibrium spreads reduce to those of Proposition 10. Dealers are not competing with each other directly on prices, but there is an indirect benefit to traders of having more dealers in the market: they improve a trader's outside option. When  $\mu$  is higher, the pool of unmatched traders is smaller, and if a trader deviates he can expect to be in a new trading relationship sooner.

All else equal, the comparative statics of estimated transaction costs with respect to the information asymmetry parameters are the same, but the effect of dealer competition also changes costs:

**Proposition 13.** *In the repeated matching game with dealer mass  $\mu$ , the estimated transaction costs  $\hat{c}$  and price impact  $\hat{\lambda}$  are given by*

$$\hat{c} = (1 - \alpha)s_L^* - \alpha s_I^* \quad \hat{\lambda} = \alpha\Delta$$

with comparative statics

$$\frac{\partial \hat{c}}{\partial \alpha} < 0 \quad \frac{\partial \hat{c}}{\partial \Delta} < 0 \quad \frac{\partial \hat{c}}{\partial \mu} < 0$$

This illustrates the second main point of the paper: in drawing conclusions about transaction costs in OTC markets, market structure matters. Whereas the equity markets literature generally assumes prices are set competitively, this assumption is not as tenable in less liquid markets. As such, the level of competition in a given asset should either be considered

directly as a determinant of transaction costs, or it should be controlled for when attributing variation in costs to other factors.

The previous results took the level of competition in a market,  $\mu$ , as exogenously given. However, there is an endogenous limit on the measure of dealers that still allows trade to occur. Once it is exceeded, market making is no longer profitable at the spreads required to maintain trader cooperation, and the market breaks down.

**Proposition 14.** *The imperfectly competitive matching game has a limit  $\bar{\mu}$  on the mass of market makers that can be supported. This limit is decreasing in the information asymmetry parameters  $\alpha$  and  $\Delta$ .*

When information asymmetry increases, via the parameters  $\alpha$  or  $\Delta$ , it is harder to keep traders cooperative so they must be given better terms of trade. When the concentration of dealers is  $\mu = \bar{\mu}$ , dealers are already earning zero profits, so they can not cede any more gains. The only way cooperation can be maintained is if the trader's outside option is reduced by decreasing the number of dealers in the market, thereby increasing the cost of returning to the matching pool.

While repeated interaction does allow gains to be achieved in a matched dealer-trader pair, removing the static-game inefficiency, asymmetric information still causes an aggregate inefficiency by limiting the viable size of the liquidity supply market. The total ex-ante gains from trade in each pairing is  $(1 - \alpha)\delta$ , so the total gains are  $\mu(1 - \alpha)\delta$ . This quantity is also decreasing in both information asymmetry parameters.

## 2.5 Conclusion and Extensions

This article demonstrates that in markets characterized by non-anonymous trade and adverse selection, a higher level of information asymmetry may actually lead to lower costs. Repeated interaction allows dealers to screen out informed trade, while superior information gives traders more market power against them, improving prices. The number of dealers in a given market also affects prices, hence inference about transaction costs in a repeated setting should, ideally, condition on market structure. Finally, assuming free entry, the maximum number of dealers in a given market is endogenously determined, and is decreasing in informed trade—repeated trade solves inefficiencies at the relationship level, but can not remove all aggregate inefficiency.

There are several extensions and improvements that could be made to the above analysis. First, non-stationary strategies will effect the dynamics of prices, but it's not clear whether they will alter measurement of average transaction costs. Second, a potentially useful extension is to consider additional quantities—when traders can transact on more than one unit, what do price schedules look like? Will they split their orders? The transaction cost literature for bond markets has identified schedules as being concave in quantity—can we match that stylized fact in this setting?

Finally, from a theoretical perspective, a useful benchmark would be to cast this model in the CARA-normal framework of Glosten (1989) and Biais, Martimort, and Rochet (2003) to at least examine the monopolist case. This framework is already used to compare the welfare implications and price schedules of static market structures (e.g. monopoly vs. oligopoly), so it may provide more insight into the conditions under which non-anonymous markets are superior to anonymous markets.



## Chapter 3

# Measuring mutual fund performance with endogenous asset prices

### 3.1 Introduction

Do mutual fund managers add value? This question has been the focus of a large empirical literature that measures mutual fund performance. In the theoretical literature, most models of mutual funds rationalize fund performance puzzles by isolating the value added by managers from the underlying financial assets traded in the economy.<sup>1</sup> Yet treating risky asset prices as an exogenous side show only make sense if mutual funds are a relatively small component of the market. Allen (2001) argues that because this is not the case—the share of corporate securities owned directly by households dropped from 90% in 1950 to less than 40% by 2000—the effects of institutions on asset prices are a first order concern. As of 2010, mutual funds alone owned 27% of U.S. equities.<sup>2</sup> In aggregate, then, investor demand for mutual funds potentially affects their demand for direct investment in risky assets.

I use a model of fund management that incorporates this pricing effect to produce two results. First, I show that while measuring fund performance against a risk adjusted benchmark indicates whether managers are skilled, it does not imply welfare is improved by the presence of mutual funds—even when these measures are positive, investors might be better off in a world without fund management. The literature on incomplete markets has shown that the introduction of new securities can reduce welfare,<sup>3</sup> so my contribution is simply to show that standard mutual fund performance measures do not capture welfare effects.

Second, I show that when pricing effects are accounted for, some existing explanations of mutual fund underperformance are harder to maintain. In models explaining fund performance puzzles such as Berk and Green (2004) and Glode (2011), fund managers control

---

<sup>1</sup>A notable exception is Savov (2010), who uses a market timing story to explain fund underperformance in an REE economy

<sup>2</sup>See Investment Company Institute (2011)

<sup>3</sup>See Duffie and Rahi (1995) for a survey of this literature.

demand for their active management technology, and hence profits, via their management fee. Assumptions on the characteristics of the payoffs being sold determine the optimal fee and the puzzle explanation. However, any effect on mutual fund demand via fees has a corresponding effect on direct demand for the underlying risky asset, affecting its price and potentially the pricing kernel. Once this joint effect is accounted for by fund managers, equilibrium fees are different, and so are performance measures used to address puzzles.

What are the puzzles people are trying to explain? Beginning with Jensen (1968), the debate initially focused on why risk-adjusted returns, after fees, were not positive if mutual fund managers had skill. Berk and Green (2004) (BG) changed the terms of the debate by showing that in a rational model of fund flows, ex-post fund performance *should* be zero. However, evidence of mutual fund underperformance after fees, including Gruber (1996) and Fama and French (2010), is not explained by their model. So far, several competing models attempt to address this negative performance puzzle. Pastor and Stambaugh (2010) attempt to show that if returns to scale are decreasing at the industry (not manager) level, and investors need to learn the parameters governing this process, they might still invest in funds despite a long history of observed underperformance. Citing evidence that mutual funds tend to perform relatively well in bad times,<sup>4</sup> Glode (2011) augments the BG model to allow for state-dependent alpha and shows that unconditional risk-adjusted returns, when measured against a passive benchmark, may be rationally negative. Whereas Glode's explanation is cross-sectional, Savov (2010) provides a market-timing explanation involving heterogeneous wealth shocks. All but the latter abstract from investor behavior and leave the underlying asset price as exogenous.

To show that pricing effects matter in explaining these puzzles, I use a simple model of fund management closest to BG and Glode (2011). The zero fund performance result in BG follows from several key assumptions: 1) completely diversifiable and idiosyncratic fund payoffs; 2) an unlimited supply of capital that flows to any positive excess return opportunity; 3) decreasing returns to scale in fund size; 4) fund manager market power. Glode assumes an exogenous pricing kernel, allows fund performance to vary with the pricing kernel (so its risk is not completely idiosyncratic/diversifiable), allows managers to produce this state-dependent alpha using a decreasing returns to scale technology, and again gives managers market power in setting fees. Glode's result is that managers optimally choose to provide alpha in high marginal utility states, so funds provide insurance, and their unconditional risk premium is therefore negative when measured against an imperfect pricing kernel proxy (such as a passive benchmark).

I allow fund payoffs to be correlated with the single risky asset in the economy (the "market" portfolio), where this correlation is exogenous – BG corresponds closely to the case where this correlation is zero, Glode to the case where it is negative. The key assumption I remove is that capital is in unlimited supply – the aggregate investor in the economy has a budget constraint that needs to be satisfied. He has to hold the outstanding supply of the risky asset and requires compensation for this risk. If fee changes shift money to/from

---

<sup>4</sup>See for example Kacperczyk, Van Nieuwerburgh, and Veldkamp (2009)

mutual funds, this changes the price of the underlying asset and hence the pricing kernel – so imposing this budget constraint is equivalent to relaxing Glode’s assumption of an exogenous pricing kernel.

Because the fund payoff is imperfectly correlated with the market portfolio, its risk is not completely diversifiable as in BG. I do not impose any decreasing returns to scale in alpha generation – this means the fund is not any less attractive as delegated investment increases, but uncertainty in alpha still limits investor willingness to invest in the fund. I do not need decreasing returns to scale to get an interior solution, and it does not qualitatively effect the results of the paper. I retain the notion of manager market power as in the existing models to give them the best chance of succeeding, and show the main result holds whether we consider a single monopolist fund manager who internalizes his effect on prices or a continuum of small managers who do not internalize this effect, but best-respond to the fees set by other managers in the economy. For the purposes of this discussion, I am being agnostic about how managers add value – we can think of their value added as coming from better stock selection or lower costs, for example.

In other areas of the delegated portfolio management literature, models *do* endogenize asset pricing effects. However, they are generally not concerned with the mutual fund performance puzzle. Instead, they focus on things like equilibrium contracting under agency frictions Cuoco and Kaniel (2006), explanations for momentum/reversal Vayanos and Woolley (2008), optimal strategies for information sale Admati and Pfleiderer (1990), excess volatility Guerrieri and Kondor (2009), etceteras.<sup>5</sup>

Given the result of this paper – that existing fund performance puzzles remain once asset prices are endogenous – the larger question remains unanswered: why would anyone rationally invest in mutual funds if they underperform in aggregate after fees? Additionally, because performance measures do not necessarily reflect the welfare gains of fund management, how can we measure the true value added by active fund management?

## 3.2 Model

Consider the following three period ( $t = 1, 2, 3$ ) model of fund management. There is a unit mass of risk averse investors with exponential utility over terminal wealth  $W$

$$U(W) = -\exp(-\rho W)$$

with initial wealth normalized to 0.

There is also a unit mass of portfolio managers who possess an identical active management technology that delivers a normally distributed payoff  $\epsilon \sim \mathcal{N}(\alpha, \sigma_\epsilon^2)$ . Managers do not have capital of their own, but they sell exposure to this payoff at a fee of  $f$  per unit. For sim-

---

<sup>5</sup>See Stracca (2006) and Bhattacharya, Dasgupta, Guembel, and Prat (2007) for more complete surveys of the delegated portfolio management literature.

plicity, this active management technology exhibits constant returns to scale—the qualitative results of the paper do not change if I use a decreasing returns to scale technology.<sup>6</sup>

There is a risky asset with a supply of one unit that pays a normally distributed dividend  $D \sim \mathcal{N}(\mu, \sigma^2)$ , and a risk free asset in infinite supply with net return normalized to  $R_f = 0$ . The risky asset and the portfolio management payoffs  $D$  and  $\epsilon$  have a correlation  $\gamma$ . Investors demand quantities  $x_S$  and  $x_F$  of the risky asset and the fund to maximize the expected utility of their terminal wealth.

The sequence of events is as follows:

- $t = 1$ : Fund managers set fee  $f$  simultaneously, taking the fees set by other managers as given and anticipating investor demand  $x_F$ . Each given manager chooses  $f$  to maximize profits  $\pi = f \cdot x_F$ .
- $t = 2$ : Investors are randomly matched with fund managers and make portfolio choices. They directly buy  $x_S$  shares of the risky asset at the market clearing price  $P$ , and  $x_F$  shares of the fund at price  $f$ .
- $t = 3$ : The asset payoff  $D$  and fund performance  $\epsilon$  are realized, and investors consume terminal wealth.

When investors choose active management, they are randomly matched with a fund manager. The fund manager therefore has “local” monopoly power over the investor he is matched with and will set fees accordingly. There is no benefit to diversification across managers, as they all have the same  $\alpha$  technology. For robustness, I later consider a single monopolist manager in charge of all delegated funds, but it does not qualitatively affect the results. In both cases, managers are able to extract some of the rents associated with their portfolio management skills, as is standard in the literature.

Agents in the model are ex-ante identical, so I focus on symmetric equilibria. Given the parameters of the economy  $\theta = (\rho, \mu, \sigma^2, \alpha, \sigma_\epsilon^2, \gamma)$ , an equilibrium is defined as follows:

**Definition 4.** An equilibrium of the economy is given by

- A pair of investor demand schedules  $(x_S^*(P, f; \theta), x_F^*(P, f; \theta))$  specifying direct and indirect investment in the risky asset, respectively, given asset price  $P$  and fund fee  $f$ .
- An asset price  $P^*(f; \theta)$  such that markets clear given symmetric fund fee  $f$ :  $x_S^*(P, f; \theta) = 1$ .
- A symmetric per-share fund fee  $f^*(\theta)$  charged by all fund managers that is a best-response for an atomistic manager given all other managers charge  $f^*$ .

---

<sup>6</sup>Berk and Green (2004) and Glode (2011) effectively need decreasing returns to scale to arrive at interior solutions; without it, mutual fund investment would grow without bound. Here, fund investment will be limited by risk in the fund payoff and investors’ budget constraints.

All of the pricing effects in the model depend on  $\mu - P$ , so I express results in terms of this premium rather than the price alone (this removes a model parameter).

To make sure that fund management is valuable, I impose

**Assumption 4.** The expected gross fund return  $\alpha$  satisfies

$$\alpha > \rho\gamma\sigma\sigma_\epsilon$$

Assumption 4 insures that active fund management is justified – for  $\alpha$  below this bound, even if the fund is free, investors will want to short it. Note that if  $\gamma < 0$ , gross  $\alpha$  can be negative and fund management is still viable, at least *before* fees, because it provides a hedging gain.

### 3.3 Equilibrium and Performance Measurement

This section solves for equilibrium quantities and derives performance measures. First note that because of the CARA-normal setup, investors equivalently maximize their certainty equivalent of wealth

$$\mathbb{E}U(W) = \mathbb{E}W - \frac{\rho}{2}\mathbf{V}ar(W)$$

If investors buy  $x_S$  units of the risky asset and  $x_F$  units of the fund, the moments of terminal wealth are given by

$$\begin{aligned} W &= x_S(D - P) + x_F(\epsilon - f) \\ \mathbb{E}W &= x_S(\mu - P) + x_F(\alpha - f) \\ \mathbf{V}ar(W) &= x_S^2\sigma^2 + x_F^2\sigma_\epsilon^2 + 2x_Sx_F\gamma\sigma\sigma_\epsilon \end{aligned}$$

Taking the symmetric fund management fee as given, first order conditions and the market clearing constraint imply

**Proposition 15.** *Given parameters  $\theta = (\rho, \sigma^2, \alpha, \sigma_\epsilon^2, \gamma)$  and management fee  $f$ , optimal prices and demand schedules are given by:*

$$\begin{aligned} \mu - P^*(f) &= \rho\sigma^2(1 - \gamma^2) + (\alpha - f)\gamma\frac{\sigma}{\sigma_\epsilon} \\ x_F^*(P, f) &= \frac{\sigma(\alpha - f) - \gamma\sigma_\epsilon(\mu - P)}{\rho\sigma\sigma_\epsilon^2(1 - \gamma^2)} = \frac{\alpha - f - \rho\gamma\sigma\sigma_\epsilon}{\rho\sigma_\epsilon^2} \\ x_S^*(P, f) &= \frac{\mu - P}{\rho\sigma^2} - \frac{\gamma\sigma_\epsilon}{\sigma}x_F = 1 \end{aligned}$$

The obvious thing to note here is that the risky asset price is a function of fund performance moments and fees—fund manager skill and behavior affect aggregate asset price and hence the pricing kernel. In the existing literature on fund manager performance, the pricing kernel is assumed to be exogenous.

Now consider a given (atomistic) fund manager's problem: he will be matched with a single investor. Let  $\bar{f}$  be the fee charged by other managers, which determines the aggregate asset price  $(\mu - P^*(\bar{f}))$ . The fee the individual manager charges,  $f$ , will directly effect investor demand for the fund. Taking  $\bar{f}$  as given, the manager's profit ( $\pi$ ) maximization problem is

$$\begin{aligned} \max_f \pi(f; \bar{f}) &= f x_F^*(P^*(\bar{f}), f) \\ &= f \frac{\sigma(\alpha - f) - \gamma \sigma_\epsilon(\mu - P^*(\bar{f}))}{\rho \sigma \sigma_\epsilon^2 (1 - \gamma^2)} \end{aligned}$$

Solving for the fixed point  $f^* = \bar{f}$  gives us the equilibrium fee:

**Proposition 16.** *The symmetric management fee charged by all managers is*

$$f^*(\theta) = \frac{(1 - \gamma^2)(\alpha - \rho \gamma \sigma \sigma_\epsilon)}{2 - \gamma^2}$$

Now that endogenous quantities are determined, what do measures of fund manager performance look like? Because of the CARA-normal setting, it is most natural to measure things on a per-share basis. Let  $S = D - P$  and  $F = \epsilon - f$  be the realized returns to one share of direct and indirect investment, respectively. The risk-adjusted excess return on the fund  $\hat{\alpha}$  can then be defined as

$$\hat{\alpha} = \mathbb{E}F - \beta_{F,S} \mathbb{E}S$$

where  $\beta_{F,S}$  regresses fund return on the market (direct/passive investment). Then, using the equilibrium quantities,

**Proposition 17.** *Ex-ante (expected) fund performance, net of fees, is given by*

$$\begin{aligned} \beta_{F,S} &= \gamma \frac{\sigma_\epsilon}{\sigma} \\ \hat{\alpha} &= \frac{(1 - \gamma^2)(\alpha - \rho \gamma \sigma \sigma_\epsilon)}{2 - \gamma^2} \end{aligned}$$

Figure 3.1 shows the equilibrium quantities. The optimal fee is non-monotonic in  $\gamma$  – as  $\gamma$  decreases from 1, the fund becomes more valuable as hedging tool, and managers can charge higher fees. Eventually, the pricing effect, which they do not account for in their individual fee, reduces the fee they can charge to 0. The symmetric equilibrium  $\hat{\alpha}$  is 0 when the fund is perfectly correlated—positively or negatively—with the risky asset payoff. The risky asset price adjusts to offset any excess return added by the managers (they do not internalize this effect), neutralizing the value they are adding.

The performance measure  $\hat{\alpha}$  is always positive under Assumption 4. This suggests that the negative performance puzzle in the literature is more difficult to explain when endogenous effects on the asset price are accounted for. For example, Glode (2011) uses an exogenous pricing kernel to show that if fund managers use a technology that pays off in bad states of the world—which corresponds to a negative correlation  $\gamma$  here—measured performance  $\hat{\alpha}$  can be negative. When the effect on the underlying risky asset (the market) is accounted for, this does not seem to be the case.

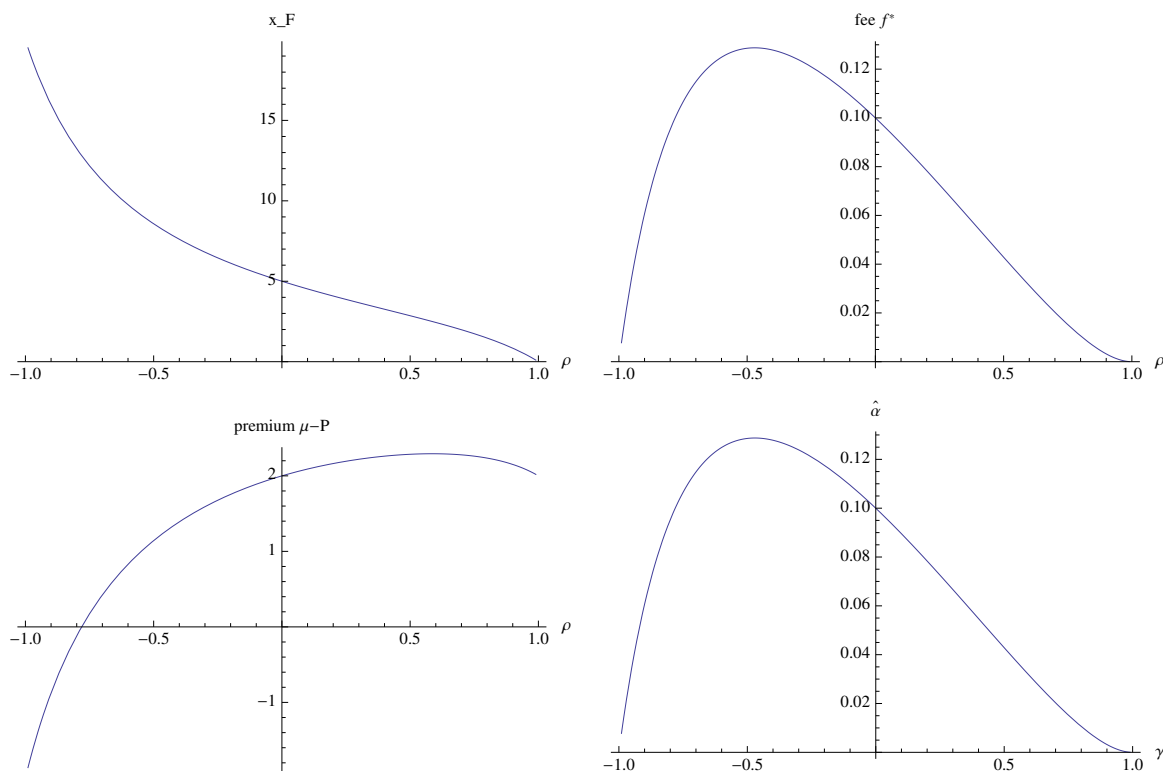


Figure 3.1: Clockwise from top left: the equilibrium fund position  $x_F^*$ ; fee  $f^*$ ; performance measure  $\hat{\alpha}$ ; risk premium  $\mu - P$ . Parameters:  $\rho = 2, \sigma = 1, \alpha = 0.2, \sigma_\epsilon = 0.1$

### 3.4 Active management does not necessarily add value

The previous section showed that under Assumption 4, which requires that fund managers are offering a desirable product, performance measures are weakly positive. I.e. after adjusting for market risk and fees, investors still earn a positive excess return. This would seem to imply that managers are adding value to the economy, but that is not necessarily the case if we consider the counter-factual: would investors be better off if fund management did not exist?

To address this, let  $V_0$  be the ex-ante utility of investors in the absence of fund management, when they invest exclusively in the risky and risk-free assets. Let  $V_I$  and  $V_M$  be the ex-ante utilities of investors and managers in the presence of fund management, so that  $V_S = V_I + V_M$  is total welfare. Let  $\Delta V_I = V_I - V_0$  and  $\Delta V_S = V_S - V_0$  be the changes in investor and social welfare associated with fund management. These values are given by

**Proposition 18.** *Ex-ante utility  $V_0$  and the incremental gains from active management are*

given by

$$\begin{aligned} V_0 &= \frac{\rho\sigma^2}{2} \\ V_M &= \frac{(1 - \gamma^2)(\alpha - \rho\gamma\sigma\sigma_\epsilon)^2}{\rho\sigma_\epsilon^2(2 - \gamma^2)^2} \\ \Delta V_I &= \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)(\alpha + \rho\gamma\sigma\sigma_\epsilon[3 - 2\gamma^2])}{2\rho\sigma_\epsilon^2(2 - \gamma^2)^2} \end{aligned}$$

Note that under Assumption 4,  $V_M > 0$ , so fund managers always earn profits. However, it is not always the case that fund management improves total welfare or make investors better off:

**Proposition 19.** *For  $\gamma > 0$ , investor and total welfare are always enhanced by the presence of fund management as long as Assumption 4 holds. For  $\gamma < 0$ , there exist thresholds  $\alpha_I$  and  $\alpha_S$  with*

$$\alpha_I > \alpha_S > \rho\gamma\sigma\sigma_\epsilon$$

such that

- $\Delta V_I < 0$  and  $\Delta V_S < 0$  for  $\alpha < \alpha_S$ .
- $\Delta V_I < 0$  and  $\Delta V_S > 0$  for  $\alpha \in (\alpha_S, \alpha_I)$ .
- $\Delta V_I > 0$  and  $\Delta V_S > 0$  for  $\alpha > \alpha_I$ .

When Assumption 4 holds, fund management adds value as long as  $\gamma > 0$ . When  $\gamma < 0$ , the fund is attractive on an individual/atomistic level, but equilibrium effects on the underlying asset price and fees are welfare reducing for  $\alpha < \alpha_S$ : managers earn profits, but these profits do not offset the loss in investor welfare relative to the benchmark case  $V_0$ . If there were no pricing effects—if we took the underlying risky asset price as exogenous and introduced fund management—welfare would be improved.

Once  $\alpha > \alpha_S$ , investors are still worse off, but manager profits exceed these losses, increasing social welfare. Finally, when  $\alpha > \alpha_I$ , fund management adds significant excess returns, improving both investor and social welfare.

These regions are shown in Figure 3.2. Note that for  $\alpha < \alpha_I$  (regions *B* and *C*), the performance measure  $\hat{\alpha}$  is still always positive, so it looks like investors are benefitting from fund management. In effect, using this type of performance measure is misleading because the benchmark itself—the risky asset price—is endogenously affected by the presence of fund management. These measures still accurately reflect manager “skill”, but they do not imply that the presence of such skill is a good thing.



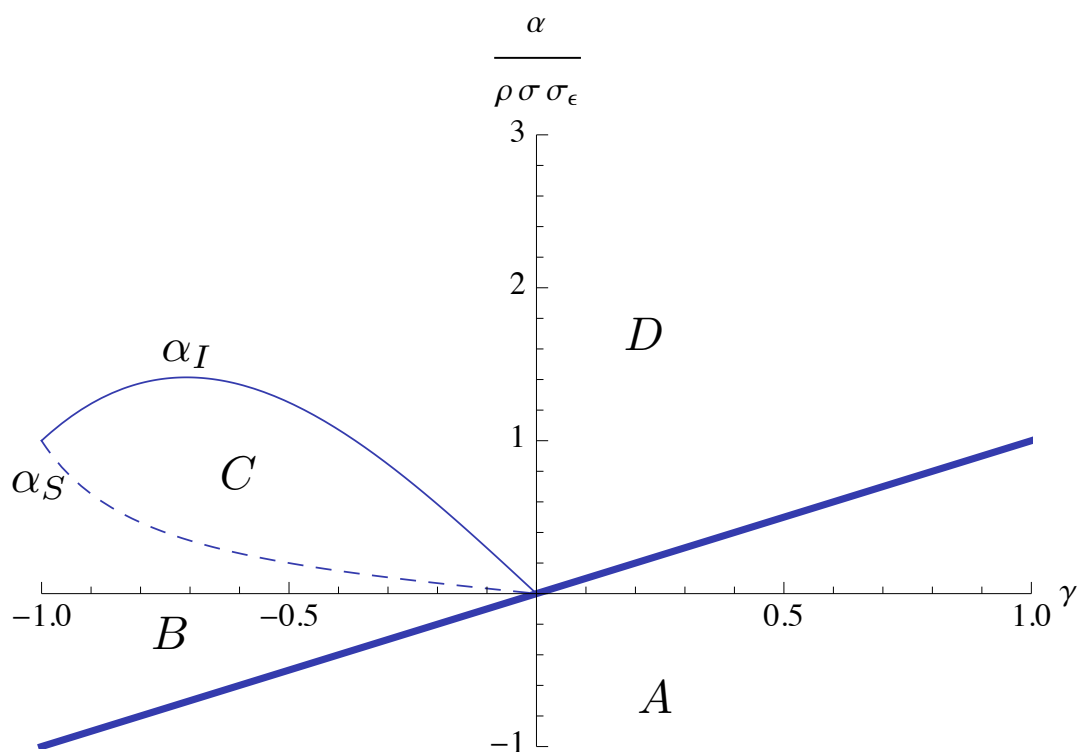


Figure 3.2: This graph shows welfare outcomes for different values of (normalized)  $\alpha$  as a function of  $\gamma$ . For  $\alpha < \alpha_0$  (region *A*), Assumption 4 is violated and fund management is not viable. In region *B*,  $\alpha < \alpha_S$  so fund management is viable but welfare reducing. In region *C*,  $\alpha \in (\alpha_S, \alpha_I)$  fund management increases total welfare but investor welfare is reduced. In region *D* investor welfare is also improved.

### 3.5 Robustness

This section considers two variation of the model—a monopolist fund manager and decreasing returns to scale in generating fund returns—and shows that the qualitative results of the paper are unchanged. Measured fund performance is generally positive and does not necessarily indicate that fund management is welfare improving.

#### Monopolist fund management

The main results of the paper assumed assumed a continuum of fund managers set fees symmetrically. Supposed instead that there is a large, monopolist fund manager. She will set a fee  $f$ , and this will effect the aggregate asset price. Relative to the continuum of managers, the monopolist will internalize her effect on prices in setting a fee.

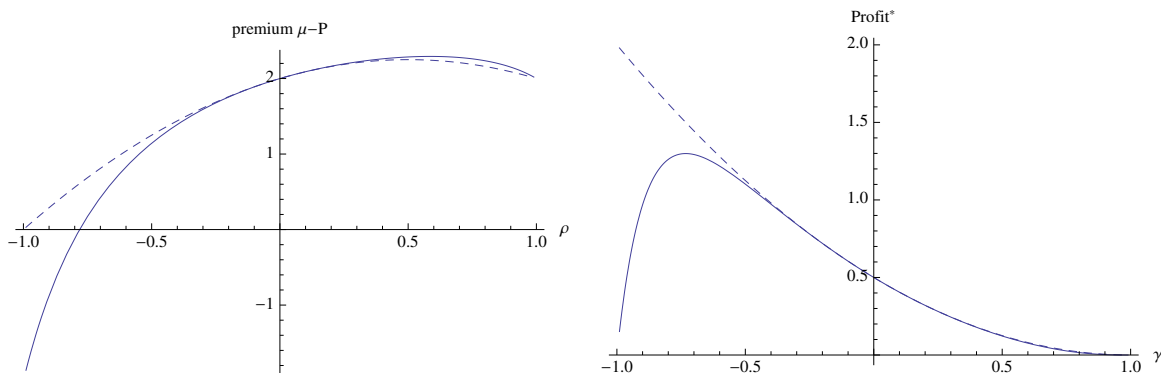


Figure 3.3: Left: the equilibrium risk premium  $\mu - P$ ; Right: equilibrium manager profit. Solid lines are the symmetric case, dashed are the monopolist case. Parameters:  $\rho = 2, \sigma = 1, \alpha = 0.2, \sigma_\epsilon = 0.1$

Conditional on a given fee, the equilibrium demands and prices in Proposition 15 are unchanged. The monopolist's profit maximization problem is then

$$\begin{aligned} \max_f \pi(f) &= f x_F^*(P^*(f), f) \\ &= f \frac{\sigma(\alpha - f) - \gamma \sigma_\epsilon (\mu - P^*(f))}{\rho \sigma \sigma_\epsilon^2 (1 - \gamma^2)} \end{aligned}$$

and the corresponding fee and performance measures are

**Proposition 20.** *The equilibrium fee and performance measure under monopolist fund management are*

$$\begin{aligned} f^*(\theta) &= \frac{\alpha - \rho \gamma \sigma \sigma_\epsilon}{2} \\ \hat{\alpha} &= \frac{(1 - \gamma^2)(\alpha - \rho \gamma \sigma \sigma_\epsilon)}{2} \end{aligned}$$

The performance measure is still positive for  $\gamma \in (-1, 1)$ . There is a difference in the monopolist fee, as shown in Figure 3.3: because the fund manager internalizes her effect on prices, she is still able to extract rents when the fund is perfectly correlated (positively or negatively) with the underlying risky asset. This was not possible when she was one of a small continuum of managers, where demand for the fund and the risky asset drive down the risky asset premium.

Modeling the fund manager as a monopolist also does not change the main welfare and measurement conclusions:

**Proposition 21.** *With a monopolist fund manager, for  $\gamma > 0$ , investor and total welfare are always enhanced by the presence of fund management as long as Assumption 4 holds.*

For  $\gamma < 0$ , there exist thresholds  $\alpha_I^m$  and  $\alpha_S^m$  with

$$\alpha_I^m > \alpha_S^m > \rho\gamma\sigma\sigma_\epsilon$$

such that

- $\Delta V_I < 0$  and  $\Delta V_S < 0$  for  $\alpha < \alpha_S^m$ .
- $\Delta V_I < 0$  and  $\Delta V_S > 0$  for  $\alpha \in (\alpha_S^m, \alpha_I^m)$ .
- $\Delta V_I > 0$  and  $\Delta V_S > 0$  for  $\alpha > \alpha_I^m$ .

As shown in Figure 3.4, while the boundaries that determine whether fund management is welfare improving shift relative to the symmetric case, the qualitative result remains: performance measures can be positive even though mutual funds reduce welfare.

## Decreasing returns to scale

Many existing mutual fund models, including Berk and Green (2004), Glode (2011), and Pastor and Stambaugh (2010), posit that fund performance decreases in the size of a fund. This assumption is partially motivated by intuition, but it also produces interior solutions to their models—without imposing it, fund demand would increase without bound.

Decreasing returns to scale were not necessary in deriving the results thus far because fund performance is risky and an investor's demand is limited by his budget constraint—in addition to the fund, investors still need to hold the underlying risky asset in the economy. Nonetheless, imposing this assumption does not change the prior results.

To incorporate decreasing returns to scale as in the existing literature, suppose that managers incur a cost  $cx_F^2/2$  in delivering  $x_F$  shares of their fund to investors. This will effect the optimal fee they decide to charge, but not the demand and price of Proposition 15, which takes the fee as given. A given manager's maximization problem, taking the equilibrium fee  $\bar{f}$  of other managers as given, is

$$\max_f \pi(f; \bar{f}) = fx_F^*(P^*(\bar{f}), f) - cx_F^*(P^*(\bar{f}), f)^2/2$$

The equilibrium fee and corresponding mutual fund performance measure are given by

**Proposition 22.** *When returns to scale are decreasing, the symmetric management fee charged by all managers is*

$$f^*(\theta) = \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)(c + \rho\sigma_\epsilon^2(1 - \gamma^2))}{c + \rho\sigma_\epsilon^2(2 - \gamma^2)}$$

$$\hat{\alpha} = \frac{\rho\sigma_\epsilon^2(1 - \gamma^2)(\alpha - \rho\gamma\sigma\sigma_\epsilon)}{c + \rho\sigma_\epsilon^2(2 - \gamma^2)}$$

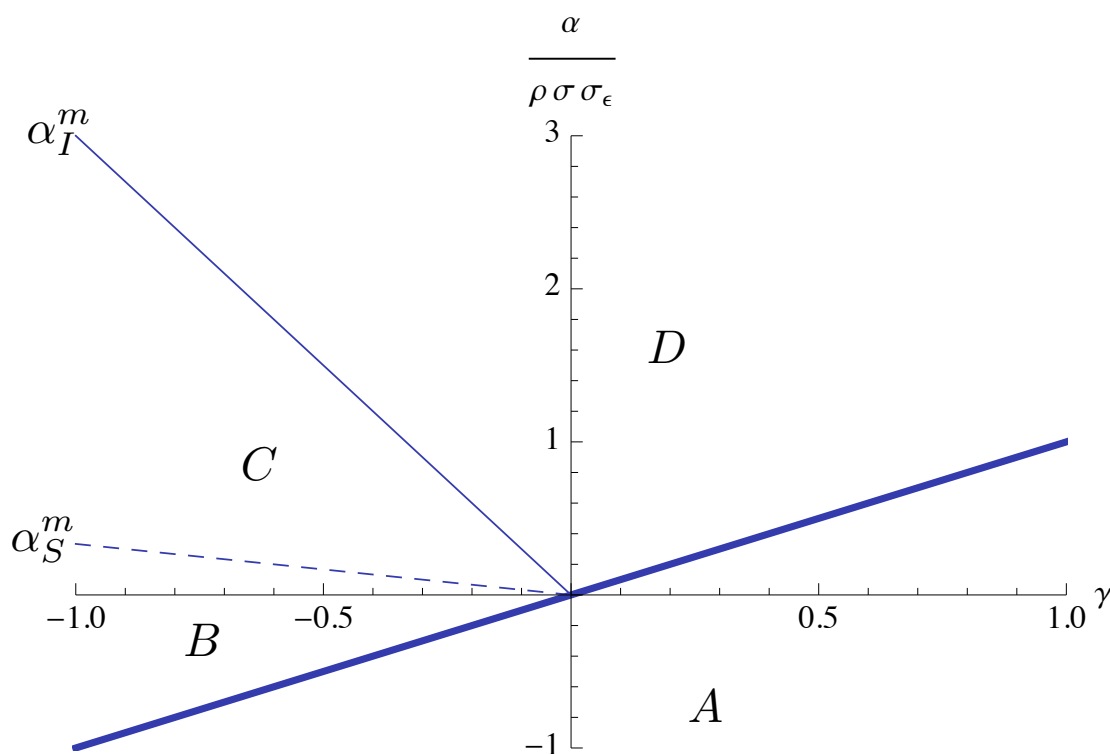


Figure 3.4: This graph shows welfare outcomes for different values of (normalized)  $\alpha$  as a function of  $\gamma$  when the fund manager is a monopolist. It corresponds to Figure 3.2 in the symmetric fee case. For  $\alpha < \alpha_0$  (region A), Assumption 4 is violated and fund management is not viable. In region B,  $\alpha < \alpha_S^m$  so fund management is viable but welfare reducing. In region C,  $\alpha \in (\alpha_S^m, \alpha_I^m)$  fund management increases total welfare but investor welfare is reduced. In region D investor welfare is also improved.

It is straightforward to show that the equilibrium fee is increasing in the cost parameter  $c$ , while measured performance is decreasing in this parameter. Managers charge higher fees to reduce demand for the fund and hence costs, and this increase in fees reduces the net return on mutual fund investment. However, measured performance is still always positive when fund management is viable (when investors do not want to short the fund).

As with the case of a monopolist manager, decreasing returns to scale shift things quantitatively but not qualitatively:

**Proposition 23.** *When fund performance exhibits decreasing returns to scale, there exist thresholds  $\alpha_I^c$  and  $\alpha_S^c$  with*

$$\alpha_I^c > \alpha_S^c > \rho\gamma\sigma\sigma_\epsilon$$

*such that the welfare results are analogous to Proposition 19. Both thresholds are increasing in cost parameter  $c$ .*

Because the main results of Proposition 19 correspond to the case where  $c = 0$ , costs simply have the effect of increasing the hurdle for fund management to be a benefit to investors and the economy as a whole, while they leave the relative relation of these hurdles unchanged.

### 3.6 Conclusion

Because delegated investment makes up a large portion of invested wealth, it potentially affects the underlying assets in an economy. Using a simple model of two aggregate assets—an underlying risky asset and a mutual fund technology—I show one potential channel for this effect: the fees and payoff characteristics of the mutual fund sector can change investor demand for the underlying risky asset.

By shifting the risky asset's price, the presence of mutual funds is not generically a good thing, and I show that standard mutual fund performance measures won't necessarily reflect this. Risk adjusted fund returns capture manager skill, but not whether mutual funds add value. This also suggests that negative observed fund performance, which French (2008) cites as a reason to eschew active management, may not be indicative of the true value added for the same reason: we don't observe the counterfactual. It's likely that other measures of value added, such as the one proposed by Berk and van Binsbergen (2011), will suffer from the same problem.

The fact that the characteristics of the mutual fund sector alter the underlying risky asset price, and hence the pricing kernel, also suggests that explanations of fund performance puzzles should account for equilibrium effects to be convincing.

# Appendix A

## Chapter 1 Proofs

*Proof of Proposition 1.* For a given bid  $B$ , let  $\mu$  be trader  $T$ 's belief about whether the asset value is high. His certainty equivalent  $\underline{u}(\mu)$  from abstaining ( $q = 0$ ) is

$$\begin{aligned}\mathbb{E}[v_t|\mu] &= v_{t-1} + (2\mu - 1)\Delta \\ \text{Var}[v_t|\mu] &= 4\mu(1 - \mu)\Delta^2 + \sigma_\epsilon^2 \\ \implies \underline{u}(\mu) &= v_{t-1} + (2\mu - 1)\Delta - \frac{\gamma}{2}[4\mu(1 - \mu)\Delta^2 + \sigma_\epsilon^2]\end{aligned}$$

First note that any equilibrium has to involve trade in the high value state  $\eta_t = \Delta$ :  $T$  is willing to trade regardless of his beliefs if  $M$ 's bid exceeds  $\underline{u}(1)$ , so  $M$  can profitably deviate from any no-trade outcome by posting this bid. Any outcome involving trade is either separating, pooling, or partially revealing (mixed).

1. **Separating Equilibrium  $\mathcal{S}^0$ :** Separation requires that  $M$  credibly reveal her information, which is only possible for bids  $B_H > v_{t-1} - \Delta$  at which trade occurs because the low type always loses money on such a bid. Hence beliefs for high bids are  $\mu([\underline{u}(1), \Delta]) = 1$  and  $M$  will extract maximum gains for these beliefs:  $B_H^* = \underline{u}(1)$ . These beliefs also need to hold for out-of-equilibrium bids in  $(-\Delta, \underline{u}(1))$ —otherwise high types have a profitable deviation by bidding on this interval, while low types do not, implying any such beliefs are inconsistent. Trade can not occur for any bid  $B_L < -\Delta$ , as high types will then have a profitable deviation posting  $B_L$  instead of  $B_H^*$ , so we can arbitrarily specify a no-trade bid  $B_L^* < \underline{u}(0)$  for which beliefs are separating. Summarizing equilibrium bids, beliefs and trading strategies:

$$B^*(\eta_t) = \begin{cases} \underline{u}(1), \eta_t = \Delta \\ < \underline{u}(0), \eta_t = -\Delta \end{cases} \quad \mu^*(B) = \begin{cases} 1, \forall B \geq \underline{u}(0) \\ 0, \text{otherwise} \end{cases} \quad q^*(B) = \begin{cases} -1, B \geq \underline{u}(1) \\ 0, \text{otherwise} \end{cases}$$

For some parameterizations, there are alternative out-of-equilibrium beliefs that support the same outcome, but the above hold generically and satisfy the “intuitive criterion” of Cho and Kreps (1987). Finally, this separating equilibrium is only feasible

if  $\underline{u}(1) > v_{t-1} - \Delta$ , as otherwise the low type can also profitably post this bid, implying these beliefs are inconsistent with separating beliefs. This implies the parameter restriction  $\gamma < \gamma_S$ , where

$$\gamma_S \equiv \frac{4\Delta}{\sigma_\epsilon^2}$$

2. **Pooling Equilibrium  $\mathcal{P}^0$ :** In a pooling equilibrium, bids are uninformative, so  $T$ 's beliefs are  $\mu = \frac{1}{2}$  for equilibrium bids.  $M$  will extract the maximum possible gains by bidding  $\underline{u}(\frac{1}{2})$  in either state:

$$B^*(\eta_t) = \underline{u}(\frac{1}{2}) \quad \mu^*(B) = \begin{cases} 1, \forall B > v_{t-1} - \Delta \\ \frac{1}{2}, \text{ otherwise} \end{cases} \quad q^*(B) = \begin{cases} -1, B \geq \underline{u}(1) \\ -1, B \in [\underline{u}(\frac{1}{2}), v_{t-1} - \Delta] \\ 0, \text{ otherwise} \end{cases}$$

Again, off-equilibrium beliefs are consistent with the intuitive criterion, and their exist inconsequential modifications to beliefs supporting the same outcome. For this equilibrium to be viable, the high type needs to be willing to post the pooling bid in both states, which is only possible if  $\underline{u}(\frac{1}{2}) \leq v_{t-1} - \Delta$  i.e.

$$\gamma \geq \gamma_P \equiv \frac{2\Delta}{\Delta^2 + \sigma_\epsilon^2}$$

3. By inspection,  $\gamma_S > \gamma_P$ .
4. Any mixed strategy equilibrium entails  $M$  signaling—posting a high bid when asset value is high—some, but not all, of the time. That implies that beliefs for any low bid  $B_L \leq v_{t-1} - \Delta$  will assign positive probability to the asset being the high type. Trade can not occur at any such bid, because the high type would always prefer to trade at  $B_L$ . But if low bids do not involve trade, the high type will strictly prefer to signal via a high bid  $B_H \geq \underline{u}(1)$  to capture some gains, implying a separating equilibrium.

□

*Proof of Corollary 1.* Note that the lower bound  $\min_\mu \underline{u}(\mu)$  provides the lowest possible bid price at which trade could occur under any belief. For example, when  $\gamma > 2/\Delta$ ,  $\underline{u}(0) > \underline{u}(\frac{1}{2})$ , and vice versa. When  $\gamma < \gamma_P$ , the pooling bid  $\underline{u}(\frac{1}{2}) > v_{t-1} - \Delta$ , so low bids  $B_L \leq v_{t-1} - \Delta$  will not result in trade, and the only feasible outcome with trade is separating, which high types will choose. When  $\gamma \geq \gamma_P$ , pooling bids are viable and will be chosen by both types. This is consistent with the beliefs because the pooling bid  $\underline{u}(\frac{1}{2}) \in [\min_\mu \underline{u}(\mu), v_{t-1} - \Delta]$ . □

*Proof of Proposition 2.* The market maker's outside option is the static separating equilibrium  $\mathcal{S}^0$  for  $\gamma < \gamma_S$ —for  $\gamma \geq \gamma_S$ , pooling is the only outcome and cannot be improved upon. Trade in a full-trade separating equilibrium occurs at both low and high bids,  $B_L \leq v_{t-1} - \Delta$

and  $B_H > v_{t-1} - \Delta$ , and can only be maintained if  $M$  will not deviate by bidding  $B_L$  when asset value is high. Let  $s_L$  and  $s_H$  be the spreads she earns on each bid; so,  $B_L = v_{t-1} - \Delta - s_L$ . Her incentive compatibility constraint is then

$$(1 - \beta)s_H + \frac{\beta}{2}(s_L + s_H) \geq (1 - \beta)(2\Delta + s_L) + \beta\frac{\gamma}{4}\sigma_\epsilon^2$$

where the last term is the discounted ex-ante gains extracted for her outside option, providing a lower bound on the high spread

$$s_H \geq \frac{s_L(2 - 3\beta) + 4(1 - \beta)\Delta + \beta\frac{\gamma}{2}\sigma_\epsilon^2}{(2 - \beta)}$$

For trade to occur, bids need to satisfy the trader's participation constraints  $B_L \geq \underline{u}(0)$  and  $B_H \geq \underline{u}(1)$ , hence both spreads  $s_L$  and  $s_H$  must be less than  $\frac{\gamma}{2}\sigma_\epsilon^2$  implying

$$s_L(2 - 3\beta) \leq (1 - \beta)(\gamma\sigma_\epsilon^2 - 4\Delta)$$

There are two cases to consider:

1.  $\beta \leq \frac{2}{3}$ : In this case,

$$s_L \leq \frac{(1 - \beta)(\gamma\sigma_\epsilon^2 - 4\Delta)}{(2 - 3\beta)}$$

but this spread has to be weakly positive, otherwise  $M$  is no better off than equilibrium  $\mathcal{S}^0$ , implying

$$\gamma \geq \frac{4\Delta}{\sigma_\epsilon^2} = \gamma_S$$

for which  $\mathcal{S}^0$  is not a credible outside option, and hence no separating improvement is possible.

2.  $\beta > \frac{2}{3}$ : In this case

$$s_L \geq \frac{(1 - \beta)(\gamma\sigma_\epsilon^2 - 4\Delta)}{(2 - 3\beta)}$$

which needs to be lower than the maximum spread  $\frac{\gamma}{2}\sigma_\epsilon^2$ . This is only possible if  $\gamma \geq \gamma_S^\infty$ , where

$$\gamma_S^\infty = \frac{8(1 - \beta)\Delta}{\beta\sigma_\epsilon^2}$$

and for any beliefs supporting separation,  $M$  maximizes profit via  $s_L^* = \frac{\gamma}{2}\sigma_\epsilon^2$ . The comparative statics of this threshold follow immediately.



Hence for  $\beta > \frac{2}{3}$ , the prices, beliefs and demands supporting the full trade separating equilibrium  $\mathcal{S}^\infty$  are

$$B^*(\eta_t; h^{t-1}) = \begin{cases} \underline{u}(1), \eta_t = \Delta \\ \underline{u}(0), \eta_t = -\Delta \end{cases} \quad \mu^*(B, h^{t-1}) = \begin{cases} 1, \forall B > v_{t-1} - \Delta \\ 0, \text{otherwise} \end{cases}$$

$$q^*(B) = \begin{cases} -1, B \geq \underline{u}(1) \\ -1, B \in [\underline{u}(0), v_{t-1} - \Delta] \\ 0, \text{otherwise} \end{cases}$$

for any history  $h^{t-1}$  which entails exclusive play of the above bidding strategy by  $M$ , and reversion to the static equilibrium  $\mathcal{S}^0$  following any period in which  $M$  deviates by bidding low when asset value is high.  $\square$

*Proof of Proposition 3.* A pooling equilibrium of the repeated game that increases trade is only possible for  $\gamma < \gamma_P$ . Above that value, the static pooling equilibrium achieves the same gains. For such  $\gamma$ ,  $\mathcal{S}^0$  is  $M$ 's outside option following any deviation.  $M$  will maintain the pooling bid, even at a loss in the low state, if future gains are high enough. For a pooling bid  $B = v_{t-1} - s$ , where  $s < \Delta$  is the spread charged,  $M$  is tempted to abstain from posting this bid in the low state, requiring that

$$(1 - \beta)(s - \Delta) + \beta s \geq (1 - \beta) \cdot 0 + \beta \frac{\gamma}{4} \sigma_\epsilon^2$$

so that the spread must satisfy

$$s \geq (1 - \beta)\Delta + \beta \frac{\gamma}{4} \sigma_\epsilon^2$$

Trade in any pooling equilibrium requires a bid  $B \geq \underline{u}(\frac{1}{2})$ , placing an upper bound on the spread:

$$(1 - \beta)\Delta + \beta \frac{\gamma}{4} \sigma_\epsilon^2 \leq \frac{\gamma}{2} (\Delta^2 + \sigma_\epsilon^2)$$

This condition is only satisfied if  $\gamma \geq \gamma_P^\infty$ , where

$$\gamma_P^\infty = \frac{4(1 - \beta)\Delta}{2\Delta^2 + (2 - \beta)\sigma_\epsilon^2}$$

It is easy to show  $\gamma_P^\infty \leq \gamma_P$ . A full specification of the actions and beliefs supporting this equilibrium ( $\mathcal{P}^\infty$ ) is given by

$$B^*(\eta_t; h^{t-1}) = \underline{u}(\frac{1}{2}) \quad \mu^*(B, h^{t-1}) = \frac{1}{2} \quad q^*(B) = \begin{cases} -1, B \geq \underline{u}(\frac{1}{2}) \\ 0, \text{otherwise} \end{cases}$$

for any history  $h^{t-1}$  which entails exclusive play of the above bidding strategy by  $M$ , and reversion to the static equilibrium  $\mathcal{S}^0$  following any period in which  $M$  deviates by bidding low when asset value is high. The specified beliefs do not necessarily survive refinements, but any more nuanced specification results in the same outcome.  $\square$

*Proof of Lemma 3.* The fact that  $\gamma_S^\infty > \gamma_P^\infty$  for all parameterizations follows from inspection of their definitions in Propositions 2 and 3. Comparing  $\gamma_S^\infty$  to the static pooling threshold  $\gamma_P$  defined in Proposition 1 produces the lower bound  $\underline{\beta}$ :

$$\underline{\beta} = \frac{4(\Delta^2 + \sigma_\epsilon^2)}{4\Delta^2 + 5\sigma_\epsilon^2}$$

□

*Proof of Proposition 18.* First consider the consumer surplus: in any pooling equilibrium,  $T$  receives  $\underline{u}(\frac{1}{2})$  for the asset, which is equal to his autarky certainty equivalent, so the consumer surplus is zero. For either of the separating equilibria,  $T$ 's ex-ante gain over autarky ( $\underline{u}(1/2)$ ) is

$$\mathbb{E}[\underline{u}(\mu(B))] - \frac{\gamma}{2} \mathbf{Var}(\underline{u}(\mu(B))) + L(\Delta^2) - \underline{u}(1/2) = L(\Delta^2)$$

In the pooling equilibrium, full gains are extracted by  $M$  for a surplus of  $\frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2)$ . In the full-trade separating equilibrium  $\mathcal{S}^\infty$ ,  $M$  extracts the residual risk sharing gains  $\frac{\gamma}{2}\sigma_\epsilon^2$  in both states of the world, while in the partial trade separating equilibrium  $\mathcal{S}^0$ , residual risk sharing gains of  $\frac{\gamma}{2}\sigma_\epsilon^2$  are lost in the low-state, which occurs half of the time, reducing ex-ante gains accordingly. □

*Proof of Proposition 5.* Negative endowment shocks  $y_t = -1$ , which are equally likely, lead to purchases at ask prices symmetric to the bids in the previous propositions, implying price processes for the various equilibria. Letting  $\Sigma$  denote the variance of price changes for each of them:

1. *Pooling equilibria  $\mathcal{P}^0, \mathcal{P}^\infty$ :* Prices are given by

$$P_t = v_{t-1} - y_t \frac{\gamma}{2} (\Delta^2 + \sigma_\epsilon^2)$$

implying price differences

$$P_t - P_{t-1} = \eta_{t-1} + \epsilon_{t-1} - (y_t - y_{t-1}) \frac{\gamma}{2} (\Delta^2 + \sigma_\epsilon^2)$$

All of the innovations are independent of each other, implying

$$\Sigma_{\mathcal{P}} = (\Delta^2 + \sigma_\epsilon^2) \left[ 1 + \frac{\gamma^2}{2} (\Delta^2 + \sigma_\epsilon^2) \right]$$

2. *Full-trade separating equilibrium  $\mathcal{S}^\infty$ :* Similarly, again using the independence of innovations, prices and changes are given by

$$P_t = v_{t-1} + \eta_t - y_t \frac{\gamma}{2} \sigma_\epsilon^2$$

$$P_t - P_{t-1} = \eta_{t-1} + \epsilon_{t-1} + (\eta_t - \eta_{t-1}) - (y_t - y_{t-1}) \frac{\gamma}{2} \sigma_\epsilon^2$$

$$\Sigma_{\mathcal{S}^\infty} = \Delta^2 + \sigma_\epsilon^2 + \frac{\gamma^2}{2} \sigma_\epsilon^4$$

3. *Partial-trade separating equilibrium*  $\mathcal{S}^0$ : For this equilibrium, trade does not occur in some states of the world, so there are periods where prices are not observed. Let  $\{\tau_n\}_{n=1}^\infty$  denote the sequence of realized trade times. Prices are then given by:

$$P_{\tau_n} = v_{\tau_{n-1}} + \eta_{\tau_n} - y_{\tau_n} \frac{\gamma}{2} \sigma_\epsilon^2$$

with differences

$$\begin{aligned} P_{\tau_n} - P_{\tau_{n-1}} &= v_{\tau_{n-1}} - v_{\tau_{n-1}-1} + (\eta_{\tau_n} - \eta_{\tau_{n-1}}) - (y_{\tau_n} - y_{\tau_{n-1}}) \frac{\gamma}{2} \sigma_\epsilon^2 \\ &= \sum_{s=\tau_{n-1}}^{\tau_n-1} (\eta_s + \epsilon_s) + (\eta_{\tau_n} - \eta_{\tau_{n-1}}) - (y_{\tau_n} - y_{\tau_{n-1}}) \frac{\gamma}{2} \sigma_\epsilon^2 \end{aligned}$$

These differences still have mean zero, so the variance is given by

$$\begin{aligned} \Sigma_{\mathcal{S}^0} &= \mathbb{E} \left\{ \sum_{s=\tau_{n-1}}^{\tau_n-1} (\eta_s + \epsilon_s) \right\}^2 + 2\mathbb{E} \left\{ (\eta_{\tau_n} - \eta_{\tau_{n-1}}) - (y_{\tau_n} - y_{\tau_{n-1}}) \frac{\gamma}{2} \sigma_\epsilon^2 \right\} \sum_{s=\tau_{n-1}}^{\tau_n-1} (\eta_s + \epsilon_s) \\ &\quad + \mathbb{E} \left\{ (\eta_{\tau_n} - \eta_{\tau_{n-1}}) - (y_{\tau_n} - y_{\tau_{n-1}}) \frac{\gamma}{2} \sigma_\epsilon^2 \right\}^2 \end{aligned}$$

It's easiest to calculate these terms separately. Conditional on being in a trade state at time  $t$ , which occurs when  $(\eta_t, y_t) \in \{(\Delta, 1), (-\Delta, -1)\}$ , let  $\mathcal{T}$  be the stopping time at which the next trade occurs. Trade occurs with probability 1/2 independent of the current state, so  $\mathbb{E}\mathcal{T} = 2$  (this is a trivially stationary, irreducible, finite-state Markov chain). Then

$$\begin{aligned} \mathbb{E} \left\{ \sum_{s=\tau_{n-1}}^{\tau_n-1} (\eta_s + \epsilon_s) \right\}^2 &= \mathbb{E} \sum_{j=1}^{\infty} \mathbf{1}_{\mathcal{T}=j} \mathbb{E} \left\{ \sum_{s=\tau_{n-1}}^{\tau_{n-1}+j} (\eta_s + \epsilon_s) \right\}^2 \\ &= \mathbb{E} \sum_{j=1}^{\infty} \mathbf{1}_{\mathcal{T}=j} \sum_{s=\tau_{n-1}}^{\tau_{n-1}+j} \mathbb{E} (\eta_s + \epsilon_s)^2 \\ &= \mathbb{E} \sum_{j=1}^{\infty} \mathbf{1}_{\mathcal{T}=j} \sum_{s=\tau_{n-1}}^{\tau_{n-1}+j} (\Delta^2 + \sigma_\epsilon^2) \\ &= (\Delta^2 + \sigma_\epsilon^2) \mathbb{E} \sum_{j=1}^{\infty} j \cdot \mathbf{1}_{\mathcal{T}=j} \\ &= (\Delta^2 + \sigma_\epsilon^2) \mathbb{E}\mathcal{T} \\ &= 2(\Delta^2 + \sigma_\epsilon^2) \end{aligned}$$

The second term is

$$\begin{aligned} 2\mathbb{E} \left\{ (\eta_{\tau_n} - \eta_{\tau_{n-1}}) - (y_{\tau_n} - y_{\tau_{n-1}}) \frac{\gamma}{2} \sigma_\epsilon^2 \right\} \sum_{s=\tau_{n-1}}^{\tau_n-1} (\eta_s + \epsilon_s) &= 2\mathbb{E} \left( -\eta_{\tau_{n-1}}^2 + \eta_{\tau_{n-1}} y_{\tau_{n-1}} \frac{\gamma}{2} \sigma_\epsilon^2 \right) \\ &= 2(-\Delta^2 + \Delta \frac{\gamma}{2} \sigma_\epsilon^2) \\ &= \Delta \gamma \sigma_\epsilon^2 - 2\Delta^2 \end{aligned}$$

The third term is

$$\begin{aligned} \mathbb{E} \left\{ (\eta_{\tau_n} - \eta_{\tau_{n-1}}) - (y_{\tau_n} - y_{\tau_{n-1}}) \frac{\gamma}{2} \sigma_\epsilon^2 \right\}^2 &= \mathbb{E}(\eta_{\tau_n} - \eta_{\tau_{n-1}})^2 - \gamma \sigma_\epsilon^2 \mathbb{E}(\eta_{\tau_n} - \eta_{\tau_{n-1}})(y_{\tau_n} - y_{\tau_{n-1}}) \\ &\quad + \frac{\gamma^2}{4} \sigma_\epsilon^4 \mathbb{E}(y_{\tau_n} - y_{\tau_{n-1}})^2 \\ &= 2\Delta^2 - 2\Delta \gamma \sigma_\epsilon^2 + \frac{\gamma^2}{2} \sigma_\epsilon^4 \\ &= \frac{1}{2} (2\Delta - \gamma \sigma_\epsilon^2)^2 \end{aligned}$$

where we've used the fact that  $\mathbb{E}\eta_{\tau_n} y_{\tau_n} = \Delta$  conditional on trade occurring. Combining these three terms:

$$\begin{aligned} \Sigma_{S^0} &= 2(\Delta^2 + \sigma_\epsilon^2) + \Delta \gamma \sigma_\epsilon^2 - 2\Delta^2 + \frac{1}{2} (2\Delta - \gamma \sigma_\epsilon^2)^2 \\ &= 2(\Delta^2 + \sigma_\epsilon^2) + \frac{\gamma^2}{2} \sigma_\epsilon^4 - \Delta \gamma \sigma_\epsilon^2 \end{aligned}$$

A similar calculation of the first-order auto-covariance  $\phi_1$  produces:

$$\begin{aligned} \mathcal{P} : \phi_1 &= -\frac{\gamma^2}{4} (\Delta^2 + \sigma_\epsilon^2)^2 \\ \mathcal{S}^\infty : \phi_1 &= -\frac{\gamma^2}{4} \sigma_\epsilon^4 \\ \mathcal{S}^0 : \phi_1 &= \frac{\gamma}{2} \Delta \sigma_\epsilon^2 - \frac{\gamma^2}{4} \sigma_\epsilon^4 \end{aligned}$$

of each price series identifies the permanent component as  $\Sigma + 2\phi_1$  and the transitory component as  $\lambda = -2\phi_1$ .  $\square$

*Proof of Proposition 6.* Consider the bid-side—this is without loss of generality—and a price-quantity offer  $(B, Q)$ , where  $Q$  is the quantity  $M$  is willing to buy at price  $B$  per unit (the total transfer). This is equivalent to a bid-schedule  $B(q)$ , where  $B(q)$  is an unacceptably low bid for  $q \neq Q$ .<sup>1</sup> For a given price-schedule  $B(q)$ , let  $\mu$  be  $T$ 's belief that the asset value is high. His certainty equivalent by selling  $q$  units  $\underline{u}(\mu, q)$  is

$$\underline{u}(\mu, q) = (1 - q)[v_{t-1} + (2\mu - 1)\Delta] + qB(q) - \frac{\gamma}{2}(1 - q)^2[4\mu(1 - \mu)\Delta^2 + \sigma_\epsilon^2]$$

<sup>1</sup>I'm abusing notation here a bit by using  $q > 0$  to indicate the quantity sold; elsewhere, I've used  $q < 0$ .

and his certainty equivalent from abstaining is then  $\underline{u}(\mu, 0)$ . Note, as in the main case, autarky is not an equilibrium:  $M$  can always offer the bid-quantity pair  $(B, Q) = (\underline{u}(1, 1), 1)$  and  $T$  will accept regardless of his beliefs.

1. **Pooling Equilibrium  $\mathcal{P}^0$ :** As in the binary ( $q \in \{0, 1\}$ ) case, risk sharing gains are maximized at  $q = 1$ , and  $M$  can extract all of these gains by offering  $\underline{u}(\frac{1}{2}, 1)$  under the pooling belief  $\mu = \frac{1}{2}$ , and a full specification of actions and beliefs supporting this outcome is

$$(B^*, Q^*) = \left( \underline{u}\left(\frac{1}{2}, 1\right), 1 \right) \quad \mu^*(B, Q) = \begin{cases} 1, \forall(B, Q) : B > v_{t-1} - \Delta \\ \frac{1}{2}, \text{ otherwise} \end{cases}$$

$$q^*(B, Q) = \begin{cases} Q, \forall(B, Q) : B \geq v_{t-1} + \Delta - (2 - Q)\frac{\gamma}{2}\sigma_\epsilon^2 \\ Q, \forall(B, Q) : B \in [v_{t-1} - \Delta - (2 - Q)\frac{\gamma}{2}(\Delta^2 + \sigma_\epsilon^2), v_{t-1} - \Delta] \\ 0, \text{ otherwise} \end{cases}$$

where the off-equilibrium beliefs and actions satisfy the intuitive criterion. For the full insurance pooling equilibrium to be viable, the pooling bid needs to be profitable in both states ( $\underline{u}(\frac{1}{2}, 1) \leq v_{t-1} - \Delta$ ), requiring that  $\gamma \geq \hat{\gamma}_P$

$$\hat{\gamma}_P \equiv \frac{2\Delta}{\Delta^2 + \sigma_\epsilon^2}$$

2. **Separating Equilibrium  $\mathcal{S}^0$ :** The high state mirrors that of Proposition 1. Let  $(B_H, Q_H)$  be the price-quantity pair offered in the high value state.  $B_H > v_{t-1} - \Delta$  signals asset value is high, and the market maker extracts all gains in this state when  $(B_H^*, Q_H^*) = (\underline{u}(1, 0), 1)$ . In the low value state, any bid-quantity pair needs to satisfy two constraints: First, an IR constraint that ensures trader  $T$  gets at least his outside option,  $\underline{u}(0, q) \geq \underline{u}(0, 0)$ , implying

$$B_L(q) \geq v_{t-1} - \Delta - (2 - q)\frac{\gamma}{2}\sigma_\epsilon^2 \quad (IR_q)$$

Second, an IC constraint ensuring that the high-type market maker is not tempted to deviate by quoting the low bid-quantity pair. His profit in the high state is  $\pi_H^* = \frac{\gamma}{2}\sigma_\epsilon^2$ , the residual risk sharing gains, so bid-quantity pairs have to satisfy

$$\begin{aligned} \pi_H &\geq q(v_{t-1} - \Delta - B_L(q)) \\ \implies B_L(q) &\geq v_{t-1} + \Delta - \frac{\gamma}{2q}\sigma_\epsilon^2 \quad (IC_q) \end{aligned}$$

For a given quantity  $q$ ,  $M$  maximizes profits when she bids the minimum amount satisfying both constraints. In the low state, her profits  $\pi_L(q)$  for a given  $q$  are therefore

$$\pi_L(q) = q \left( v_{t-1} - \Delta - \max \left\{ v_{t-1} - \Delta - (2 - q)\frac{\gamma}{2}\sigma_\epsilon^2, v_{t-1} + \Delta - \frac{\gamma}{2q}\sigma_\epsilon^2 \right\} \right)$$

For small  $q$ , the  $IR_q$  constraint binds. It's straightforward to show that  $\pi_L(q)$  is increasing on  $q$  when this is the case, so  $M$  increases her offer. Solving for the  $q$  that equalizes the constraints shows that there is a single crossing point at

$$q^* = 1 - \frac{\sqrt{\Delta^2 + \Delta\gamma\sigma_\epsilon^2} - \Delta}{\frac{\gamma}{2}\sigma_\epsilon^2} \in (0, 1)$$

after which  $IC_q$  determines the bid.  $\pi_L(q)$  is decreasing as  $q$  increases for this case, so  $M$  will optimally bid  $B_L(q^*)$ . Define the set of incentive compatible bids  $\mathcal{B}$  as

$$\mathcal{B} = \{(B, Q) : B \geq v_{t-1} + \Delta - \frac{\gamma}{2Q}\sigma_\epsilon^2\}$$

A full equilibrium specification is given by

$$\begin{aligned} (B^*, Q^*)(\eta_t) &= \begin{cases} (\underline{u}(1, 0), 1), \eta_t = \Delta \\ (v_{t-1} - \Delta - (2 - q^*)\frac{\gamma}{2}\sigma_\epsilon^2, q^*), \eta_t = -\Delta \end{cases} \\ \mu^*(B, Q) &= \begin{cases} 1, & \forall (B, Q) : B > v_{t-1} - \Delta \cup (B, Q) \notin \mathcal{B} \\ 0, & \forall (B, Q) : B \leq v_{t-1} - \Delta \cap (B, Q) \in \mathcal{B} \end{cases} \\ q^*(B, Q) &= \begin{cases} Q, \forall (B, Q) : B \geq v_{t-1} + \Delta - (2 - Q)\frac{\gamma}{2}\sigma_\epsilon^2 \\ Q, \forall (B, Q) : B \leq v_{t-1} - \Delta \cap (B, Q) \in \mathcal{B} \\ 0, \text{ otherwise} \end{cases} \end{aligned}$$

For some parameterizations, there are more complex out-of-equilibrium belief specifications that support the same outcome—they do not change prices and quantities. Unlike in the binary case, this separating equilibrium always exists: the low type  $M$  is never tempted to post the high-type bid, even if the bid  $B_H < v_{t-1} - \Delta$  so that is profitable.

□

*Proof of Proposition 7.* Most of the logic here mirrors Proposition 1, so I omit details on belief specifications, etc. Let  $\mu$  be trader  $T$ 's belief about whether the asset value is high. His certainty equivalents  $\underline{u}(\mu)$  and  $\bar{u}(\mu)$  from abstaining ( $q = 0$ ) following positive and negative endowment shocks, respectively, are

$$\begin{aligned} \mathbb{E}[v_t|\mu] &= v_{t-1} + \Delta_L + p(\Delta_H - \Delta_L) \\ \mathbf{Var}[v_t|\mu] &= \mu(1 - \mu)(\Delta_H - \Delta_L)^2 + \sigma_\epsilon^2 \\ \implies \underline{u}(\mu) &= v_{t-1} + \Delta_L + p(\Delta_H - \Delta_L) - \frac{\gamma}{2}[\mu(1 - \mu)(\Delta_H - \Delta_L)^2 + \sigma_\epsilon^2] \\ \bar{u}(\mu) &= -[v_{t-1} + \Delta_L + p(\Delta_H - \Delta_L)] - \frac{\gamma}{2}[\mu(1 - \mu)(\Delta_H - \Delta_L)^2 + \sigma_\epsilon^2] \end{aligned}$$

1. **Separating Equilibrium  $\mathcal{S}^0$** : Separation requires that  $M$  credibly reveal her information, which is only possible for bids  $B_H > v_{t-1} + \Delta_L$  and asks  $A_L < v_{t-1} + \Delta_H$ . Because the other types lose money at these quotes, beliefs are  $\mu(\underline{u}(1), \Delta_H) = 1$  for high bids and  $\mu(\bar{u}(0), \Delta_H) = 0$  for low asks.  $M$  will extract maximum gains for these beliefs:  $B_H^* = \underline{u}(1)$ ,  $A_L^* = -\bar{u}(0)$ . Trade can not occur in any separating equilibrium for low bids  $B_L < v_{t-1} + \Delta_L$  or high asks  $A_H > v_{t-1} + \Delta_H$ .

As before, the separating bid quotes are only feasible if  $\underline{u}(1) > v_{t-1} + \Delta_L$ . Similarly, separating ask quotes are only feasible  $A_L^* < v_{t-1} + \Delta_H$ . Each of these constraints implies the same parameter restriction  $\gamma < \gamma_S$ , where

$$\gamma_S \equiv \frac{2(\Delta_H - \Delta_L)}{\sigma_\epsilon^2}$$

2. **Bid-side Pooling Equilibrium  $\mathcal{P}_B^0$** : For  $M$  to commit to a pooling bid in either state, it is necessary that  $\underline{u}(p) \leq v_{t-1} + \Delta_L$ , requiring that  $\gamma \geq \gamma_{P,B}$ , where

$$\gamma_{P,B} \equiv \frac{2p(\Delta_H - \Delta_L)}{\sigma_\epsilon^2 + p(1-p)(\Delta_H - \Delta_L)^2}$$

3. **Ask-side Pooling Equilibrium  $\mathcal{P}_A^0$** : Similarly, on the ask-side, pooling requires  $-\bar{u}(p) \geq v_{t-1} + \Delta_H$ , implying that  $\gamma \geq \gamma_{P,A}$ , where

$$\gamma_{P,A} \equiv \frac{2(1-p)(\Delta_H - \Delta_L)}{\sigma_\epsilon^2 + p(1-p)(\Delta_H - \Delta_L)^2}$$

4. When  $\gamma$  exceeds both pooling thresholds, then obviously pooling on both sides of the market is feasible.
5. Follows from inspection of the thresholds.
6. Follows from inspection of the thresholds.
7. The logic is the same as Proposition 1: any mixed equilibrium implies partially revealing prices in some states of the world, but for induced under any of those prices,  $M$  will deviate from his mixed strategy.

□

*Proof of Proposition 8.* The logic follows Propositions 2-3 and Lemma 2, so many details are omitted. First consider the separating improvements on the bid-side: the easiest improvement to support gives  $M$  all of the residual risk sharing gains on every trade. His tempted deviation is to bid low when the asset has a high value, which triggers the static separating outcome, requiring

$$(1 - \beta) \frac{\gamma}{2} \sigma_\epsilon^2 + \beta \frac{\gamma}{2} \sigma_\epsilon^2 \geq (1 - \beta) (\Delta_H - \Delta_L + \frac{\gamma}{2} \sigma_\epsilon^2) + \beta \frac{\gamma}{4} \sigma_\epsilon^2$$

implying that  $\gamma \geq \gamma_S^\infty$ , where

$$\gamma_S^\infty \equiv \frac{4(1-\beta)(\Delta_H - \Delta_L)}{\beta\sigma_\epsilon^2}$$

Note the constraint is identical on the ask-side, as the temptation to deviate is independent of the high-state probability  $p$ , so the separating threshold is the same.

For a pooling bid, let  $s$  denote the spread between the trader's prior expectation and the bid:  $B = \mathbb{E}[v_t|p] - s$ . In a pooling equilibrium,  $s$  is the market maker's ex-ante expected per-period profit, but she loses money on this spread in the low-state when  $\gamma < \gamma_{P,B}$ . She is tempted to deviate by abstaining from trade in this state because she knows the asset value is low:

$$\begin{aligned} (1-\beta)(v_{t-1} + \Delta_L - B) + \beta s &\geq (1-\beta) \cdot 0 + \beta \frac{\gamma}{4} \sigma_\epsilon^2 \\ \Rightarrow s &\geq (1-\beta)p(\Delta_H - \Delta_L) + \beta \frac{\gamma}{4} \sigma_\epsilon^2 \end{aligned}$$

In order for trade to occur, the bid has to exceed the trader's certainty equivalent given his prior:  $B \geq \underline{u}(p)$ , implying that

$$s \leq \frac{\gamma}{2} [\mu(1-\mu)(\Delta_H - \Delta_L)^2 + \sigma_\epsilon^2]$$

which combined with the previous IC constraint requires that  $\gamma \geq \gamma_{P,B}^\infty$ , where

$$\gamma_{P,B}^\infty \equiv \frac{4(1-\beta)p(\Delta_H - \Delta_L)}{(2-\beta)\sigma_\epsilon^2 + 2p(1-p)(\Delta_H - \Delta_L)^2}$$

The logic on the ask side is symmetric with  $p \rightarrow (1-p)$ , so

$$\gamma_{P,A}^\infty \equiv \frac{4(1-\beta)(1-p)(\Delta_H - \Delta_L)}{(2-\beta)\sigma_\epsilon^2 + 2p(1-p)(\Delta_H - \Delta_L)^2}$$

That both of these pooling thresholds are lower than  $\gamma_S^\infty$  follows by inspection.  $\square$



# Appendix B

## Chapter 2 Proofs

*Proof of Proposition 9.* First note that it must be the case that  $s^*(0) = s^*(1)$  – trader  $i$  can not credibly signal her trading motive and will always choose the narrowest spread – so  $m$  is choosing a single spread, call it  $s$ . Suppose  $\delta < \Delta$ . First note that for any  $s \in (\delta, \Delta)$ ,  $i$ 's best response is to trade when informed ( $\hat{q}(\eta = 1) = 1$ ) and abstain when liquidity driven ( $\hat{q}(\eta = 0) = 0$ ), implying  $m$ 's expected payoff is  $V_t^m = (1 - \alpha)0 + \alpha(s - \Delta) < 0$ . If  $s \geq \Delta$ ,  $i$ 's best response is to abstain from trade in both states ( $\hat{q}(\eta = 0) = \hat{q}(\eta = 1) = 0$ ), implying an expected payoff of  $V_t^m = 0$ , which clearly dominates  $s \in (\delta, \Delta)$ . For  $s \in (0, \delta)$ ,  $i$ 's best response is to always trade ( $\hat{q}(\eta = 0) = \hat{q}(\eta = 1) = 1$ ), with expected payoffs  $V_t^m = s - \alpha\Delta$  and  $V_t^i = (1 - \alpha)\delta + \alpha\Delta - s$ .

When  $\delta < \alpha\Delta$ , any  $s \in (0, \delta)$  leads to a negative expected payoff for  $m$ , and is dominated by the zero-payoff actions  $s \geq \Delta$ , result (1). When  $\delta \in (\alpha\Delta, \Delta)$ ,  $m$ 's most profitable action is to set  $s = \delta$  for an expected payoff profile  $V_t^m = \delta - \alpha\Delta > 0$ ,  $V_t^i = \alpha(\Delta - \delta) > 0$ , implying result (2).

When  $\delta > \Delta$ , any  $s \in (\Delta, \delta)$  allows  $m$  to screen out losses from informed trade while earning positive profits in the hedging state:  $V_t^m = (1 - \alpha)s$ . Any  $s > \delta$  leads to no trade, so  $m$ 's profits are maximized at  $s^* = \delta$ . Trader  $i$  abstains from trade in the information state ( $\hat{q}(\eta = 1) = 0$ ), and trades (indifferently) in the liquidity state ( $\hat{q}(\eta = 0) = 1$ ), for a payoff of  $V_t^i = 0$ , result (3).  $\square$

*Proof of Proposition 10.* We'll ignore the  $IR$  constraint and verify it holds later. Because the objective function is linear, we are simply looking for the maximum  $s_L$  and minimum  $s_I$  such that the  $IC$  constraints are satisfied. Assume both  $IC$  constraints bind: this occurs when  $\Delta - s_I > s_L$  and  $s_L > \delta$ . Solving the  $IC$ 's simultaneously produces the resulting spreads  $s_L^*$  and  $s_I^*$ . Now it remains to verify that there are not better solutions where either constraint binds (at least one has to bind at any optimum). Suppose only  $IC_1$  binds – then  $m$  will choose spreads such that the  $IR$  binds, producing  $s_L^* = \delta + \alpha(\Delta - \delta)$  and  $s_I^* = (1 - \alpha)(\Delta - \delta)$ . But checking  $IC_2$  shows it is violated, a contradiction. A similar argument shows it can't be

the case that only  $IC_2$  binds. Value functions are given by:

$$\begin{aligned} V_i^1 &= \frac{\alpha(1-\beta)(\Delta-\delta)}{1+\alpha\beta} \\ V_m^1 &= (1-\alpha)\delta - V_i^1 \\ &= \frac{(1-\alpha^2\beta)\delta - \alpha(1-\beta)\Delta}{1+\alpha\beta} \end{aligned}$$

To be profitable, it must be the case that

$$V_m^1 \geq 0 \iff \delta > \underline{\delta}$$

which defines the parameter region for which gains from trade can be achieved via repeated play.  $\square$

*Proof of Proposition 11.* Regressing price changes  $p_t - p_{t-1}$  on  $x_t$  and  $x_{t-1}$ , using the independence of  $x_t$  and  $x_{t-1}$ , produces estimates

$$\begin{aligned} (\hat{\lambda} + \hat{c}) &= \frac{\mathbf{Cov}(p_t - p_{t-1}, x_t)}{\mathbf{Var}(x_t)} \\ -\hat{c} &= \frac{\mathbf{Cov}(p_t - p_{t-1}, x_{t-1})}{\mathbf{Var}(x_t)} \end{aligned}$$

The observed price in the reduced form specification  $p_t$  corresponds to the true model price  $P_t$ , while the order flow indicator  $x_t$  corresponds to the true model order flow  $y_t$ . Order flow  $y_t$  is a symmetric binary variable, so its variance is 1. The true evolution of prices changes as a function of the primitive shocks is

$$\begin{aligned} P_t - P_{t-1} &= v_{t-1} - v_{t-2} + (\Delta - s_I^*)(\eta_t y_t - \eta_{t-1} y_{t-1}) + s_L^*[(1 - \eta_t)y_t - (1 - \eta_{t-1})y_{t-1}] \\ &= (\Delta - s_I^*)\eta_t y_t + s_L^*(1 - \eta_t)y_t + s_I^*\eta_{t-1}y_{t-1} - s_L^*(1 - \eta_{t-1})y_{t-1} \end{aligned}$$

Because both price changes and  $y_t$  have a mean of zero, their covariances are equal to their cross moments.  $y_t$  and  $y_{t-1}$  are only related to contemporary terms because everything in the model is i.i.d., so

$$\begin{aligned} (\hat{\lambda} + \hat{c}) &= \mathbb{E}(P_t - P_{t-1})y_t = (\Delta - s_I^*)\mathbb{E}\eta_t y_t^2 + s_L^*\mathbb{E}(1 - \eta_t)y_t^2 \\ -\hat{c} &= \mathbb{E}(P_t - P_{t-1})y_{t-1} = s_I^*\mathbb{E}\eta_{t-1}y_{t-1}^2 - s_L^*\mathbb{E}(1 - \eta_{t-1})y_{t-1}^2 \end{aligned}$$

$y_t$  and  $\eta_t$  are both independent binary variables, and it's straightforward to show that

$$\mathbb{E}\eta_t y_t^2 = \alpha \quad \mathbb{E}(1 - \eta_t)y_t^2 = 1 - \alpha$$

so the regression coefficients are

$$\begin{aligned} (\hat{\lambda} + \hat{c}) &= \alpha(\Delta - s_I^*) + (1 - \alpha)s_L^* \\ -\hat{c} &= \alpha s_I^* - s_L^*(1 - \alpha) \end{aligned}$$

Solving for  $\hat{\lambda}$  and  $\hat{c}$  produces the estimates in the proposition. For the comparative statics, substitute in the equilibrium spreads and differentiate:

$$\begin{aligned}\hat{c} &= (1 - \alpha)s_L^* - \alpha s_I^* \\ &= \frac{\delta(1 - \alpha^2\beta) - \alpha\Delta(1 - \beta)}{1 + \alpha\beta} \\ \frac{\partial \hat{c}}{\partial \alpha} &= -\frac{\beta[1 + \alpha(2 + \alpha\beta)]\delta + (1 - \beta)\Delta}{1 + \alpha\beta} < 0 \\ \frac{\partial \hat{c}}{\partial \Delta} &= \frac{-\alpha(1 - \beta)}{1 + \alpha\beta} < 0\end{aligned}$$

□

*Proof of Proposition 12.* First note that a given  $(m, i)$  pairing,  $m$  will solve the same problem as a monopolist, with an adjustment for  $i$ 's improved outside option. Letting  $V_i^0$  be the continuation value of a trader entering the matching pool,  $m$ 's problem is

$$\max_{s_L, s_I} (1 - \alpha)s_L - \alpha s_I \quad (P')$$

such that:

$$(1 - \beta)s_I + \beta[\alpha s_I + (1 - \alpha)(\delta - s_L)] \geq (1 - \beta)(\Delta - s_L) + \beta \cdot V_i^0 \quad (IC_1)$$

$$(1 - \beta)(\delta - s_L) + \beta[\alpha s_I + (1 - \alpha)(\delta - s_L)] \geq \beta V_i^0 \quad (IC_2)$$

$$\alpha s_I + (1 - \alpha)(\delta - s_L) \geq 0 \quad (IR)$$

As in proposition 10, both  $IC$  constraints bind at the optimum, and solving for them taking  $V_i^0$  as exogenous gives

$$s_L^* = \frac{\delta + \alpha\beta - \beta V_i^0}{1 + \alpha\beta} \quad s_I^* = \frac{(1 - \alpha\beta)(\Delta - \delta) + 2\beta V_i^0}{1 + \alpha\beta}$$

Let  $V_i^\mu$  be the continuation value of a trader in a relationship where spreads  $s_I^*$  and  $s_L^*$  are being charged on the two types of trades, and cooperation is maintained:

$$V_i^\mu = (1 - \alpha)(\delta - s_L^*) + \alpha s_I^*$$

For an unmatched trader at the start of a period, if  $p$  is the probability a trader in the pool is matched with a dealer, his ex-ante utility is

$$\begin{aligned}V_i^0 &= pV_i^\mu + (1 - p)\beta V_i^0 \\ \implies V_i^0 &= \frac{p}{1 - (1 - p)\beta} V_i^\mu\end{aligned}$$

because with probability  $(1 - p)$  he rolls into next period's matching pool. To derive the matching probability, note that the probability of a match at the start of the period is

$$p = \frac{\text{mass of market makers in the pool}}{\text{mass of traders in the pool}}$$

New market makers only enter the pool as a replacement for those that died the previous period, which have mass  $(1 - \beta)\mu$ . Any unmatched traders from the previous period (or their replacements) add a mass of  $(1 - \mu)$  to this periods pool. Any traders replacing those from dead relationships in the previous period result in an additional  $(1 - \beta)\mu$  traders in the pool, so that

$$p = \frac{(1 - \beta)\mu}{1 - \mu + (1 - \beta)\mu} = \frac{(1 - \beta)\mu}{1 - \beta\mu}$$

Substituting this and  $V_i^\mu$  gives us the value of  $V_i^0$  as a function of the spreads,

$$\begin{aligned} V_i^0 &= \frac{p}{1 - (1 - p)\beta} V_i^\mu \\ &= \mu V_i^\mu \\ &= (1 - \alpha)(\delta - s_L^*) + \alpha s_I^* \end{aligned}$$

In any equilibrium, this endogenous outside option needs to be consistent with the optimal spreads being charged, so plugging  $V_i^0$  back into the expressions for  $s_I^*$  and  $s_L^*$  produces the equilibrium spreads as a function of the primitives:

$$s_L^* = \frac{\delta + \alpha\beta - \beta\mu(\delta + \alpha\Delta)}{1 + \beta[\alpha - (1 + \alpha)\mu]} \quad s_I^* = \frac{[1 - \alpha\beta - \beta\mu(1 - \alpha)](\Delta - \delta)}{1 + \beta[\alpha - (1 + \alpha)\mu]}$$

Note that these match the monopolist case when  $\mu = 0$ , as expected.  $\square$

*Proof of Proposition 13.* All that's changed in this case are the spreads, as a function of the outside option due to market maker concentration  $\mu$ , so transaction cost estimate  $\hat{c}$  becomes

$$\begin{aligned} \hat{c} &= (1 - \alpha)s_L^* - \alpha s_I^* \\ &= \frac{\delta[1 - \beta\mu - \alpha^2\beta(1 - \mu)] - (1 - \beta)\alpha\Delta}{1 + \beta[\alpha - (1 + \alpha)\mu]} \end{aligned}$$

and the comparative statics of  $\hat{c}$  w.r.t.  $\alpha$ ,  $\Delta$ , and  $\mu$  are

$$\begin{aligned} \frac{\partial \hat{c}}{\partial \alpha} &= -\frac{\beta\delta(1 - \mu)[1 + \alpha(2 + \alpha\beta) - (1 + \alpha)^2\beta\mu] + (1 - \beta)(1 - \beta\mu)\Delta}{\{1 + \beta[\alpha - (1 + \alpha)\mu]\}^2} < 0 \\ \frac{\partial \hat{c}}{\partial \Delta} &= -\frac{(1 - \beta)\alpha}{1 + \beta[\alpha - (1 + \alpha)\mu]} < 0 \\ \frac{\partial \hat{c}}{\partial \mu} &= -\frac{\alpha(1 + \alpha)\beta(1 - \beta)(\Delta - \delta)}{\{1 + \beta[\alpha - (1 + \alpha)\mu]\}^2} < 0 \end{aligned}$$

where the first two reduce to the monopolist case when  $\mu = 0$ .  $\square$

*Proof of Proposition 14.* To solve this, note that in order for the market to be viable, market makers  $m$  need to be weakly profitable. Their ex-ante utility entering the game, assuming  $\mu < 1$ , is

$$\begin{aligned} V_m &= (1 - \alpha)s_L^* - \alpha s_I^* \\ &= \frac{\delta[1 - \beta\mu - \alpha^2\beta(1 - \mu)] - (1 - \beta)\alpha\Delta}{1 + \beta[\alpha - (1 + \alpha)\mu]} \end{aligned}$$

Note this is equivalent to the average transaction cost they are able to extract. Solving  $V_m \geq 0$  for  $\mu$  produces the upper bound

$$\bar{\mu} = \frac{\delta(1 - \alpha^2\beta) - (1 - \beta)\alpha\Delta}{\beta\delta(1 - \alpha^2)}$$

It's straight forward to check that  $\bar{\mu} \in (0, 1)$  if and only if  $\delta \in (\underline{\delta}, \alpha\Delta)$  as defined in Proposition 10: markets at least have to support a monopolist dealer. The comparative static w.r.t  $\alpha$  is

$$\frac{\partial \bar{\mu}}{\partial \alpha} = \frac{(1 - \beta)(2\alpha\delta - (1 + \alpha^2)\Delta)}{(1 - \alpha^2)^2\beta\delta}$$

which is negative if  $2\alpha\delta - (1 + \alpha^2)\Delta < 0$ , and this is the case for all  $\delta < \alpha\Delta$  (assumption 3). The other comparative static is

$$\frac{\partial \bar{\mu}}{\partial \Delta} = -\frac{\alpha(1 - \beta)}{(1 - \alpha^2)\beta\delta} < 0$$

□

# Appendix C

## Chapter 3 Proofs

*Proof of Proposition 15.* The first order conditions of the investor's expected utility are:

$$\begin{aligned}\frac{\partial \mathbb{E}U(W)}{\partial x_S} : 0 &= (\mu - P) - \rho\sigma^2 x_S - x_F \rho \gamma \sigma \sigma_\epsilon \\ \frac{\partial \mathbb{E}U(W)}{\partial x_F} : 0 &= (\alpha - f) - \rho\sigma_\epsilon^2 x_F - x_S \rho \gamma \sigma \sigma_\epsilon\end{aligned}$$

Solving for  $x_S$  and  $x_F$  gives the demands as a function of the premium  $\mu - P$ .  $\mu - P$  follows by imposing market clearing:  $x_S = 1$ . Substituting this premium into  $x_F$  gives the equilibrium demand as a function of the primitives  $\theta$  and the management fee  $f$ .  $\square$

*Proof of Proposition 16.* Taking the symmetric fee  $\bar{f}$  charged by other managers as given, and substituting the equilibrium risk premium  $\mu - P$  as a function of that fee, the manager's FOC w.r.t. her individual fee  $f$  is

$$\frac{\partial \pi(f; \bar{f})}{\partial f} : 0 = \frac{\alpha(1 - \gamma^2) - 2f + \gamma[\gamma\bar{f} - \rho\sigma\sigma_\epsilon(1 - \gamma^2)]}{\rho\sigma_\epsilon^2(1 - \gamma^2)}$$

In any symmetric equilibrium, it must be the case that the best response  $f = \bar{f}$ . Making that substitution and solving for  $f$  produces the symmetric equilibrium fee  $f^*$ .  $\square$

*Proof of Proposition 17.* To calculate  $\beta_{F,S}$ , we first need the covariance between fund and stock returns:

$$\begin{aligned}\mathbf{Cov}(F, S) &= \mathbf{Cov}(\epsilon - f, D - P) \\ &= \mathbf{Cov}(\epsilon, D) \\ &= \gamma\sigma\sigma_\epsilon\end{aligned}$$

The variance of  $S$  is  $\sigma^2$ , so  $\beta_{F,S} = \frac{\text{Cov}(F,S)}{\sigma^2} = \gamma \frac{\sigma_\epsilon}{\sigma}$ . The performance measure  $\hat{\alpha}$  is then

$$\begin{aligned}\hat{\alpha} &= \mathbb{E}F - \beta_{F,S}\mathbb{E}S \\ &= \alpha - f^* - \gamma \frac{\sigma_\epsilon}{\sigma} (\mu - P(f^*)) \\ &= \alpha - \frac{(1 - \gamma^2)(\alpha - \rho\gamma\sigma\sigma_\epsilon)}{2 - \gamma^2} - \gamma \frac{\sigma_\epsilon}{\sigma} \left[ \rho\sigma^2(1 - \gamma^2) + (\alpha - f)\gamma \frac{\sigma}{\sigma_\epsilon} \right] \\ &= \frac{(1 - \gamma^2)(\alpha - \rho\gamma\sigma\sigma_\epsilon)}{2 - \gamma^2}\end{aligned}$$

□

*Proof of Proposition 18.* For  $V_0$ , note that in a world without the mutual fund, investors simply choose  $x_S$  to maximize

$$\begin{aligned}\mathbb{E}U(W) &= \mathbb{E}W - \frac{\rho}{2} \mathbf{Var}(W) \\ &= x_S(\mu - P) - \frac{\rho}{2} x_S^2 \sigma^2\end{aligned}$$

and the first order condition implies the standard risk-return tradeoff

$$x_S^* = \frac{\mu - P}{\rho\sigma^2}$$

Market clearing requires  $x_S^* = 1$ , which implies the equilibrium risk premium is  $\mu - P = \rho\sigma^2$ . Substituting these back into the objective produces  $\mathbb{E}U(W) = \rho\sigma^2/2$ , which is  $V_0$ .

Once the fund is introduced, the manager's value is the fee multiplied by the portfolio choice:

$$V_M = f^* \cdot x_F^*$$

so substituting in the equilibrium values from Propositions 15 and 16 and simplifying produces  $V_M$ .

Similarly, substituting the equilibrium values of  $x_S^*$ ,  $x_F^*$ ,  $\mu - P$  and  $f^*$  into the investor's utility

$$\mathbb{E}U(W) = x_S(\mu - P) + x_F(\alpha - f) - \frac{\rho}{2} [x_S^2\sigma^2 + x_F^2\sigma_\epsilon^2 + 2x_Sx_F\gamma\sigma\sigma_\epsilon]$$

produces the investors ex-ante certainty equivalent  $V_I$ . Subtracting  $V_0$  from  $V_I$  gives the change in investor utility due to the presence of mutual funds. □

*Proof of Proposition 19.* First consider the change in investor welfare  $\Delta V_I$ . The denominator is positive, and the first term in the numerator is always positive under Assumption 4, so

$$\begin{aligned}\Delta V_I > 0 &\iff \alpha + \rho\gamma\sigma\sigma_\epsilon[3 - 2\gamma^2] > 0 \\ &\iff \alpha > -(3 - 2\gamma^2)\rho\gamma\sigma\sigma_\epsilon\end{aligned}$$

When  $\gamma > 0$ , Assumption 4 insures that this constraint is met. When  $\gamma < 0$ , this constraint binds, so it defines the threshold

$$\alpha_I \equiv -(3 - 2\gamma^2)\rho\gamma\sigma\sigma_\epsilon$$

that applies for  $\gamma < 0$ . First note that the change in total welfare in the economy  $\Delta V_S$  is

$$\begin{aligned} \Delta V_S &= V_M + \Delta V_I \\ &= \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)(\alpha[3 - 2\gamma^2] + \rho\gamma\sigma\sigma_\epsilon)}{2\rho\sigma_\epsilon^2(2 - \gamma^2)^2} \end{aligned}$$

The denominator and first numerator term are again positive, so

$$\begin{aligned} \Delta V_S > 0 &\iff \alpha[3 - 2\gamma^2] + \rho\gamma\sigma\sigma_\epsilon > 0 \\ &\iff \alpha > -\frac{\rho\gamma\sigma\sigma_\epsilon}{(3 - 2\gamma^2)} \end{aligned}$$

which again is satisfied for  $\gamma > 0$ , but when  $\gamma < 0$  it binds, defining the threshold

$$\alpha_S \equiv -(3 - 2\gamma^2)\rho\gamma\sigma\sigma_\epsilon$$

It's straightforward to show  $\alpha_I > \alpha_S$ . □

*Proof of Proposition 20.* The monopolist's fee  $f$  directly affects the premium  $\mu - P(f)$ , so she internalizes its effect, and substituting in this premium gives

$$\begin{aligned} \max_f \pi(f) &= f x_F^*(P^*(f), f) \\ &= f \frac{\sigma(\alpha - f) - \gamma\sigma_\epsilon(\mu - P^*(f))}{\rho\sigma\sigma_\epsilon^2(1 - \gamma^2)} \\ &= f \frac{\alpha - f - \rho\gamma\sigma\sigma_\epsilon}{\rho\sigma_\epsilon^2} \end{aligned}$$

Solving the first order condition w.r.t.  $f$  produces the optimal monopolist fee. Investor demand for the fund and the risky asset premium change under a monopolist, they are obtained by substituting the optimal fee into the equilibrium values of Proposition 15. The performance measure is still

$$\begin{aligned} \hat{\alpha} &= \mathbb{E}F - \beta_{F,S}\mathbb{E}S \\ &= \alpha - f^* - \beta_{F,S}(\mu - P(f^*)) \end{aligned}$$

The value of  $\beta_{F,S}$  is unchanged from Proposition 17. Using the update values of the fee  $f^*$  and premium  $\mu - P(f^*)$  produces the performance measure for the monopolist case. □



*Proof of Proposition 21.* The proof here is identical to Proposition 19, except the monopolist fee is used to determine ex-ante utility and the corresponding boundaries. I omit the details and just provide the quantities.

$$\begin{aligned}
V_M &= \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)^2}{4\rho\sigma_\epsilon^2} \\
\Delta V_I &= \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)(\alpha + 3\rho\gamma\sigma\sigma_\epsilon)}{8\rho\sigma_\epsilon^2} \\
\Delta V_S &= \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)(3\alpha + \rho\gamma\sigma\sigma_\epsilon)}{8\rho\sigma_\epsilon^2} \\
\alpha_I &\equiv -3\gamma\rho\sigma\sigma_\epsilon \\
\alpha_S &\equiv -\frac{1}{3}\gamma\rho\sigma\sigma_\epsilon
\end{aligned}$$

□

*Proof of Proposition 22.* As in Proposition 16, the equilibrium fee is obtained by taking the fee of other managers,  $\bar{f}$ , as given and maximizing profits

$$\max_f \pi(f; \bar{f}) = fx_F^*(P^*(\bar{f}), f) - cx_F^*(P^*(\bar{f}), f)^2/2$$

with respect to  $f$ . I omit the algebraic details, but as in Proposition 16, the first order condition for the above problem is set to 0, and the equilibrium condition  $f = \bar{f}$  is imposed to solve for the symmetric fee  $f^*$ . The fund performance measure  $\hat{\alpha}$  is derived as before, with modifications to the risk premium  $\mu - P(f^*)$  and fund demand  $x_F^*$  to account for the effects of costs on the fee. □

*Proof of Proposition 23.* The algebra with decreasing returns to scale is more complicated, but ex-ante utilities and boundaries are computed as in Propositions 19 and 21. The resulting values are

$$\begin{aligned}
V_M &= \frac{(\alpha - \rho\gamma\sigma\sigma_\epsilon)^2 [c + 2\rho\sigma_\epsilon^2(1 - \gamma^2)]}{2[c + \rho\sigma_\epsilon^2(2 - \gamma^2)]^2} \\
\Delta V_I &= \frac{\rho\sigma_\epsilon(\alpha - \rho\gamma\sigma\sigma_\epsilon)}{2[c + \rho\sigma_\epsilon^2(2 - \gamma^2)]^2} [2c\gamma\sigma + \alpha\sigma_\epsilon + \rho\gamma\sigma\sigma_\epsilon^2(3 - 2\gamma^2)] \\
\Delta V_S &= V_M + \Delta V_I \\
\alpha_I &\equiv -\gamma \left[ \frac{2c\sigma}{\sigma_\epsilon} + \rho\sigma\sigma_\epsilon(3 - 2\gamma^2) \right] \\
\alpha_S &\equiv -\gamma \frac{\rho\sigma\sigma_\epsilon(c + \rho\sigma_\epsilon^2)}{c + \rho\sigma_\epsilon^2(3 - 2\gamma^2)}
\end{aligned}$$

In terms of comparative statics, recall that these thresholds apply for  $\gamma < 0$ .  $\alpha_I$  is obviously increasing in  $c$ , and it is straightforward to show that  $\alpha_S$  is as well. □

# Bibliography

- Admati, Anat R., and Paul Pfleiderer, 1990, Direct and indirect sale of information, *Econometrica* 58, 901–928.
- Allen, Franklin, 2001, Presidential address: Do financial institutions matter?, *The Journal of Finance* 56, 1165–1175.
- Bao, Jack, Jun Pan, and Jiang Wang, 2009, Liquidity of corporate bonds, *Working Paper*.
- Battalio, Robert, Andrew Ellul, and Robert Jennings, 2007, Reputation effects in trading on the new york stock exchange, *The Journal of Finance* 62, 1243–1271.
- Benveniste, Lawrence M., Alan J. Marcus, and William J. Wilhelm, 1992, What’s special about the specialist?, *Journal of Financial Economics* 32, 61–86.
- Berk, J.B., and J.H. van Binsbergen, 2011, Measuring economic rents in the mutual fund industry, *Working Paper*.
- Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.
- Bernhardt, Dan, Vladimir Dvoracek, Eric Hughson, and Ingrid M. Werner, 2005, Why do larger orders receive discounts on the london stock exchange?, *The Review of Financial Studies* 18, 1343–1368.
- Bernhardt, Dan, and Eric Hughson, 1997, Splitting orders, *The Review of Financial Studies* 10, 69–101.
- Bessembinder, Hendrik, William Maxwell, and Kumar Venkataraman, 2006, Market transparency, liquidity externalities, and institutional trading costs in corporate bonds, *Journal of Financial Economics* 82, 251–288.
- Bhattacharya, Sudipto, Amil Dasgupta, Alexander Guembel, and Andrea Prat, 2007, Incentives in funds management: A literature overview, *Working Paper*.
- Biais, Bruno, David Martimort, and Jean-Charles Rochet, 2003, Competing mechanisms in a common value environment, *Econometrica* 68, 799–837.

- Cao, H. Henry, Martin D.D. Evans, and Richard K. Lyons, 2006, Inventory information, *Journal of Business* 79, 325–364.
- Carlin, Bruce Ian, Miguel Sousa Lobo, and S. Viswanathan, 2007, Episodic liquidity crises: Cooperative and predatory trading, *The Journal of Finance* 62, 2235–2274.
- Chaboud, Alain, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega, 2009, Rise of the machines: Algorithmic trading in the foreign exchange market, *Working Paper, Federal Reserve Board*.
- Cho, In-Koo, and David M. Kreps, 1987, Signaling games and stable equilibria, *The Quarterly Journal of Economics* 102, 179–222.
- Cuoco, Domenico, and Ron Kaniel, 2006, Equilibrium prices in the presence of delegated portfolio management, *Working Paper*.
- Dennert, Jürgen, 1993, Price competition between market makers, *The Review of Economic Studies* 60, 735–751.
- Desgranges, Gabriel, and Thierry Foucault, 2005, Reputation-based pricing and price improvements, *Journal of Economics and Business* 57, 493–527.
- Duffie, Darrell, Nicolae Garleanu, and Lasse Heje Pedersen, 2005, Over-the-counter markets, *Econometrica* 73, 1815–1847.
- , 2007, Valuation in over-the-counter markets, *The Review of Financial Studies* 20, 1865–1900.
- Duffie, Darrell, Gaston Giroux, and Gustavo Manso, 2010, Information percolation, *American Economic Journal: Microeconomics* 2.
- Duffie, D., and R. Rahi, 1995, Financial market innovation and security design: An introduction, *Journal of Economic Theory* 65, 1–42.
- Edmans, Alex, 2009, Blockholder trading, market efficiency, and managerial myopia, *The Journal of Finance* LXIV.
- Edwards, Amy K., Lawrence E. Harris, and Michael S. Piwowar, 2007, Corporate bond market transaction costs and transparency, *The Journal of Finance* 62, 1421–1451.
- Fama, Eugene F., and Kenneth R. French, 2010, Luck versus skill in the cross-section of mutual fund returns, *The Journal of Finance* LXV, 1915–1947.
- French, Kenneth R., 2008, Presidential address: The cost of active investing, *The Journal of Finance* LXIII, 1537–1573.

- Glode, V., 2011, Why mutual funds underperform, *Journal of Financial Economics* 99, 546–559.
- Glosten, Lawrence R., 1989, Insider trading, liquidity, and the role of the monopolist specialist, *The Journal of Business* 62, 211–235.
- , and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.
- Goettler, Ronald, Christine Parlour, and Uday Rajan, 2009, Informed traders and limit order markets, *Journal of Financial Economics* 93, 67–87.
- Goldstein, Michael A., Edith S. Hotchkiss, and Erik R. Sirri, 2007, Transparency and liquidity: A controlled experiment on corporate bonds, *The Review of Financial Studies* 20, 235–273.
- Golosov, Michael, Guido Lorenzoni, and Aleh Tsyvinski, 2009, Decentralized trading with private information, *Working Paper*.
- Gomes, Armando, 2000, Going public without governance: Managerial reputation effects, *The Journal of Finance* LV, 615–646.
- Gould, John P., and Robert E. Verrecchia, 1985, The information content of specialist pricing, *Journal of Political Economy* 93, 66–83.
- Green, Richard C., Burton Hollifield, and Norman Schurhoff, 2007, Financial intermediation and the costs of trading in an opaque market, *The Review of Financial Studies* 20, 275–314.
- Green, Richard C., Dan Li, and Norman Schurhoff, 2010, Price discovery in illiquid markets: Do financial asset prices rise faster than they fall?, *The Journal of Finance* LXV, 1669–1702.
- Grinblatt, Mark S., and Stephen A. Ross, 1985, Market power in a securities market with endogenous information, *The Quarterly Journal of Economics* 100, 1143–1167.
- Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *The Journal of Finance* 51, 783–810.
- Guerrieri, Veronica, and Peter Kondor, 2009, Fund managers, career concerns, and asset price volatility, *Working Paper*.
- Hasbrouck, Joel, 1993, Assessing the quality of a security market: a new approach to transaction-cost measurement, *Review of Financial Studies* 6, 191–212.
- Hobson, Jeremy, 2010, Why investors trade with goldman, *NPR's Marketplace*.

- Hotchkiss, Edith S., and Tavy Ronen, 2002, The informational efficiency of the corporate bond market: An intraday analysis, *The Review of Financial Studies* 15, 1325–1354.
- Huang, Jennifer, and Jiang Wang, 2010, Market liquidity, asset prices, and welfare, *Journal of Financial Economics* 95, 107 – 127.
- Huang, Roger D., and Hans R. Stoll, 1997, The components of the bid-ask spread: A general approach, *The Review of Financial Studies* 10, 995–1034.
- Investment Company Institute, 2011, 2011 investment company fact book, .
- Jensen, Michael C., 1968, The performance of mutual funds in the period 1945-1964, *The Journal of Finance* 23, 389–416.
- Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp, 2009, Attention allocation over the business cycle, Discussion paper National Bureau of Economic Research.
- Laffont, Jean-Jacques, and Eric S. Maskin, 1990, The efficient market hypothesis and insider trading on the stock market, *Journal of Political Economy* 98, 70–93.
- Leland, Hayne E., 1992, Insider trading: Should it be prohibited?, *Journal of Political Economy* 100, 859–887.
- Madhavan, Ananth, Matthew Richardson, and Mark Roomans, 1997, Why do security prices change? a transaction-level analysis of nyse stocks, *The Review of Financial Studies* 10, 1035–1064.
- Mailath, George, Masahiro Okuno-Fujiwara, and Andrew Postlewaite, 1993, Belief-based refinements in signalling games, *Journal of Economic Theory* 60, 241.
- Mailath, G.J., and L. Samuelson, 2006, *Repeated games and reputations: long-run relationships* (Oxford University Press, USA).
- Pastor, Lubos, and Robert F. Stambaugh, 2010, On the size of the active management industry, *Working Paper*.
- Permanent Subcommittee On Investigations, 2010, Wall street and the financial crisis: The role of investment banks, Discussion paper U.S. Senate Washington, DC.
- Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of Finance* 39, 1127–1139.
- Savov, Alexi, 2010, Free for a fee: The hidden cost of index fund investing, *Working Paper*.
- Securities and Exchange Commission, 2010, Securities and exchange commission v. goldman, sachs co. and fabrice tourre, 10 civ. 3229, Discussion paper Litigation Release No. 21489.

- Seppi, Duane J., 1990, Equilibrium block trading and asymmetric information, *The Journal of Finance* 45, 73–94.
- Stein, Jeremy C., 1988, Takeover threats and managerial myopia, *The Journal of Political Economy* 96, 61–80.
- Stracca, Livio, 2006, Delegated portfolio management: A survey of the theoretical literature, *Journal of Economic Surveys* 20, 823–848.
- Taibbi, Matt, 2009, The great american bubble machine, *Rolling Stone* 1082-1083.
- The Financial Services Authority, 2009, Reforming remuneration practices in financial services, Discussion paper.
- Vayanos, D., and P. Woolley, 2008, An institutional theory of momentum and reversal, Discussion paper National Bureau of Economic Research.