Middlemen in Limit Order Markets[1]

Boyan Jovanovic and Albert J. Menkveld

June 20, 2016

Middlemen in Limit Order Markets

Abstract

A limit order market enables an early seller to trade with a late buyer by leaving a price quote. Information arrival in the interim period creates adverse selection risk for the seller and therefore hampers trade. Entry of high-frequency traders (HFTs) might restore trade as their machines can refresh quotes quickly on (hard) information. Empirically, HFT entry reduced adverse selection by 23% and increased trade by 17%. Model calibration shows that one percentage point more of the gains from trade were realized. Finally, we show that a well-designed double auction raises this to ten percentage points.

For online appendix, please click: http://goo.gl/40Cnyd.

1 Introduction

Automation has changed the structure of securities markets in recent decades. Trading floors were replaced by electronic venues. Investors can now trade without intermediaries by submitting "limit orders," i.e., price quotes to either buy or sell a security. The trading venue matches incoming orders sequentially. It will try to match any incoming new order against a stock of unexecuted orders. If there are multiple matches, it will pick the one that is most beneficial to the new order and conclude a trade. If there is no match, the new order will be added to the stock of unexecuted orders (see Parlour and Seppi, 2008, for a literature survey).

Figure 1: Market structure change. This graph illustrates two important trends in equity markets in the second half of the 2000s. The first series plotted here is the NYSE market share in NYSE-listed stocks. The second series is the market share of high-frequency traders in overall equity market volume. The first series was constructed based on data from Barclays Capital Equity Research. The second series was taken from "High Frequency Trading: Evolution and the Future," a report published by Capgemini.
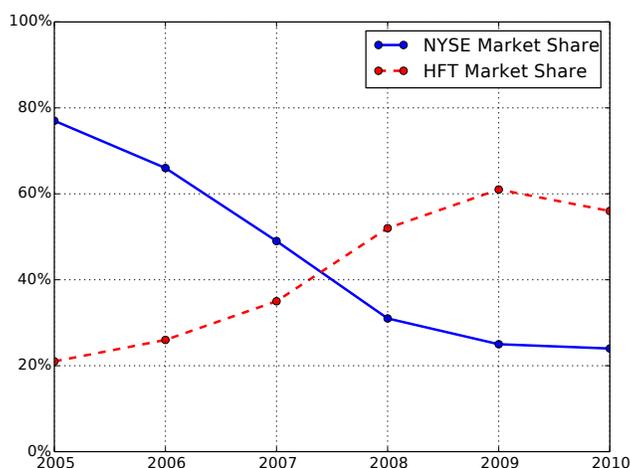
Figure 1 illustrates perhaps the most dramatic example of the migration towards limit order markets. The New York Stock Exchange (NYSE) ran a floor-based market in the late 2000s. Its market share in NYSE-listed equity dropped from 77% in 2005 to 24% in 2010. Volume migrated largely towards limit order markets.

Intermediation did not disappear, however. Human market makers such as the NYSE specialist were replaced by extremely fast machines: high-frequency traders (HFTs). They are a type of

intermediaries as they establish and liquidate positions at "very short time frames," SEC (2010, p. 45). Figure 1 illustrates that their market share in equities grew from 21% in 2005 to 56% in 2010.

High-frequency traders have the ability to parse large amounts of information almost instantly. Yet, classic models of intermediation have the market maker at an informational disadvantage (e.g., Glosten and Milgrom, 1985; Kyle, 1985). What happens when this friction is turned on its head? Can better informed middlemen improve welfare?

Model outline. Let us outline intuitively the approach that we take. The model involves a seller who values an asset at $x + z$ and a buyer who values it at $y + z$. The private values x and y are drawn independently from a common distribution, and z is private information of the buyer. It helps to think of this as combining two different models. In the first model, z is public information (or unknown by both). In the second model, z is private information, but it is common knowledge that there are gains from trade, $y > x$. Let us briefly discuss each model.

First, if z were common knowledge, this would strongly resemble the Myerson and Satterthwaite (1983) environment. Ideally, the buyer would get the good if and only if $y > x$, but we know that there is no mechanism that delivers that outcome. Obviously, limit orders are a special case of mechanisms, so they will not be able to deliver that outcome.

Second, if z is private information but $y > x > 0$, then the first best would dictate that trade always occurs. But limit orders cannot implement the first best if the support of z, say $[z_1, z_2]$, is sufficiently large, i.e., if $z_2 - z_1 > y - x$. The question then is whether some other mechanism could implement the first best in this environment. Middlemen seem to be exactly the right mechanism. More precisely, suppose in the second environment that there are competitive middlemen who know y and z but value the good only at z. In equilibrium, middlemen would post competitive limit orders, and buy at the price $y + z$, and then sell the asset for that price. This competitive sector then achieves the first best (and gives all the rents to the seller).

Our paper tackles a hybrid of these two environments, augmented by the fact that neither the buyer nor the middlemen know z perfectly, and middlemen do not know y (augmentation dictated

2

by empirics). We know from the first environment that there is no way to achieve the first best. The second environment gives us reasons to believe that middlemen will be able to improve upon equilibrium allocations, and our estimates suggest that this is indeed the case. We also show, however, that a well-designed double auction can do even better.

Main findings.   The model and its calibration yield the following main insights. First, the more of the common value z is machine-readable "hard information," the more useful a competitive middleman is.[1] Instead of quoting himself, the seller will pass the security off to a middleman who can refresh his price quote on hard information. A middleman therefore is not adversely selected by the buyer on this information. The cost of going down this route however is that when the buyer's valuation is below the middleman quote then the security will be stuck with the middleman who might value it less than the seller. In a competitive equilibrium, the middleman will pass the expected cost and benefit (that they experience in the subgame with the buyer) to the seller.[2] The seller thus passes off the security to the middleman only if the "rents" of following this path are higher than the expected value of posting a price himself. It turns out that he only passes it off in states where his private value is low and there is lots of hard information.

Second, although the seller gets to choose whether or not to accept a middleman bid, he could in fact be worse off after middleman entry. The best case for middleman entry (in terms of welfare) is when the buyer is fully informed on the hard information that arrived in the interim period, i.e., the period between seller departure and buyer arrival. Middleman entry benefits trade in this case as it removes some of the adverse selection risk. The worst case is when the buyer remains uninformed on the hard information that arrived. In this case, the seller was not adverse selected on hard information before middleman entry. After entry he suddenly faces adverse selection risk.

---

[1]We refer to Petersen (2004) for an elaborate discussion on hard versus soft information. Examples of hard information in our case are price changes in the market index, in same-industry stocks, or a signal obtained from a machine parsing news releases. Examples of soft information are the quality of a new management team or the value of a patent. In finance one would say that the seller's price quote is not necessarily semi-strong efficient whereas a middleman price quote is (Fama, 1970).

[2]Our model only considers the competitive case and whenever we refer to middleman in a singular tense we assume a competitive middleman. This assumption is reasonable as limit order markets feature Bertrand price competition as will become clear in the model discussion.

Middleman entry hurts trade in this case. The model accounts for intermediate cases by letting the buyer learn hard information with some probability λ.

Third, model calibration illustrates the importance of the "hidden cost" associated with middleman entry mentioned in the second point. A seller faces a potentially larger threat of being adversely selected when he decides to post his own quote. In a way, the middleman bid in the first round becomes an "offer he cannot refuse." This observation explains why the model calibration yields a modest welfare increase of at most one percent (expressed in terms of expected gains from trade) while improvements in standard trade statistics are sizeable: a reduction in adverse selection cost of 23% and an increase in trading frequency of 17%. Thus, investors experience a lower bid-ask spread (due to less adverse selection cost) after middleman entry but are, at the same time, effectively forced out of earning the spread as their prices face the threat of being "sniped" by middlemen. This finding warns against over-interpreting large spread reductions and volume increases often associated with HFT entry as evidence of strongly improved "market quality."

Fourth, the empirical analysis of middleman participation provides support for one of the model's most "idiosyncratic" predictions: Middleman participation is higher on days with relatively more hard information.

Fifth, at the calibrated parameters, middleman entry in a limit order market does worse than replacing it with well-designed periodic double auctions. Middleman entry achieves one percentage point more in terms of gains from trade, whereas the double auction could raise this to ten percentage points more. This finding speaks to a recent literature that proposes frequent batch auctions as a potentially better market design than limit order markets (Farmer and Skouras, 2012; Budish, Cramton, and Shim, 2015; Wah and Wellman, 2013). In the double auction alternative, the seller's ask price remains hidden from the buyer when the latter decides his bid price. If the auction design is such that the transaction price is "close enough" to the bid when the bid exceeds the ask, then double auctions do better in terms of welfare than middleman entry into limit order markets. A transaction price close to the most informed investor's price by design, appears to reduce the adverse selection friction.

Finally, we analyze a monopoly middleman version of the model and calibrate it to the same data. Such analysis of the other polar case — monopoly instead of perfect competition — yields further insight into what it means to add a middleman to a limit order market. The change is non-trivial. In the competitive case, middlemen Bertrand compete and reveal their information through their bids and pass on future gains from trade, all in a relatively trivial sense. The monopoly model is more complex as the middleman effectively controls the beliefs of the buyer and seller through how he bids. In equilibrium, the bidding function needs to satisfy an ordinary differential equation. Calibrating the solution to the data suggests that, contrary to the competitive case, the entry of a monopoly middleman led to a modest welfare reduction: two percentage points. The data however seem to favor the competitive model as it is better able to match the sizeable entry effects.

High-frequency trading literature. The paper adds to a recent and fast growing theoretical literature on high-frequency trading (see, e.g., Menkveld, 2016, for a review). Some models emphasize the speed advantage of HFTs in the trading process (e.g., Cartea and Penalva, 2012; Aït-Sahalia and Saglam, 2014; Foucault, Hombert, and Roşu, 2016; Roşu, 2016; Hoffmann, 2014; Menkveld and Zoican, 2016). Others focus on the arms race element of competition among HFTs (e.g., Budish, Cramton, and Shim, 2015; Biais, Foucault, and Moinas, 2015; Glode and Opp, 2016). None of the models endogenizes the response of non-HFTs to HFT entry as we do. The only exception is Hoffmann (2014) who models HFTs as existing traders investing into faster market access. Our model analyzes HFT entry as the arrival of a new middleman who is informed on the "hard" part of common value innovations. We further compare HFT entry to an alternative market design: a double auction instead of a limit order market.[3]

---

[3]The calibrated welfare change and the result that HFTs participate more at times of relatively more "hard" information is our unique contribution to a rapidly developing empirical literature on high-frequency trading, and on algorithmic trading more generally (e.g., Hendershott, Jones, and Menkveld, 2011; Baron, Brogaard, and Kirilenko, 2012; Breckenfelder, 2013; Easley, López de Prado, and O'Hara, 2013; Hagströmer and Nordén, 2013; Hasbrouck and Saar, 2013; Boehmer, Fong, and Wu, 2014; Chaboud *et al.*, 2014; Kirilenko *et al.*, 2014; Brogaard, Hendershott, and Riordan, 2014; Boehmer, Li, and Saar, 2015).

Middlemen in search markets literature. More broadly still, our findings contribute to the literature on middlemen in search markets. Our middlemen do not speed up the meeting process, a role that is central to the models of Rubinstein and Wolinsky (1987) and Masters (2007). They also do not have information about private values, which creates an evaluative role for middlemen in Li (1996). Rather, they are informed about common values. The models closest to our model are therefore those in which some investors have a signal on the common value, while others do not. There are the centralized market-clearing models (Grossman and Stiglitz, 1980) and models of bilateral trade (e.g., Golosov, Lorenzoni, and Tsyvinki, 2014; Duffie, Malamud, and Manso, 2009). None of these models captures the essence of the most prevalent modern market structure: A limit order market in which agents can post price quotes themselves or consume price quotes of others.

Regulation. The SEC aimed at starting a public debate on high-frequency trading as part of their 2010 market structure review (SEC, 2010). It followed up recently with a discussion document focused on HFT only (SEC, 2014). Jones (2013) summarizes some commonly voiced regulatory proposals: order-level audit data, transaction tax, order cancellation or excess message fees, and a minimum order exposure time. With the exception of audit data, our analysis suggests that all measures are potentially harmful. Order cancellation or excess message fees in particular, as they would hurt a middleman's ability to reduce overall adverse selection cost. It would make it costly for him to continuously update his quote on the arrival of hard information.

2 Model

The model features two investors whose stay in the market does not overlap: an early (potential) seller (**S**) who owns one unit of the security and a late (potential) buyer (**B**) who does not (Grossman and Miller, 1988). At some point, middlemen or machines (**M**) are introduced in the market that, unlike investors, have zero opportunity cost and are therefore in the market permanently. The

agents (**S**, **B**, and **M**) have available to them a limit order market to potentially effectuate a trade and therefore a transfer of the security. All agents in the model are risk-neutral. An overview of the model's notation and the purpose of the various variables is included as Appendix A.

S values the security at $x + z$, **B** values it at $y + z$. The private values x and y are drawn independently from a common distribution with CDF F. The common value z consists of two independent components:

$$z = z^h + z^s.$$

The superscripts refer to machine-processable "hard" information (h) and, its complement, "soft" information (s) that can only be processed by humans (Petersen, 2004). Let

$$\alpha = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_s^2} \tag{1}$$

denote the relative importance of hard information. The CDFs of z, z^h, and z^s are G, G^h, and G^s, respectively.

Middlemen (**M**) have zero private value for the security. In other words, each **M**'s valuation of the security is just z. **M** always know z^h but, unlike investors, do not have the ability to parse z^s.

The informational structure of the game is depicted in Panel A of Figure 2. Private values are observed only by the agents themselves. The hard common value z^h is observed by **M** and, with probability λ, by **B**. The soft common value z^s is observed only by **B**. The latecomer, **B**, is better informed partly because of the passage of time. In fact, it is likely that **B** enters *because* he has seen information that may be useful to him; this is a selection effect that is left unmodeled here.

2.1   The game without middlemen

This section develops the game without **M**. It consists of two periods:

1. **S** observes (only) x, posts an ask p^S, and leaves.

Figure 2: Structure of the game.   This figure illustrates the game between the seller **S**, the middleman **M**, and the buyer **B**. Panel A illustrates the informational structure. Panel B depicts the various steps in the game.

Panel A: information structure
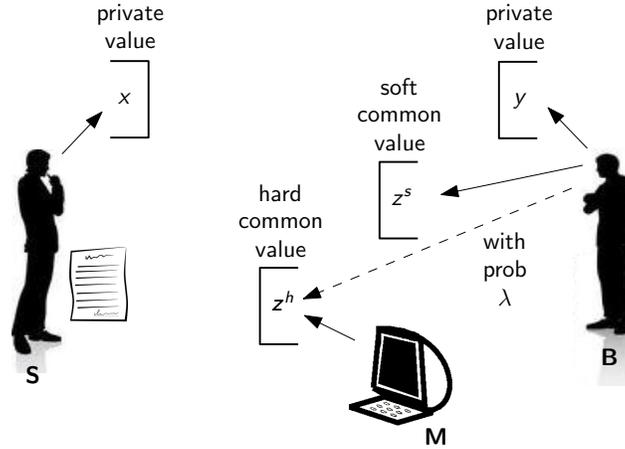
Panel B: steps in the game
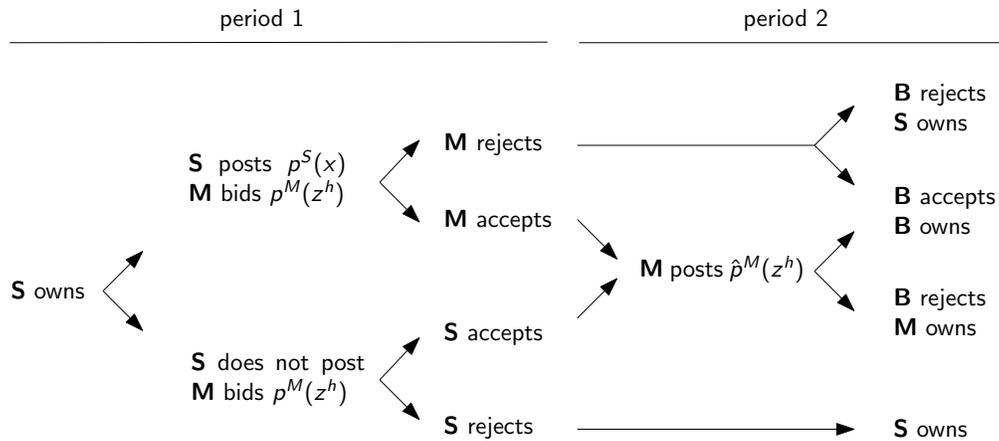
2. **B** arrives, observes y and z^s for sure and z^h with probability λ, and decides on whether or not to accept the posted ask price.

Note that a bid by **B** is excluded as by the time that **B** can post one, **S** is gone.

As **S** can only condition on his private value x, his ask price is uncorrelated with z. **B** accepts $p^S = p$ iff

$$\begin{cases} y + z > p & \text{with probability } \lambda, \\ y + z^s > p & \text{with probability } 1 - \lambda. \end{cases} \tag{2}$$

S receives p in the event of a sale and $x + z$ otherwise. Conditioning on y, **S**'s expected return therefore is:

$$\lambda\left(p\left[1 - G(p - y)\right] + \int_{-\infty}^{p-y}(x + z)g(z)\,dz\right) +$$
$$(1 - \lambda)\left(p\left[1 - G^s(p - y)\right] + \int_{-\infty}^{p-y}(x + z)g^s(z)\,dz\right), \tag{3}$$

where $g(z) = \int_{-\infty}^{\infty}g^h(z - z')g^s(z')\,dz'$ is the density of z, g^h is the density of z^h, and g^s the density of z^s. Since (3) is linear in probabilities, **S** chooses p as follows:

$$v\left(x, g^\lambda\right) = \max_p \int_{-\infty}^{\infty}\left(p\left(1 - G^\lambda(p - y)\right) + \int_{-\infty}^{p-y}(x + z)g^\lambda(z)\,dz\right)dF(y), \tag{4}$$

where, for all z, $G^\lambda = \lambda G + (1 - \lambda)G^s$ and $g^\lambda = \lambda g + (1 - \lambda)g^s$.

Note that in (2), **B** conditions on the full z or on z^s only and he therefore imposes adverse selection cost on **S**, i.e., $E(z \mid \text{trade}) > 0$. The presence of z lowers **S**'s utility and if z is important enough, it can eliminate trade altogether:

Adverse selection and trade with normally distributed x, y and z.   Suppose $x \sim N(1, 1)$, $y \sim N(1, 1)$, and $z \sim N\left(0, \sigma^2\right)$. We shall now show that if we raise the parameter σ relative to the unit variance of x and y, this worsens the adverse-selection problem and as σ becomes large enough, trade disappears. That is, let

$$\phi(u) = \frac{1}{\sqrt{2\pi}}e^{-u^2/2} \quad \text{and} \quad \Phi(u) = \int_{-\infty}^{u}\phi(s)\,ds$$

denote the standard normal density and cumulative density, respectively. Further, let $p(x)$ be **S**'s optimal ask price. The probability that the offer of a type-x seller is accepted is

$$\tau(x) = \int_{-\infty}^{\infty}\left[1 - \lambda\Phi\left(\frac{p - y}{\sigma_h + \sigma_s}\right) - (1 - \lambda)\Phi\left(\frac{p - y}{\sigma_s}\right)\right]\phi(y - 1)\,dy. \tag{5}$$

9

We claim that it converges to zero as $\sigma \to \infty$. Since y and z are independent, their sum $s = (y + z)$ is normal with variance $1 + \sigma^2$. Then

$$F(y) = \Phi(y - 1), \quad f(x) = \phi(x - 1), \quad G(z) = \Phi\left(\frac{z}{\sigma}\right), \quad \text{and} \quad g(z) = \frac{1}{\sigma}\phi\left(\frac{z}{\sigma}\right). \tag{6}$$

for $\sigma \in \{\sigma_s, \sigma_s + \sigma_h\}$. We can then write **S**'s problem in (4) as

$$\max_p \int_{-\infty}^{\infty} \left\{ \begin{array}{l} \left[1 - \lambda\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) - (1 - \lambda)\Phi\left(\frac{p-y}{\sigma_s}\right)\right]p \\ + \int_{-\infty}^{p-y}(x + z)\left[\frac{\lambda}{\sigma_h+\sigma_s}\phi\left(\frac{z}{\sigma_h+\sigma_s}\right) + \frac{1-\lambda}{\sigma_s}\phi\left(\frac{z}{\sigma_s}\right)\right]dz \end{array} \right\} \phi(y - 1)\, dy. \tag{7}$$

PROPOSITION 1. ———————————————————————————————*("No trade theorem")*

Under the normality assumptions in (6), if $\lambda > 0$,

$$(i) \quad \lim_{\sigma \to \infty} \tau(x) = 0 \ \text{ for each } x \in \mathbb{R} \qquad and \qquad (ii) \quad \frac{\partial v}{\partial \sigma} < 0 \qquad (8)$$

for $\sigma \in \{\sigma_s, \sigma_h\}$.

The proofs of the lemmas and propositions are all in in Appendix B.

When is there no adverse selection cost? Naturally, there is no adverse selection cost to **S** if there is no common value, i.e., $\sigma_z^2 = \sigma_h^2 + \sigma_s^2 = 0$. Empirically more relevant, however, is the case:

$$\sigma_s^2 = \lambda = 0, \ \text{coupled with } \sigma_h^2 > 0. \qquad (9)$$

In this case, only z^h matters but neither **S** nor **B** is aware of it. In this case, there is no adverse selection between the seller and the buyer to begin with; the decision rule (2) reduces to "buy iff $y > p$", and (4) reduces to:

$$\max_p \left(p \left(1 - F(p) \right) + x F(p) \right). \qquad (10)$$

If **M** is not in the market, the ask price that maximizes this expression is an equilibrium.

How might M introduce adverse selection cost? Suppose condition (9) holds so that the equilibrium $p^{\mathrm{S}}(x)$ solves (10). This equilibrium is vulnerable to the entry of **M** as, knowing z^h, the machines will buy the security whenever $z^h > p$, leaving **S** or **B** with the security in states where $z^h < p$. Therefore **M** introduce adverse selection in the market and reduce trade and welfare.

2.2 The game with middlemen

High-frequency traders are modeled as a group of competitive middlemen. Surely they will adversely select investors' price quotes, but competition also forces them to place bid quotes them-

selves (through buy limit orders). Their quotes, however, are not vulnerable to adverse selection on z^h as all **M** know z^h. By incorporate its realization into their bid, **M** serve to inform both **B** and **S** about z^h.

The game is structured in such a way that **S** can only take one action while he is in the market.[4] Panel B of Figure 2 illustrates the following sequential structure of the game.

Period 1. **S** arrives and can take only *one* action: either post an ask or accept a bid.

1. **S** observes x and might post an ask p^S. Simultaneously, **M** post competitive bids p^M.

2. If **S** did not post an ask he can accept an **M**'s bid. If he rejects it, the game ends.

3. If **S** did post an ask (and therefore foregoes being able accept an **M**'s bid), **M** accept or reject it.

4. **S** leaves.

Period 2. **B** arrives.

1. **B** observes z^h which is revealed through **M**'s bids. **B** also observes y and z^s.

2. If an **M** bought the security from **S**, it[5] now tries to sell it to **B**. Its ask is $\hat{p}^M$.

3. If **M** rejected the ask that **S** posted, the ask is now available for **B**, and **B** takes it iff $y+z > p^S$.

Under these assumptions, **M** compete each other out of all rents that derive from knowing z^h and from being fast. Competitive bids by **M** in period 1.1 allow both **S** and **B** to infer z^h and, as a result, unless **S** has posted an ask, the adverse selection on z^h disappears. One important consequence of this is that, regardless of λ, **B** learns z^h and therefore gets to know both components of z. Therefore, λ affects only the game in which **M** are absent.

---

[4]This rules out that **S** observes **M**'s bid, learns z^h, turns around, and posts a more informed ask price himself. The proposed structure is a reduced form of a more elaborate model that adds an intermediate stage where **M** learns z^h and updates his quote. We chose the current structure for reasons of parsimony.

[5]Throughout, a machine is referred to as "it" while an investor is referred to as "he".

2.2.1 Second period sub-game between **M** and **B**

As **B** infers z^h from competitive bidding by **M** in period 1, this hard information is common knowledge between them in the second period sub-game when **M** owns the security and posts an ask price $\hat{p}^M(z^h)$ (shortened to $\hat{p}^M$) for **B** to consider. **B** accepts the offer iff

$$y + z^s + z^h > \hat{p}^M. \tag{11}$$

Now let

$$p = \hat{p}^M - z^h. \tag{12}$$

Then from (11), trade takes place iff

$$z^s > p - y. \tag{13}$$

The *ex ante* value of the security to **M** therefore is $z^h + v^M$, where

$$v^M = \max_p \int_{-\infty}^{\infty} \left(p \left[1 - G^s (p - y) \right] + \int_{-\infty}^{p-y} z g^s (z) \, dz \right) dF(y). \tag{14}$$

is the profit that **M** can expect to extract from **B**. Note that in doing this **M** faces adverse selection cost itself induced by **B**'s private knowledge of z^s. Also, v^M does not depend on g^h and its value is therefore independent of σ_h^2. Moreover, a comparison of the RHSs of (4) and (14) reveals that they only differ in terms of the distribution $g(.)$. In fact: $v^M = v(0, g^s)$. Consequently, when $\sigma_h = 0$, an **M** that owns the security is in the same situation as **S** was in the game without **M**.

2.2.2 First period choice for S to post an ask price or not

S is ignorant about both z^s and z^h and posting an ask price therefore exposes him to adverse selection on both these common value components. Therefore he may prefer not to post, but he cannot avoid adverse selection altogether as even (instead of posting) accepting **M**'s bid exposes him indirectly. The competitive bids by **M** pass on to **S** the revenue that **M** expects in the second period

sub-game with **B**, but also **M**'s expected cost in that **M** is adversely selected on z^s by **B**. In the end, **S** has two options to trade with **B**: he either finds him through the upper branch of the game's tree by posting himself (see panel B of Figure 2) or through the lower branch by accepting **M**'s bid. As will become clear, the second route has the benefit of removing adverse selection on z^h but has the cost of paying a welfare loss when the machine ends up holding the security (as opposed to **S** himself). We now analyze the value of both branches of the tree.

If S does not post an ask (lower branch). If **S** does not post an ask, he knows the competitive **M**'s bid is:

$$p^M = z^h + v^M, \tag{15}$$

which therefore reveals z^h to him.

S now has the option to accept or reject. Having learned z^h, he accepts iff

$$x + z^h + E\left(z^s\right) < z^h + v^M \quad \Leftrightarrow \quad x < v^M, \tag{16}$$

since $E\left(z^s\right) = 0$. **M** therefore obtains the security iff (16) holds. **S**'s expected utility conditional on x therefore equals:

$$U\left(x\right) = \max\left(x, v^M\right). \tag{17}$$

Differentiating $U\left(x\right)$ for $x \neq v^M$ yields:

$$U'\left(x\right) = I_{\left\{x > v^M\right\}} = \Pr\left(\text{no trade} \mid x\right), \tag{18}$$

where $I_{\{.\}}$ is the indicator function.

If S does post an ask (upper branch). If, instead, the seller decides to post an ask price, he is fully exposed to adverse selection on hard information. Let $V(x)$ denote the value to him of this

option. If **S** posts $p^S(x)$, it will be accepted if either an **M** or **B** values the security more, i.e., if:

$$p^S(x) < \max\left(z^h + v^M, y + z^s + z^h\right)$$
$$= z^h + \max\left(v^M, y + z^s\right) \tag{19}$$
$$= z^h + \xi\left(y + z^s\right).$$

M are faster than **B** and therefore have the first take on the offer. If $p^S < z^h + v^M$, the **M** that is lucky enough to snap up the security earns a positive expected rent. Thus, if p^S is in the money for **M**, one of them will snap it up. If not, the ask p^S remains in the order book and is available to **B** when he shows up. **B** will accept the offer if $p^S < y + z^s + z^h$.

Leading up to **S**'s optimization problem, condition (19) is more easily expressed as: $z^h > p^S - \max\left(v^M, y + z^s\right)$. Anticipating **M**'s and **B**'s response, **S** chooses p^S to solve:

$$V(x) = \max_p \int \left(\begin{array}{l} p\left[1 - G^h\left(p - \max\left(v^M, y + z^s\right)\right)\right] + \\ \int_{-\infty}^{p - \max\left(v^M, y + z^s\right)}\left(x + z^s + z^h\right)dG^h\left(z^h\right) \end{array} \right) dG^s\left(z^s\right) dF\left(y\right). \tag{20}$$

Differentiating (20), using the envelope theorem, and simplifying yields:

$$V'(x) = \Pr\left(z^h + \xi\left(y + z^s\right) \le p^S(x)\right) > 0, \tag{21}$$

where $p^S(x)$ is the optimal policy for **S.**

This expression is analogous to (18), i.e., the derivative $V'(x)$ is equal to the probability of no trade conditional on x:

$$V'(x) = \int G^h\left(p^S(x) - \xi\left(y + z^s\right)\right)dG^s\left(z^s\right)dF\left(y\right). \tag{22}$$

Moreover, $\lim_{x \to \infty} V'(x) = 1$ and $V'(x) < 1$.

If $p^S(x)$ is increasing, V is convex and if $p^S(x)$ is differentiable, we can use the chain rule to

15

conclude that:

$$V''(x) = \frac{\partial p^S(x)}{\partial x} \cdot \int g^h \left(p^S(x) - \xi(y + z^s) \right) dG^s(z^s) \, dF(y) > 0,$$

That $\frac{\partial p^S(x)}{\partial x} > 0$ follows from the FOC to (20) which (after a cancellation of $p \times g^h \left[p - \max \left(v^M, y + z^s \right) \right]$) reads

$$0 = \int \left(\begin{array}{c} 1 - G^h \left(p - \max \left(v^M, y + z^s \right) \right) + \\ \left[x + z^s - \max \left(v^M, y + z^s \right) \right] g^h \left[p - \max \left(v^M, y + z^s \right) \right] \end{array} \right) dG^s(z^s) \, dF(y). \qquad (23)$$

A rise in x raises the RHS of (23) and the second-order conditions (which follow from the log-concavity of the normal density) imply that p^S must rise.

S's decision whether or not to post.   This choice problem is illustrated in the top-left panel of Figures 3, 4, and 7, each time for a different parameter set. As the figures show, the optimal decision depends on x. **S** posts iff

$$x \in A, \quad A = \{x | V(x) \geq U(x)\}. \qquad (24)$$

As $V(x) > x$, (24) holds for $x \geq v^M$. An intersection of V and U can therefore only occur at a value of x below v^M. Moreover, (22) reveals that $V' > 0$ so that any such intersection must (i) be unique and (ii) must occur at a value of x, say a, that is smaller than v^M so that

$$v^M = V(a). \qquad (25)$$

The set A therefore becomes: $A = [a, \infty)$. Finally, as $V(x) > x$, (25) implies:

$$x \notin A \implies x < v^M, \qquad (26)$$

which leads to

16

PROPOSITION 2. ——————————————————————————————*(Order posting strategy seller)*

(*i*) *If* **S** *does not post then* **S** *accepts* **M**'s *bid with probability 1.*

(*ii*) *If* $z^h \sim N\left(0, \sigma_h^2\right)$, *then as* σ_h^2 *rises, a rises and A shrinks.*

———————————————————————————————————————

Part (*i*) of this proposition effectively removes the very lowest branch from the game tree (Panel B of Figure 2). Part (*ii*) says that, *ceteris paribus*, **S** is more likely to "use" **M** to get to **B** when there is (relatively) more hard information, *ergo*, **S** benefits from **M**'s technology. The analog of part (*ii*) for σ_s does *not* hold. A change in σ_s reduces both $V(x)$ and v^M and the net effect on a is therefore unclear.

LEMMA 1. ——————————————————————*(Seller self-protects against adverse selection)*

For each x,

$$\lim_{\sigma_h \to \infty} p^S(x) = +\infty \tag{27}$$

and

$$\lim_{\sigma_h \to \infty} V(x) = x. \tag{28}$$

———————————————————————————————————————

An immediate result from this lemma is:

PROPOSITION 3. —————————————————————— *(Maximum middlemen participation)*

$$\lim_{\sigma_h \to \infty} a = v^M,$$

where the convergence of a is monotonic. As σ_h *rises,* **M**'s *trade participation gets arbitrarily close to 100%.*

———————————————————————————————————————

Corollary. An increase in α, the relative proportion of hard information, implies more **M** participation in trades.

This rather idiosyncratic prediction of the model will be taken to the data in Section 1.3.

## 2.3   A graphical illustration of the effect of M entry

Before developing some more results in order to connect the model to the data, this section illustrates how the entry of middlemen affects trading through two sets of graphs. They are meant to compare and contrast two extreme cases: welfare-reducing **M** and welfare-enhancing **M**. The key parameters that drive this distinction are α, the relative size of hard information (or, how much of an informational edge do machines have) and λ, the likelihood that **B** is aware of hard information (or, how much of an informational friction was there to begin with). We vary latter parameter to develop the extremes. Figures 3 and 4 illustrate the **M** and no-**M** case for $\lambda = 0$ and $\lambda = 1$, respectively. Note that changes in λ do not affect the **M** game because the **M** bid $(z^h + p)$ (see (12)) which fully reveals z^h regardless of λ.

The parameter choices are mostly arbitrary and are picked to maximize visual differences. $\theta = 3$ and $\alpha = 0.9$. The private value and common value distribution, F and G, are taken to be Gaussian. Private value variance is set to scale all welfare results relative to the first-best gains-from-trade, i.e., $\sigma = \sqrt{\pi}$ yields (Clark, 1961, eqn. (2)):

$$\text{``first-best''} - \text{``autarky''} = E \max (x, y) - E (x) = \frac{\sigma}{\sqrt{\pi}} = 1, \tag{29}$$

so that (x, y, z) are now measured in "gains-from-trade" units. The private value mean μ is set to 4 in order to avoid the artificial welfare gain that occurs when introducing an additional agent (**M**) into the economy. That is, the probability that **M** values the security more than both **S** and **B**

Figure 3: Comparative statics: welfare-reducing middlemen ($\lambda = 0$) The case of welfare-reducing middlemen is summarized in four graphs. The solid blue line illustrates the benchmark case of no middlemen (**M**) and the dashed red line illustrates the case where **M** are present. The default parameter values for these graphs are: $\lambda=0$, $\mu=4$, $\sigma^2=\pi$, $\theta=3$, and $\alpha=0.9$. First-best welfare is 1 by construction; welfare in the no-**M** case is 67%; welfare in the **M** case is 34% (i.e., a reduction of 49% relative to the **M** case). The top-left graph illustrates for what private value x the seller (**S**) decides to post a price himself rather than wait for a bid by **M**, i.e., the region in which $V(x)$ exceeds $U(x)$. The top-right graph illustrates the ask price **S** posts in case he decides to post a price. The bottom-left graph plots the probability of a transfer of the security from **S** to **B**. The bottom-right contour plot graphs the probability of an **S**-**B** transfer as a function of x and y which are the private values of the **S** and **B**, respectively.
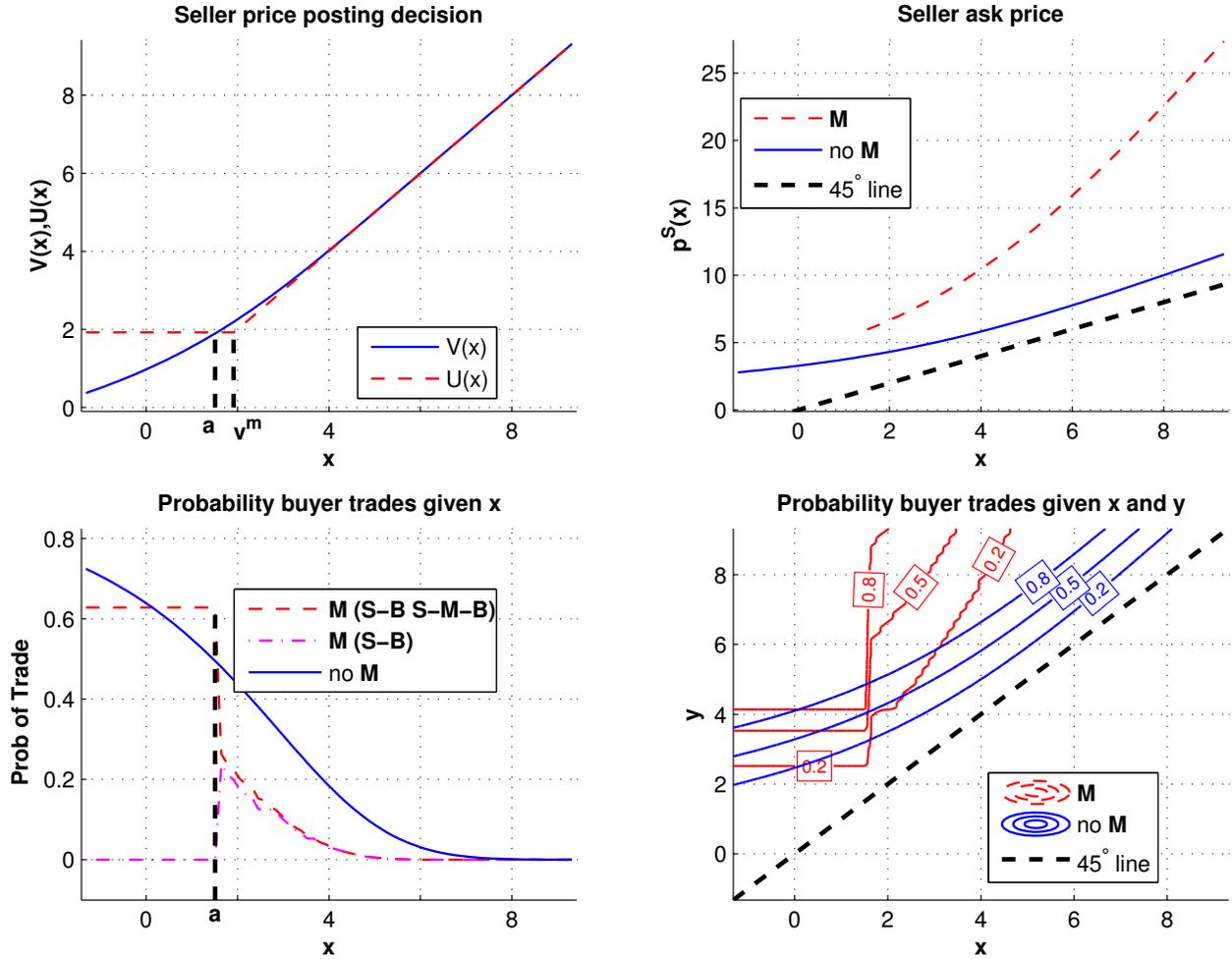
becomes very small:

$$\Pr\left\{\max\left(x,y\right) < 0\right\} = \left(\Phi\left(-\frac{\mu}{\sqrt{\pi}}\right)\right)^2 = 0.00038, \tag{30}$$

where Φ is the standard normal CDF.

Figure 4: Comparative statics: welfare-enhancing middlemen ($\lambda = 1$) The case of welfare-enhancing middlemen is summarized in four graphs; they are the same as the graphs in Table 3 except that λ is 1 instead 0. The solid blue line illustrates the benchmark case of no middlemen (**M**) and the dashed red line illustrates the case where **M** are present. The default parameter values for these graphs are: $\lambda=1$, $\mu=4$, $\sigma^2=\pi$, $\theta=3$, and $\alpha=0.9$. First-best welfare is 1 by construction; welfare in the no-**M** case is 29%; welfare in the **M** case is 34% (i.e., an increase of 17% relative to the **M** case). The top-left graph illustrates for what private value x the seller (**S**) decides to post a price himself rather than wait for a bid by **M**, i.e., the region in which $V(x)$ exceeds $U(x)$. The top-right graph illustrates the ask price **S** posts in case he decides to post a price. The bottom-left graph plots the probability of a transfer of the security from **S** to **B**. The bottom-right contour plot graphs the probability of an **S**-**B** transfer as a function of x and y which are the private values of the **S** and **B**, respectively.
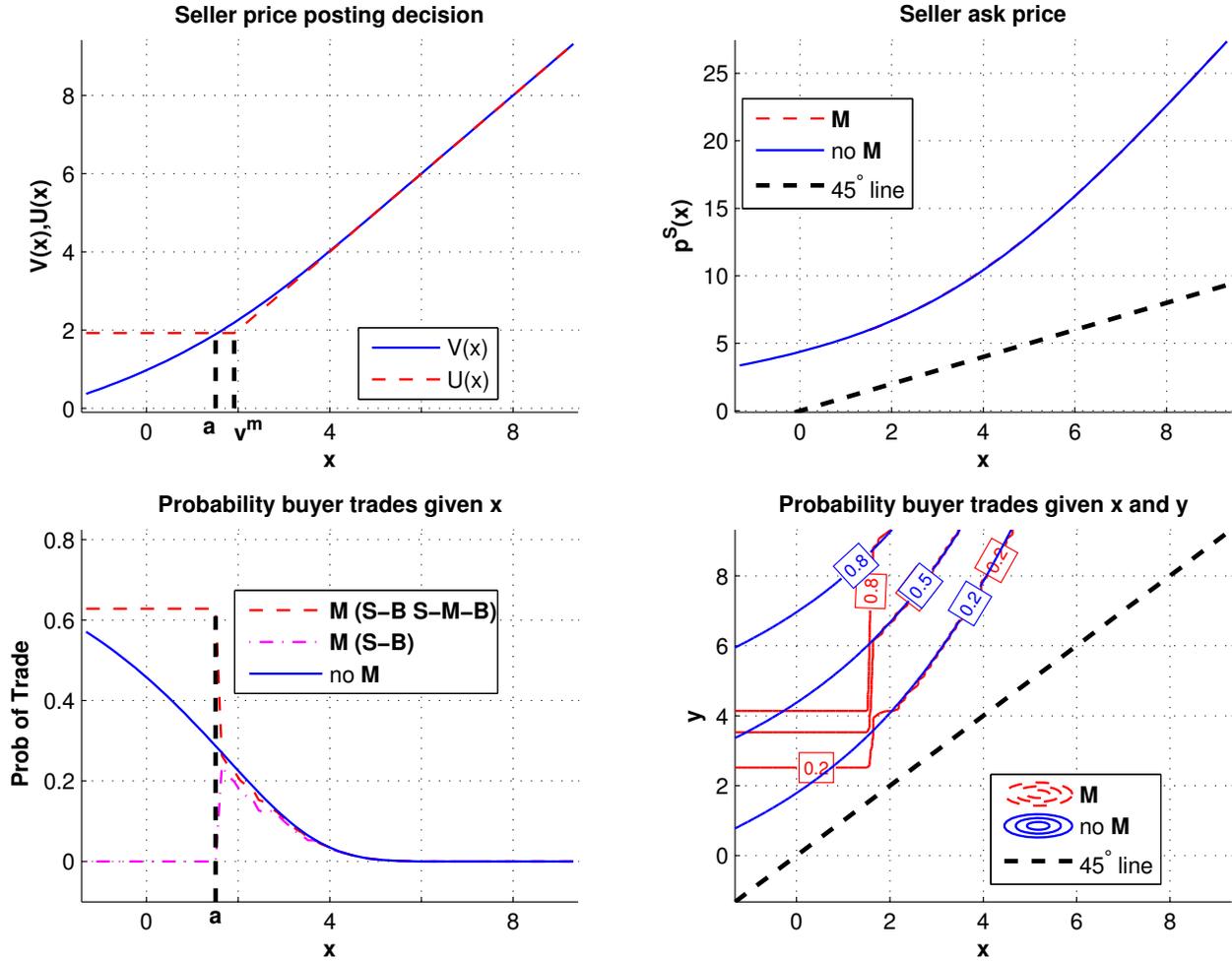

Comparing Figures 3 and 4 yields the following observations.

Top-left panel. $U(x)$ and $V(x)$, and therefore a, remain unchanged when changing the value of λ as the **M** game is unaffected by how much of z^h the buyer knows *ex ante* (as, in the **M** game he infers z^h from **M**'s competitive bids).

20

Top-right panel. If in the game with **M**, **S** decides to post an ask price, i.e., $x \geq a$, the price schedule $p^S(x)$ also remains unchanged across the two extreme λ regimes. In the no-**M** game, however, the price schedule is considerably higher for the fully-informed buyer case as **S** protects himself from being adversely selected. So much so in fact that this price schedule seems to coincide with the **M** case price schedule (red dashed line and blue solid line can hardly be told apart in Figure 4). This is not surprising as, for $\lambda = 1$, **S** is up against fully informed counterparties in both cases. The correspondence is, however, not perfect because the optimization that p^S solves in the no-**M** and in the **M** case ((4) and (20), respectively) are not quite the same; the probability of a trade at any fixed p^S is, for most values of x, slightly higher in the **M** case (because **B**'s acceptance set is unchanged but now also **M** are on the demand side). As a result, p^S is mostly slightly higher in the **M** case although this only becomes visible when one zooms into the curves.

Bottom-left panel. Again, as the **M** case is unaffected by λ the only difference across the two λ regimes is that the probability of trade shifts down (uniformly across all x) in the better informed **B** regime (i.e., $\lambda = 1$). The graph illustrates how some trade is restored in this regime when **M** enter. Not in the region where **S** posts a price himself ($x \geq a$) as his prices are slightly higher than in the no-**M** case (cf. top-right panel). Instead, trade is restored in the region where **S** decides to pass the security off to **M** to get it to **B** ($x < a$). The red dashed line that captures "successful" trades (**S**→**B** and **S**→**M**→**B**) in the **M** case is lower than the blue solid line in the no-**M** case when **S** posts himself ($x \geq a$) and higher when he does not ($x < a$). In fact, in the no-posting region, when **S** accepts **M**'s bid for sure, **M** resells to **B** 63% of the time. Notice that when **S** posts a price, most of the transfers are direct **S** $\rightarrow$ **B** rather than via a middleman **S** $\rightarrow$ **M** $\rightarrow$ **B** (as the purple dash-dot-dash line almost reaches up to the dashed red line).[6]

In summary, **S** seems to consciously use **M** to get the security across to **B** when **S**'s private value is low (i.e., no price quote, accept **M**'s bid for sure) but effectively keeps **M** out when **S**'s private value is high. The reason is that it becomes prohibitively costly to travel via **M** on a high private value as the welfare loss to him when **M** becomes the final owner is high (as he would have realized a high

---

[6]Not shown here is the probability that **M** becomes the final owner in case he snaps up the security by hitting **S**'s price quote; this probability is the same order of magnitude as the probability that **M** snaps up the security and resells to **B** (i.e., **S** $\rightarrow$ **M** $\rightarrow$ **B**).

private value himself had he still had the security); he regrets selling in such case but, obviously, cannot get the security back as he has left the market. The via-**M** detour's ultimate trade-off is: Reduce adverse selection vs. forego private value realization.

Bottom-right panel. The socially desirable trades are those for which the buyer's valuation exceeds the seller's valuation, i.e., $y > x$. A social planner would therefore with certainty move the security from **S** to **B** above the 45 degree line in the (x, y) plane. The bottom-right panel illustrates the actual probability of trade in no-**M** and in the **M** game. Absent middlemen, this probability is much higher in the uninformed-buyer regime ($\lambda = 0$) relative to the informed-buyer regime ($\lambda = 1$). With middlemen, the contour plots are equal across both regimes.

Welfare and λ. The bottom-right panels illustrate that middleman entry reduces the likelihood of trade in the uninformed buyer regime ($\lambda = 0$), but increase it in the informed buyer regime. Trade is a proxy of welfare in this case, but it is not a substitute for it. In this trading game welfare is naturally defined as the gains-from-trade, i.e., the expected realized private value of the final owner:

$$\int x I_{\{\mathbf{S} \text{ owns}\}} + y I_{\{\mathbf{B} \text{ owns}\}} dF(x) \, dF(y) - \mu, \tag{31}$$

where μ is the expected autarky value.[7] The maximum gains-from-trade are equal to one by construction. With middlemen, welfare is 0.34 in both regimes. Without middlemen, it is 0.67 in the uninformed-buyer regime and 0.28 in the informed-buyer regime. Welfare is thus substantially reduced when middlemen are introduced in the $\lambda = 0$ regime (adverse selection is aggravated) whereas it is increased on middleman entry in the $\lambda = 1$ regime (adverse selection is mitigated). Inspection of intermediate levels of λ show that welfare is approximately linear in λ. The critical value of λ above which **M** entry starts to increase welfare is 0.79.

---

[7]**M**'s private value is zero and therefore does not appear in this expression.

2.4 The effect of M entry on trade, adverse selection cost, and welfare

This subsection calculates the (expected) trade frequency, the cost of adverse selection on price quotes, and welfare, all as a function of the model's primitive parameters. The first two variables will be used to match their equivalents in the data. The reason to match these two is that (i) they capture the model's most salient features (trade as it proxies welfare — realized gains-from-trade — and adverse selection as it captures the model's main friction) and (ii) their sample counterparts are easily computed, standard trade variables widely used in the empirical microstructure literature.

2.4.1 The game without middlemen

Trades. The expected number of trades in the no-**M** case is:

$$T_{NM} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(1 - G^{\lambda}\left[p^{S}(x) - y\right]\right) dF(x) dF(y),$$

where NM is the label for the no-**M** case.

Trade-weighted adverse selection. Mathematically, the expected adverse selection is $E(z \mid \text{trade})$. In the no-**M** case, it is:

$$\frac{1}{T_{\text{NM}}} \int z \left[1 - F(p(x) - z)\right] dG^{\lambda}(z) dF(x). \tag{32}$$

Welfare. Welfare is the expectation of the sum of each investors' private value interacted with a dummy that is one if the investor becomes the final owner, zero otherwise:

$$W_{NM} = E\left(x I_{\{\mathbf{S} \text{ owns}\}} + y I_{\{\mathbf{B} \text{ owns}\}}\right)$$
$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \left[x G^{\lambda}\left[p^{S}(x) - y\right] + y \left(1 - G^{\lambda}\left[p^{S}(x) - y\right]\right)\right] dF(x) dF(y)\right). \tag{33}$$

The common value z drops out as it is a transfer.

2.4.2 The game with middlemen

Trades. If **S** posts an ask price (i.e., $x \geq a$), **M** has the first take on it (as **M** is fast) and gets the security iff the price is below **M**'s reservation value, i.e., iff

$$p^S < z^h + v^M. \tag{34}$$

As $p^S(x)$ is strictly increasing in x and thus invertible, the probability that **M** gets the security is $F\left(\left(p^S\right)^{-1}\left(z^h + v^M\right)\right)$, where $\left(p^S\right)^{-1}$ denotes the inverse function of $p^S(x)$.

If **S** posts an ask ($x \geq a$), the number of trades is:

$$T_A\left(x, y, z^h, z^s\right) = \begin{cases} 0 & \text{if} \quad p^S(x) > \max\left(z^h + v^M, y + z^h + z^s\right), \\ 1 & \text{if} \quad \begin{cases} p^S(x) > z^h + v^M \quad \text{and} \quad p^S(x) < y + z^h + z^s, \\ p^S(x) < z^h + v^M \quad \text{and} \quad \hat{p}^M > y + z^h + z^s, \end{cases} \\ 2 & \text{if} \quad p^S(x) < z^h + v^M \quad \text{and} \quad \hat{p}^M < y + z^h + z^s. \end{cases} \tag{35}$$

If **S** does not post an ask ($x < a$), the number of trades is:

$$T_A\left(x, y, z^h, z^s\right) = \begin{cases} 0 & \text{if} \quad x > v^M, \\ 1 & \text{if} \quad x < v^M \quad \text{and} \quad \hat{p}^M > y + z^h + z^s, \\ 2 & \text{if} \quad x < v^M \quad \text{and} \quad \hat{p}^M < y + z^h + z^s. \end{cases} \tag{36}$$

The expected total number trades therefore is the (unconditional) expectation over the four mutually independent shocks:

$$T_M = E\left(T_A + T_A\right) = \int \left(T_A + T_A\right) dF(x) \, dF(y) \, dG^h\left(z^h\right) dG^s(z^s),$$

where M indicates the **M** case. The only trades that **M** do not participate in are those for which $z^h + v^M < p^S(x) < y + z^h + z^s$.

Trade-weighted adverse selection. In the **M** game, there is no adverse selection when $x < a$ and **S** accepts **M**'s bid, but on all other trades there is adverse selection irrespective of whether **M** or **S** posted the ask price. The expected adverse selection cost ($E(z \mid \text{trade})$) therefore is:

$$\frac{1}{T_M} \int \left\{ \begin{array}{c} F(a)\left(z^s + z^h\right) I_{\{y+z^s+z^h>p+z^h\}} \\ + \int_a^\infty \left(z^s + z^h\right) \left[I_{\{p^S(x)<z^h+\max\left(v^M, y+z^s\right)\}} + I_{\{p^S(x)<z^h+v^M \cap p<y+z^s\}} \right] dF(x) \end{array} \right\} dG^s dG^h dF. \quad (37)$$

Welfare.   The common value z itself drops out of the welfare equation (as all agents value it equally) but its two components critically determine who ends up with the security. Welfare itself is a weighted average of welfare when **S** posts ($x \geq a$) and welfare when **S** does not post ($x < a$). Let us start with the easier second case. The expected welfare of for $x < a$ is:

$$W_{\text{NP}} = \int_{\{\sim A \cap (z^h+v^M, \infty)\}} x \, dF(x) \, dG^h\left(z^h\right) +$$
$$\int_{\{\sim A \cap \left(-\infty, z^h+v^M\right)\}} dF(x) \int_{-\infty}^\infty \left(\int_{p-z^s} y \, dF(y) \right) dG^s(z^s) \, dG^h\left(z^h\right), \quad (38)$$

where the label *NP* indicates that this is the no-posting region (vs. *P* used below for the posting region) and p is defined in (14). If **M** ends up with the security welfare is zero.

The contribution to welfare when **S** posts an ask ($x \geq a$) is:

$$W_{\text{P}} = \int_{\left\{A \cap \left(p^S(x)>\max\left(z^h+v^M, y+z^s+z^h\right)\right)\right\}} x \, dF(x) \, dF(y) \, dG^s(z^s) \, dG^h\left(z^h\right) + D,$$

where D (given in (39) below) is the expected welfare obtained when **S** posts and then parts with the security, which happens if $x \in \left\{ A \cap \left(p^S(x) < \max\left(z^h + v^M, y + z^s + z^h\right)\right) \right\}$. Welfare, y, is collected in the following two mutually exclusive events:

$$\left(p^S(x) \leq z^h + v^M \quad \text{and} \quad y + z^s + z^h \geq \hat{p}^M\right) = D_1\left(x, y, z^h, z^s\right),$$

with $\hat{p}^M = z^h + p$ and with p defined in (12) or

$$z^h + v^M < p^S(x) \leq y + z^s + z^h = D_2\left(x, y, z^h, z^s\right).$$

Therefore:

$$D = \int y I_{\{A \cap (D_1 \cup D_2)\}} dF(x)\, dF(y)\, dG^h\left(z^h\right) dG^s\left(z^s\right). \tag{39}$$

Thus, welfare for the case with middlemen is:

$$W_M = W_P + W_{NP}. \tag{40}$$

3   Model calibration

The empirical part of the paper is an event study on the introduction of middlemen — high-frequency traders — in a standard limit order market. The objective is three-fold.

First, we search for the arrival of middlemen around a discrete change in the trading environment which, as a result, became more HFT-friendly. We discover the arrival of a large HFT and provide some trade statistics to characterize its behavior. This part serves to verify the assumptions that have gone into the theoretical model.

Second, we study whether an important idiosyncratic prediction of the model is supported by the data. Time variation in the sample is used to verify whether middlemen participate more on high α days, i.e., days with a high proportion of hard information (relative to total information).

Third, we gauge the net welfare effect of middleman arrival based on a calibration of the model. A middleman treatment effect is identified through a diff-in-diff analysis. The post-entry trade sample is compared to a pre-entry sample and this time differential is compared with the time differential of a sample that did not see the entry of middlemen (as it did not experience the discrete environment change). The focus variables in this diff-in-diff analysis are the key trading variables that have a natural equivalent in the model: trade frequency and the adverse selection cost of posting prices. The model's deep parameters are calibrated to match the changes in these variables. Once calibrated, the model-implied welfare effect of middleman entry is calculated.

The results on these three objectives are presented in the next three subsections. The first two

subsections are relatively brief. They summarize the key findings. The key empirical facts that enter the model calibration are reported in the third subsection, which mostly focuses on the actual calibration to these empirical facts. The full set of empirical results is in Section 1 of the online appendix (http://goo.gl/40Cnyd) and is structured the same way, i.e., three subsections lay out the results that speak to the three stated objectives.

3.1   Identifying a middleman and characterizing its trading behavior

The middleman identified in the data has the following features. First, the middleman trades mostly passive. For four out of five of his trades he was the limit order that was taken out of the book. This is in line with theory as he endogenously participates in the trading game only through the quotes he submits (see the discussion of Figures 3 and 4). If the seller chooses not to take the middleman quote but post a price himself instead, then he sets his ask price so high that the middleman rarely adversely selects it. Second, middleman quotes seem a lot more responsive to hard information arrival than non-middleman quotes. These two findings correspond to analysis done in subsections 1.2.1 and 1.2.2 of the online appendix, respectively.

3.2   Hard information and middleman participation

This subsection exploits variation in the stock-day panel dataset to study one of the theoretical model's main predictions: investors endogenously choose to involve middlemen more when α is high, i.e., hard information is relatively important (see Proposition 3 and its Corollary). The natural way to test this prediction is to plot middleman participation in trades against variation in the relative size of hard information. The α proxy for a stock-day is how important index information is for total information on that day for that stock, i.e., the R^2 of an intraday single-factor CAPM regression. A straightforward OLS regression is impossible due to an errors-in-variables problem. A cointegration model is used to identify and remove any transitory ("errors") effects. Details are

27

Figure 5: Scatter plot of middleman activity versus amount of hard information. This figure contains a scatter plot of a proxy for the relative size of hard information, α, and middleman participation in trades for all stock-day observations in the sample. In the data, an important source of hard information is the index-futures market. A proxy for α is obtained by essentially regressing a stock's transaction return on the contemporaneous index-futures return. This regression is implemented through a cointegration model so as to identify and remove transitory effects in both series (details are in the online appendix). The proxy for α is defined as $\frac{var(\hat{P}_m(\Delta f))}{var(\Delta f)}$. For each stock, both series are demeaned to remove cross-sectional heterogeneity, i.e., the "fixed effect" is removed. The scatterplot also graphs a regression line and the calibrated model's implied relationship. This implied relationship is re-centered to match the data, i.e., the average single-factor CAPM R^2 in the data (0.39 c.f. Table 4) is subtracted from the α series and its corresponding **M** participation value is subtracted from the **M** participation series. Finally, we note that the covariation between the α proxy and middleman participation was not used in the calibration.
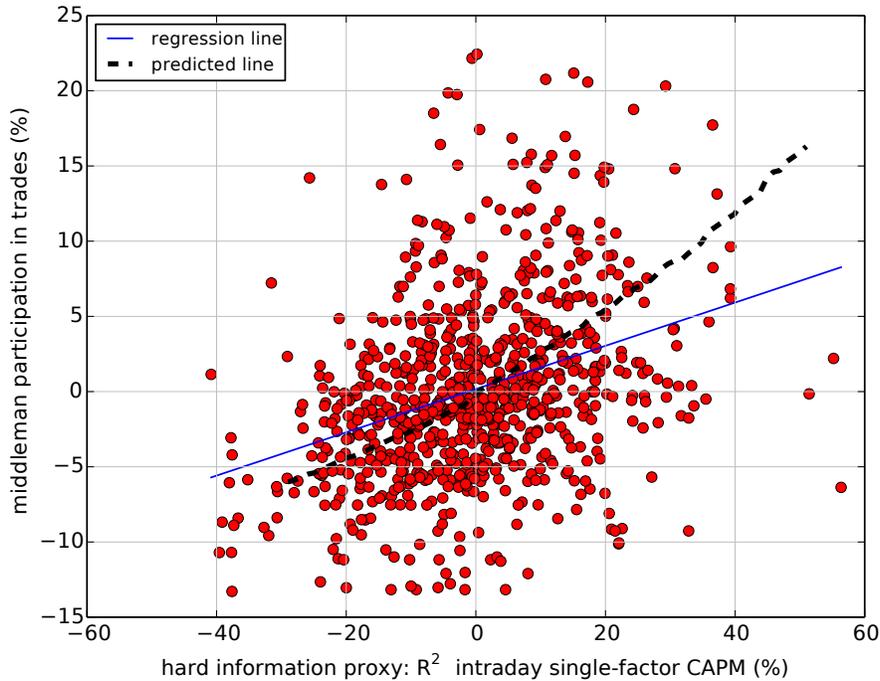
in the online appendix.

The scatter plot in Figure 5 provides empirical support for the model's prediction that middlemen participate more when hard information is relatively more important. The intraday R^2 shows substantial variation through time as its range is almost 100%. The variation in middleman participation is also substantial as its range is roughly 40%. The scatter cloud indicates a positive correlation which is supported by the regression line. The relationship is economically meaningful as moving from a no-index information day to a full-index information day has the middleman participate almost 15 percentage points more. This change is somewhat less than the model-implied

change (dashed line) when varying only α and keeping the other parameters at their calibrated values (see Section 3.3).

3.3 Model calibration to identify welfare effect of middleman entry

This subsection first calibrates the theoretical model to fit the data's salient features and then uses the calibration to compute the net welfare effect of middleman entry. Consistent with Figures 3 and 4, investors' private values are assumed to be independent and normally distributed: $x, y \sim N(\mu, \sigma^2)$ where $\sigma^2 = \pi$ (so that all results are in units of the maximum gains-from-trade, see Section 2.3). The following four parameters are estimated:

- $\lambda \in [0, 1]$: the probability of **B** being informed about hard information in the no-**M** case.

- $\alpha \in [0.39, 1]$: the relative size of hard (common) value information relative to total common value information. The lower bound is empirically motivated as it is derived from **M**'s ability to parse the index-futures market information instantly (see Section 1.2.2); the index is, on average, 39% of intraday stock returns. Middlemen might use additional sources of hard information which makes the 0.39 a lower bound.

- $\theta = \sigma_z^2/\pi \geq 0$: common value variance relative to private value variance (the latter is fixed at π).

- μ: the mean of both investors' private value distribution; it matters only in the **M** case as it determines how much more investors value the security relative to middlemen (which have zero private value for the security).

The targeted values for these parameters (and what has been achieved) are summarized as (the tables referenced here are in the online appendix):

	Targeted	Reason	Achieved
Δ adverse selection cost ($E(z \mid \text{trade})$)	-23%	Table 6, line 4	-19%
Δ trade frequency	$+17\%$	Table 6, line 7	$+22\%$
Middlemen participation	$\geq 14.3\%$	Table 2, line 10	28%
Hard-information fraction (α)	≥ 0.39	Table 4, line 3-4	0.39

The criterion function that was minimized is:

$$\text{Criterion Function} = (\Delta \text{ adverse selection cost} + 0.23)^2 + (\Delta \text{ trade frequency} - 0.17)^2, \quad (41)$$

subject to the middlemen participation and hard-information constraints (see target diagram above). The parameter estimates are:

	Pre-set	Calibrated			
Parameter	σ^2	λ	α	θ	μ
Value	3.14	1	0.39	0.58	3.35

M entry and welfare. At the calibrated parameter values, middleman entry raises welfare by 1% of to the maximum gains-from-trade.[8]

How certain can we be of the λ's high value which is at the root of the welfare gain? That is, the discussion surrounding Figures 3 and 4 illustrates that middlemen can only be benevolent if there is an adverse selection problem to begin with, i.e., **B** is informed. It turns out that when λ is small and θ large, the model explains the data almost equally well. Figure 6 illustrates this point by plotting the contours of the criterion function (41) in (λ, θ) space. The minimized criterion is 0.003, in the bottom-right region. Similar but slightly higher values are obtained in the top-left region. In this region, however, the market functions well even without **M** because **B** is unaware

---

[8]At the calibrated parameter values, the direct effect on welfare of adding **M** as an additional candidate agent to hold the security (see the calculation in (30)) remains negligible as the probablity that **M** has the highest value for the security is $\left(\Phi\left(-\frac{3.35}{\sqrt{\pi}}\right)\right)^2 = 0.0092$. Therefore any welfare effect is due to **M**'s role in intermediation.

Figure 6: Contour plot of the criterion function used in the calibration. This figure depicts the contour plot of the criterion function that was minimized to calibrate the theoretical model (see discussion around (41)). The parameter estimates obtained this way are: λ=1.00, μ=3.35, σ^2=π, θ=0.58, and α=0.39. They are represented by the yellow dot where the minimized criterion is 0.003. The contour plot illustrates how the criterion function depends on the model's two main parameters: λ which represents how likely it is that the late investor knows about hard information and θ which represents how large common value variance is relative to private value variance.
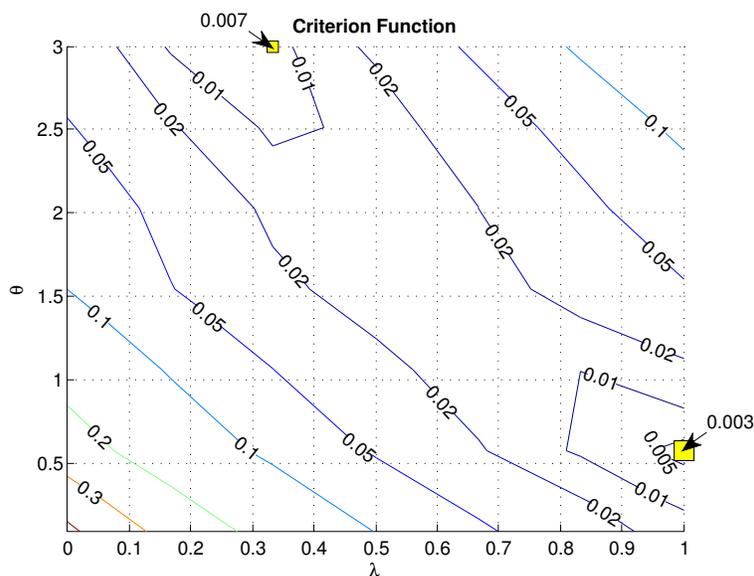

of z^h and, as has become clear from the Figure 3 discussion, this is a situation where **M** reduce welfare. Our welfare conclusion that **M** raise welfare can therefore only be tentative. On the other hand, in the top-left region, θ is unrealistically high given the evidence in Hollifield *et al.* (2006).[9] We therefore take our estimate seriously and conclude that it is indeed quite likely that **M** did raise welfare.

Figure 7 illustrates the model in other dimensions and, essentially, replicates the Figures 3 and 4 for the calibrated parameters. Not surprisingly, it resembles Figure 4 more than Figure 3.

Welfare change and technology cost. This final part compares the model-implied welfare change due to middleman entry with what middlemen invest in order to be part of the game. Such cost was

---

[9]The authors estimate a structural limit order model for stocks at the Vancouver Stock Exchange. Their estimate of investors' private value standard deviation is 21% (Table VII). This is very high relative to the standard deviation of the *full day* stock return — common value innovation — which, in their sample, is estimated at 7% (Table II). Their results therefore imply a θ value that is well below 1.

Figure 7: Comparative statics calibrated model. These graphs illustrate the middlemen equilibrium for the cali-brated paramaters; the same graphs are the same as in Figure 3 and 4 except that the parameters are calibrated. The solid blue line illustrates the benchmark case of no middlemen (**M**) and the dashed red line illustrates the case where **M** are present. The parameter values for these graphs are: $\lambda=1.00$, $\mu=3.35$, $\sigma^2=\pi$, $\theta=0.58$, and $\alpha=0.39$. First-best welfare is 1 by construction; welfare in the no-**M** case is 59% of what first-best would achieve; welfare in the **M** case is 60% (i.e., an increase of 1% relative to the **M** case). **M** participation rate is 24%. The top-left graph illustrates for what private value x the seller (**S**) decides to post a price himself rather than wait for a bid by **M**, i.e., the region in which $V(x)$ exceeds $U(x)$. The top-right graph illustrates the ask price **S** posts in case he decides to post a price. The bottom-left graph plots the probability of a transfer of the security from **S** to **B**. The bottom-right contour plot graphs the probability of an **S**-**B** transfer as a function of x and y which are the private values of the **S** and **B**, respectively.
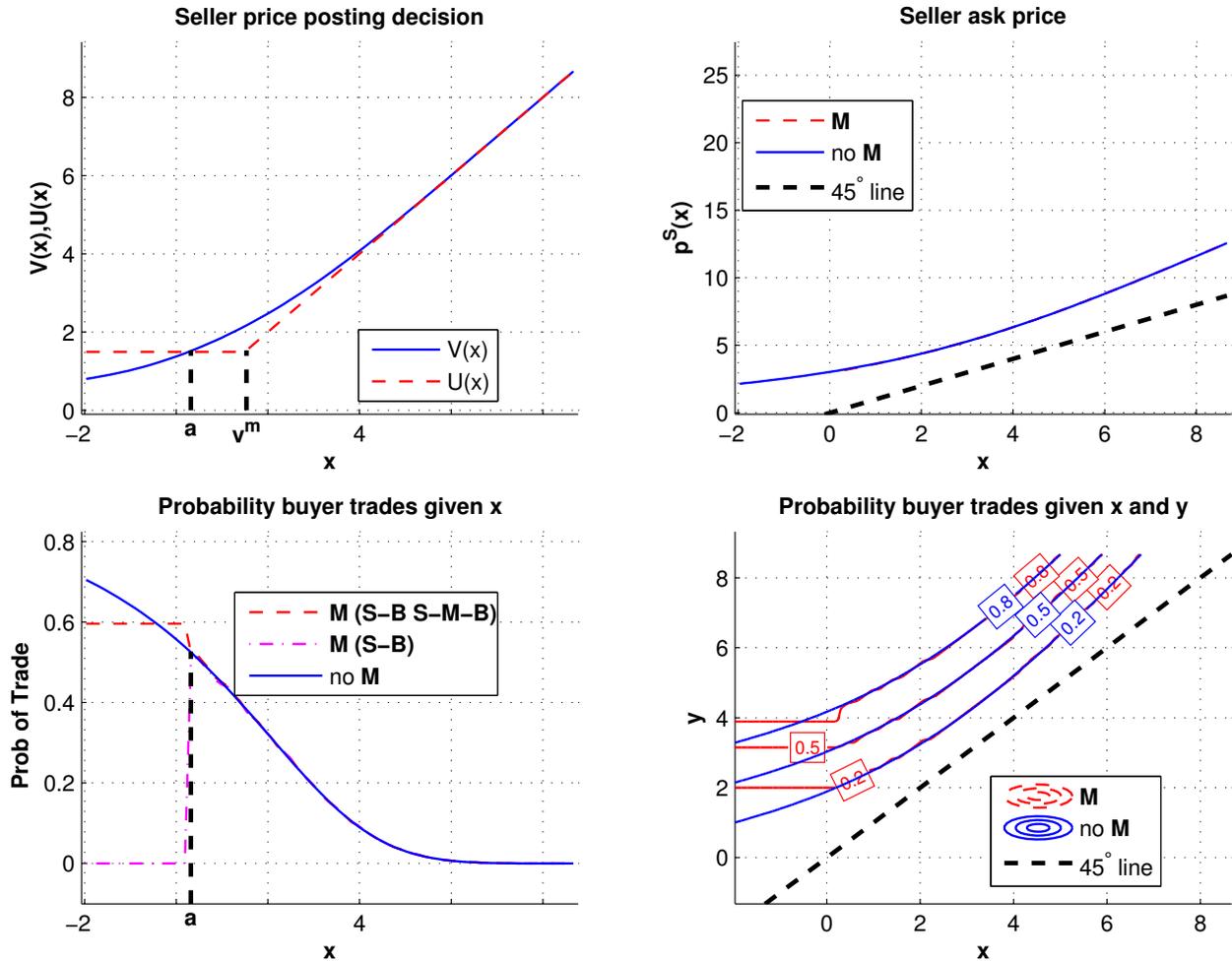
out of scope thus far but should be accounted for in welfare judgments. To compare investment cost with the calibrated welfare change, we need to also express it relative to expected gains-from-trade (which was normalized to one in the model). Figure 8 plots welfare change due to **M** entry along with technology cost as a function of the fraction of hard information (α). The technology

Figure 8: Welfare change and technology cost. This graph plots welfare change due to middleman entry along with the technology cost middlemen incur to participate in the game. Both series are expressed as a percentage of the total gain-from-trade (which was normalized to one in the model). They are plotted as a function of the proportion of hard information in the common value innovation occurs between the seller and buyer arrival: α. The technology cost is calculated based on estimates from Budish, Cramton, and Shim (2015) and securities markets volume data. Total cost depends on the number of middlemen participating (N). In addition to the Budish, Cramton, and Shim (2015) estimate of 20 middlemen, we also consider $N=1$, 50, and 100. The welfare change is calculated based on the model calibration: $\lambda=1.00$, $\mu=3.35$, $\sigma^2=\pi$, $\theta=0.58$, and $\alpha=0.39$.
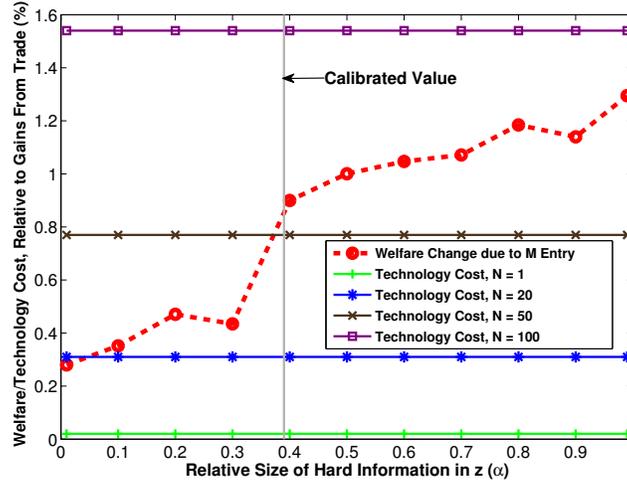
cost estimate is based on the following data:

- Middlemen spend \$50 million annually on technology. This estimate was obtained from Budish, Cramton, and Shim (2015).

- Results are obtained for $N=20$ middlemen (the Budish et al. estimate), but also for $N=1$, 50, or 100.

- World-wide equity volume is approximately 5% of total volume (i.e., equity, options, futures, and foreign exchange).

- Total world-wide equity volume is €6.4 billion. The numbers thus far deliver the technology cost per euro traded.

- The gains-from-trade per euro traded is derived from the calibration. Panel B of Table 4 shows that the variance of z is 10.12 bps^2. θ is calibrated to be 0.58 which implies that the variance of private value variance is 10.12/0.58=17.45 bps^2. The private value variance in

the model was set to π so that expected gains-from-trade equal one. The expected gains-from-trade per euro traded in the data therefore is $\sqrt{17.45/\pi}$=2.36 bps.

The graph indicates that for the calibrated value of hard information (α=0.39) and the number of middlemen assumed in Budish et al. (N=20), middleman entry benefits trading as the net social gain is 0.90-0.77=0.13% of the total gains-from-trade. This result holds for any α above 0.05. The social value of middlemen is upward sloping in hard information, as that is the part of the friction in trade between the seller and the buyer that they can help overcome. We further find that entry of middlemen is always beneficial in the case of a single middleman, only for α above roughly a third for 50 middlemen, and never for 100 middlemen.

Finally, what does the calibrated model imply for the critical value of λ beyond which middleman entry benefits trading? Such value exists as the welfare change due to **M** entry increases monotonically in the probability that the buyer observes hard information (λ). Disregarding technology cost the critical value is 0.85. For the twenty middleman technology cost case, it rises to 0.90. The beneficial result of **M** entry therefore hinges critically on there being a severe adverse selection friction on hard information between the seller and the buyer to begin with (absent middlemen). The calibration shows that there is, as λ was estimated to be one.

4   Double auction as an alternative mechanism

How well does middleman mediated trading compare to other feasible mechanisms? We now analyze a mechanism that is feasible as long as the designer can transfer the asset and the payment after **S** is no longer present. This mechanism is a double auction, and we describe it in more detail.

A feasible mechanism is the Chatterjee and Samuelson (1983) type of double auction which Myerson and Satterthwaite (1983) proved was *ex-ante* optimal in the case where there were no common values and where the private values were uniformly distributed. Since we have a common value

and since (x, y, z) are normally distributed, we do not know if a double auction is the optimal mechanism here, but it does outperform the middleman mechanism. Let a be the seller's ask and b the buyer's bid. The balanced-budget mechanism transfers the asset from **S** to **B** iff $a \leq b$ and in the event of a transfer, **S** receives from **B** the payment

$$p^k_{a,b} = (1 - k) a + kb. \tag{42}$$

The take-it-or-leave-it case that we studied under the no-**M** regime arises if $k = 0$.

Since we estimate $\lambda = 1$, we shall only analyze this case. Then **B**'s strategy depends only on $y + z$. Let $b(y + z)$ be the optimal bid. Then the optimal ask $a(x)$ is

$$a(x) = \arg\max_a \left\{ \int_{\{a \leq b(y+z)\}} p^k_{a,b(y+z)} dF(y) \, dG(z) + \int_{\{a > b(y+z)\}} (x + z) \, dF(y) \, dG(z) \right\} \tag{43}$$

and the optimal bid is

$$b(y + z) = \arg\max_b \left\{ \int_{\{a(x) \leq b\}} \left(y + z - p^k_{a(x)b} \right) dF(x) \right\}. \tag{44}$$

For each $k \in [0, 1]$ (on some grid) we calculate the optimal strategies $a(x)$ and $b(y + z)$. Let $\eta = y + z$. The CDF of η therefore is $\Psi(\eta) = \Phi\left(\frac{\eta - \mu_y}{\sqrt{\sigma^2 + \sigma_z^2}} \right)$, where Φ is the standard normal CDF. The density is

$$\psi(\eta) = \frac{1}{\sqrt{\sigma^2 + \sigma_z^2}} \phi\left(\frac{\eta - \mu_y}{\sqrt{\sigma^2 + \sigma_z^2}} \right) = \frac{1}{\sqrt{2\pi \left(\sigma^2 + \sigma_z^2 \right)}} \exp\left(\frac{\left(\eta - \mu_y \right)^2}{2 \left(\sigma^2 + \sigma_z^2 \right)} \right).$$

Then (43) and (44) simplify to

$$a(x) = \arg\max_a \left\{ \int_{\{a \leq b(\eta)\}} p^k_{a,b(\eta)} \psi(\eta) \, d\eta + \int_{\{a > b(y+z)\}} (x + z) f(y) \frac{1}{\sigma_z} \phi(z) \, dy \, dz \right\} \tag{45}$$

(note that z still appears on the RHS) and

$$b(\eta) = \arg\max_b \left\{ \int_{\{a(x)\leq b\}} \left(\eta - p_{a(x),b}^k\right) dF(x) \right\} \tag{46}$$

The FOCs generate two ODEs from which the optimal strategies can be calculated numerically. An alternative is to approximate the integrals numerically and optimize these using a numerical maximizer. We implemented the latter method in Matlab. Further details are in Appendix C.

Welfare under the mechanism. Since $p_{a,b}^k$ is a transfer, this function enters total welfare only through its effect on the probability of trade which occurs on the set of points (x, y, z) at which $a(x) \leq b(y + z)$. We suppress from the notation the dependence of the strategies on the parameter k. Moreover, x, y, and z are independent of each other and, hence, so are x and η. Total *ex-ante* welfare is

$$U_k = \int_{\{a(x)\leq b(y+z)\}} y f(x) f(y) \frac{1}{\sigma_z} \phi\left(\frac{z}{\sigma_z}\right) dx dy dz + \int_{\{a(x)>b(\eta)\}} x \psi(\eta) f(x) d\eta dx$$

where $\phi(s)$ is the standard normal density. Since z is a common value, it does not affect welfare conditional on a trade; it affects welfare only indirectly by affecting the probability of trade. The welfare levels of **S** and **B**, however, do include the transfer. So, we write $U_k = U_k^{\mathbf{S}} + U_k^{\mathbf{B}}$ where the welfare of **S** is

$$U_k^{\mathbf{S}} = \int_{\{a(x)\leq b(\eta)\}} p_{a(x),b(\eta)}^k f(x) \psi(\eta) dx d\eta + \int_{\{a(x)>b(y+z)\}} (x + z) f(x) f(y) \frac{1}{\sigma_z} \phi\left(\frac{z}{\sigma_z}\right) dx dy dz$$

and the welfare of **B** is

$$U_k^{\mathbf{B}} = \int_{\{a(x)\leq b(y+z)\}} \left(y + z - p_{a,b(y+z)}^k\right) f(x) f(y) \frac{1}{\sigma_z} \phi\left(\frac{z}{\sigma_z}\right) dx dy dz.$$

The counterpart of the gains from trade in (31) is the welfare that the mechanism delivers relative to financial autarky: $u_k = U_k - \mu_x$. This is to be compared to the welfare gains in the no-**M** and **M**

cases respectively, i.e., to W_M and W_{NM}, defined in (33) and (40).

Figure 9: Double auction as an alternative mechanism. This graph illustrates auction as an alternative mechanism. The seller posts an ask price and the buyer posts a bid price. The key difference with the baseline (continuous) limit order market is that the seller's ask remains hidden for the buyer who can therefore not condition on it. The auctioneer will only effectuate trade if the buyer's bid is above the seller's ask. The transaction price equals $(1 - k)a + kb$ where a is the ask price and b is the bid price. There are no middlemen in the auction. The graphs illustrate the outcome of the auction at the calibrated parameters. The top-left graph plots welfare for the auction and compares it to the outcome of the baseline model, both the middleman and the no-middleman case. The top-right graph decomposes this welfare into buyer and seller utility. The bottom-left graph illustrates the strategies of both the seller and the buyer for $k = 0$. The seller's ask depends on her private value draw: x. The buyer's bid depends on the common value innovation he see plus his private value: $\eta = z + y$. The bottom-right graph replots this graph for $k = 1$. The calibrated parameter values are: λ=1.00, μ=3.35, σ^2=π, θ=0.58, and α=0.39.
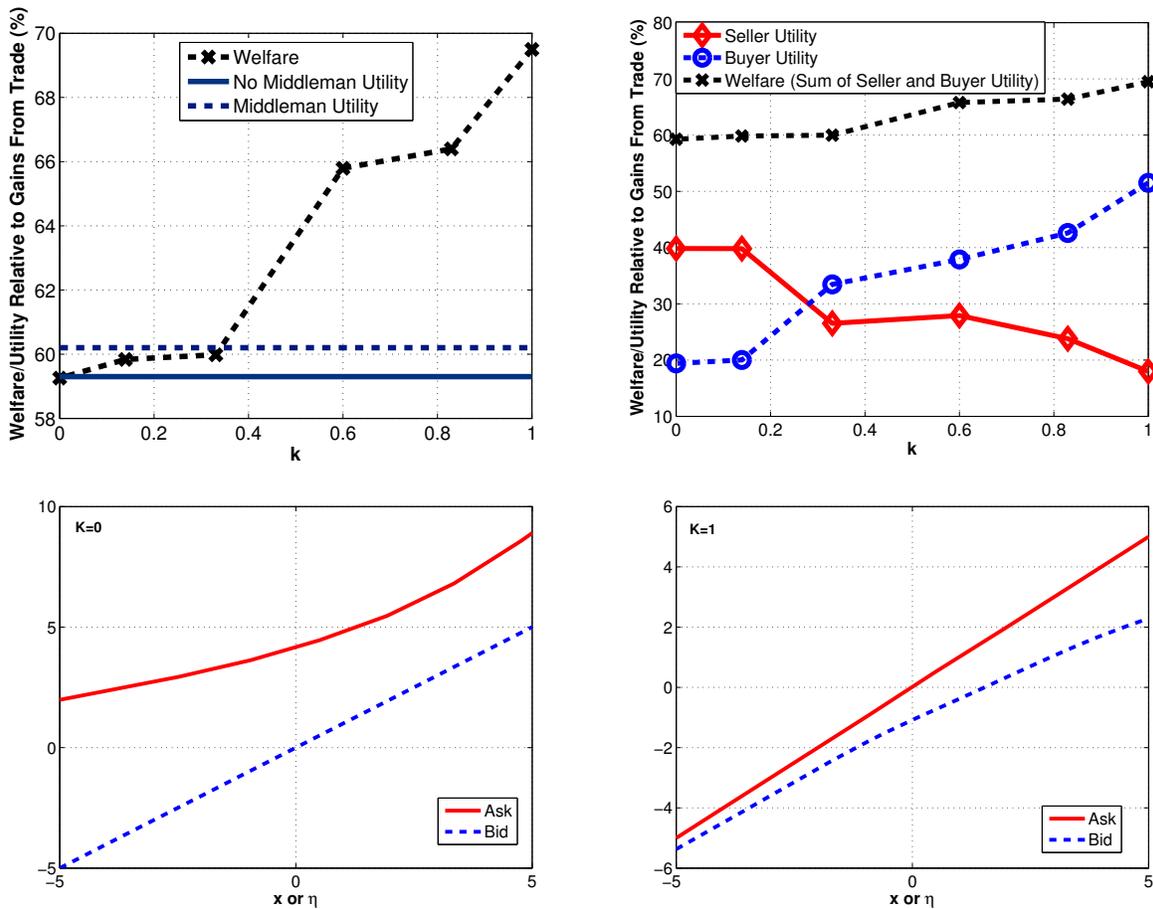
The double auction strategies and the outcome are plotted in Figure 9. They are computed as a function of the transaction price parameter k, see (42). The primitive parameters are set at their calibrated values as reported on p. 30.

First off, at $k = 0$, **S** makes a take-it-or-leave-it offer, and we already studied this under the no-**M** regime. Based on that, we should have

$$u_0 = W_{\text{NM}} = 0.59 \tag{47}$$

and we see in top-left graph that this is indeed the case. We further verified that **B**'s strategy in the $k = 0$ case is to set $b(\eta) = \eta$. At $k = 1$ we also verified that the algorithm yields the outcome in which **B** makes a take-it-or-leave-it offer and in which **S**'s strategy is to set $a(x) = x$. The bottom-left and the bottom-right graphs report both players' strategies for these two limiting cases.

The top-right graph shows that a transaction price closer to the bid (higher k) benefits the buyer. This result seems counter-intuitive as trade only takes place when the bid is above the ask and, keeping strategies fixed, this would hurt **B**. He will have to pay a higher price. But, price-posting strategies change and **S** starts to post more aggressively (lower asks) as the transaction price depends less on the ask. In the extreme case of $k = 1$ **S** posts his reservation value. **B**, on the other hand, shades his bid more as the transaction price depends more on his price. The net effect is that **B** benefits. This result is consistent with Chatterjee and Samuelson (1983, p. 844) who analyze the case without a common value. With information asymmetry about the common value though, our analysis reveals that welfare (the sum of **B** and **S** utility) increases the closer the transaction price is to the limit order price of the informed agent, in this case **B**. This result differs from Chatterjee and Samuelson (1983) who show that welfare is highest for the split-the-difference case, i.e., $k = 1/2$.

Our double-action results contribute to the debate on whether continuous markets should be replaced by frequent batch auctions (Farmer and Skouras, 2012; Budish, Cramton, and Shim, 2015; Wah and Wellman, 2013).[10] In particular, our results suggest that if multiple transaction prices effectuate the same trade, one should consider picking the one that is closest to the most informed agent. This would be good for welfare. Trade history might be used to determine how informed each participant is.

---

[10]One of the largest trade venues for currencies, ICPS's EBS, took a step in this direction. It started to batch incoming orders and randomize their arrival sequence. See "Exclusive: EBS Take New Step to Rein in High-Frequency Traders," Wanfeng Zhou and Nick Olivari, Reuters, August 23, 2013.

5 Robustness: Monopoly middleman model and calibration

The data used to calibrate the competitive model featured entry of a middleman with a participation in trade of 14.3% (see Table 2). Our data are imperfect and therefore the fact that we could only identify one middleman in the data does not rule out that others entered as well. Given the large trade share of the identified middleman it seems prudent to also develop the monopoly middleman model, calibrate its equilibrium outcome, and recompute the welfare result. We take up such task in this section. The additional benefit of analyzing this other polar case — perfect monopoly instead of perfect competition — is that it yields further insight into what it means to add a middleman to a limit order market.

5.1 Monopoly middleman model

The key differentiating feature of the monopoly middleman model is that **M** now "controls" the beliefs on z^h of both the seller and the buyer through his bid function. In the competitive case he could not do so as Bertrand bidding forced him to reveal his information on z^h by bidding $z^h + v^m$ where v^m is a constant (see (15)). In the monopoly model, however, a separating equilibrium where each value z^h leads to a different bid can only exist if for any *given* value of z^h that **M** privately observes, he bids $p(z^h)$ and does not want to deviate. It turns out that such separating equilibrium exists. The problem is formally defined in Section 2 of the online appendix (http://goo.gl/40Cnyd) which shows that the **M**'s FOC yields the following differential equation for **M**'s bid $p(z)$ (the superscript h for z is omitted in the interest of legibility).

$$0 = F(p - z)\left(\frac{\partial V}{\partial Z}(z, z)\frac{\mathrm{d}z}{\mathrm{d}p} - 1\right) + f(p - z)\left(\frac{\mathrm{d}z}{\mathrm{d}p} - 1\right)(p - V(z, z)) \tag{48}$$

where, as **M**'s incentive to signal to the buyer is proportional to $1 - \lambda$,

$$\frac{\partial V}{\partial Z}(z, z) = \int_{-\infty}^{\infty}(1 - \lambda)yg^s(p - y - z)\,\mathrm{d}F(y). \tag{49}$$

39

The differential equation reveals the trade-off that the middleman faces. Suppose he observes that the true value of z^h is z'. He will bid $p(z')$ and considers deviating. The upside of raising the bid price by, say, dp is that the buyer will believe it is worth more and **M** will thus earn an additional $(\partial V/\partial Z) \times (dz/dp) \times dp$ minus the dp he pays more, but all of this has to be multiplied by the probability that **M** gets the asset from the seller which is $F(p-z)$. This corresponds, apart from the dp factor which will drop out, to the left-most term in (48).

The downside of raising the bid price by dp is that the seller will believe that the asset is worth more and is therefore less likely to sell. This probability of selling is reduced by $f(p-z) \times (dz/dp-1) \times dp$ (i.e., the "probability" that the seller's private value draw x falls is in the marginal region, $f(p-z)$, times how much value the seller believes is in the deal: $(dz/dp-1) \times dp$). In this case, **M** foregoes the value of having the asset which is: $p - V(z,z)$. The product of all these terms corresponds to the right-most term in (48), again apart from the dp factor which drops out.

5.2   Calibration monopoly model

We calibrate the monopoly middleman model to the data by using an extended criterion function. In the competitive model, the model-implied middleman participation rate had to exceed a lower bound of 14.3%. This 14.3% is how much the identified middleman participated in trade (given the imperfect data there could have been other middlemen who we failed to identify). In the monopoly model however the implied participation rate should equal the observed one (since by assumption now the identified middleman must be the monopoly one). Any deviation therefore is penalized by adding the squared difference to the objective function of (41) which therefore becomes:

$$
\begin{aligned}
\text{Criterion Function} \quad = \quad & (\Delta \text{ adverse selection cost} + 0.23)^2 + (\Delta \text{ trade frequency} - 0.17)^2 + \\
& + (\Delta \text{ middleman participation} - 0.143)^2 \qquad\qquad (50)
\end{aligned}
$$

and is again subject to the hard information constraint: $\alpha \geq 0.39$.

The calibration minimizes the criterion function over the parameter space $(\lambda, \alpha, \theta, \mu, p_0)$ subject to

hard information constraint. The parameter vector is the same as in the competitive case except for p_0 which was added (explanation below). We search for a (global) minimum over a grid of the parameter space. This grid is constructed by, for each parameter in the parameter vector, picking seven points that we consider reasonable. The grid search is numerically costly but avoids converging to a local minimum (potentially the case for a steepest descent type algorithm).For each value of the paramater vector we compute the bid price function that is determined by the differential equation of (48) along with a parameter p_0 that sets the bid price at $z^h = 0$ (i.e., $p(0) = p_0$).[11]

The calibration yields the following parameter vector (competitive case is added for comparison)

Parameter	Pre-set σ^2	Calibrated λ	α	θ	μ	p_0
Competitive **M**	3.14	1	0.39	0.58	3.35	
Monopoly **M**	3.14	0.67	0.39	2.52	4.00	1.33

and the following fit

	Targeted	Achieved Competitive **M**	Monopoly **M**
Δ adverse selection cost ($E(z \mid \text{trade})$)	-23%	-19%	-12%
Δ trade frequency	$+17\%$	$+22\%$	$+17\%$
Middlemen participation	$\geq 14.3\%$	28%	31%
Hard-information fraction (α)	≥ 0.39	0.39	0.39

Overall, the results shows that the monopoly middleman fit is worse than the competitive middleman fit. The monopoly middleman model seems to lack the power to generate the large effects

---

[11]We need the additional parameter p_0 as z^h is Gaussian and therefore has infinite support: $(-\infty, +\infty)$. We cannot apply the usual trick of evaluating the function value at either the lower or upper boundary. Instead, we add p_0 as a parameter to optimize over.

of middleman entry witnessed in the data. In particular, it falls far short (relative to the competitive middleman case) in generating the observed large adverse-selection cost reduction and trade frequency increase after middleman entry.

Comparing the parameter values across the two calibrations generates some insight into to why the monopoly case does worse. The most striking difference is that the monopoly calibration makes the seller experience a lot more information asymmetry, four times as much (i.e., $\theta = \sigma_z^2/\sigma_x^2$ is 2.52 instead of 0.58). This is arguably due to the monopoly middleman exploiting his position to take a cut of the total gains from trade, which makes the seller less eager to use him (by accepting his bid) to get the asset to the buyer. Jacking up information asymmetry then creates stronger incentives for the seller to use the middleman (who effectively reduces adverse selection cost between the seller and the buyer just as in the competitive case).

Another large difference is that in the monopoly case the buyer is less than fully aware of the hard information arrival, only in two thirds of the cases does the buyer observe the hard information the middleman always sees. In the competitive calibration the buyer always observed this information (i.e., the monopoly λ is 0.67 versus a competitive λ of 1). We speculate that this further incentivizes the seller to use the middleman post-entry. In other words, pre-entry the middleman is not around and there is therefore less information asymmetry between the seller and the buyer (as the buyer only observes the hard information in two thirds of the cases). Figure 10 illustrates the monopoly calibration and shows that post-entry the seller indeed posts a higher ask price (top-right graph).

We next compare welfare for the monopoly model before and after middleman entry. That is, we compute welfare for the model without a middleman and with a middleman at the monopoly model calibrated parameters. We find that in this case middleman entry leads to a modest welfare reduction of 2%. This contrasts with the 0.9% welfare increase suggested by the competitive case (see Section 3.3). Comparing these numbers one has to bear in mind that the competitive model fits the data substantially better.

Figure 10: Calibrated model monopoly middleman. These graphs illustrate the *monopoly* middlemen equilibrium for the calibrated parameters; the graphs are the same as in Figure 3, 4, and Figure 7 except that the this time they are derived from calibrating an alternative model (i.e., monopoly **M**) to the data. The solid blue line illustrates the benchmark case of no middlemen (**M**) and the dashed red line illustrates the case where **M** are present. The parameter values for these graphs are: λ=0.67, μ=4.00, σ^2=π, θ=2.52, and α=0.39. First-best welfare is 1 by construction; welfare in the no-**M** case is 35% of what first-best would achieve; welfare in the **M** case is 33% (i.e., a decrease of 2% relative to the **M** case). **M** participation rate is 31%. The top-left graph illustrates for what private value x the seller (**S**) decides to post a price himself rather than wait for a bid by **M**, i.e., the region in which $V(x)$ exceeds $U(x)$. The top-right graph illustrates the ask price **S** posts in case he decides to post a price. The bottom-left graph plots the probability of a transfer of the security from **S** to **B**. The bottom-right contour plot graphs the probability of an **S**-**B** transfer as a function of x and y which are the private values of the **S** and **B**, respectively.
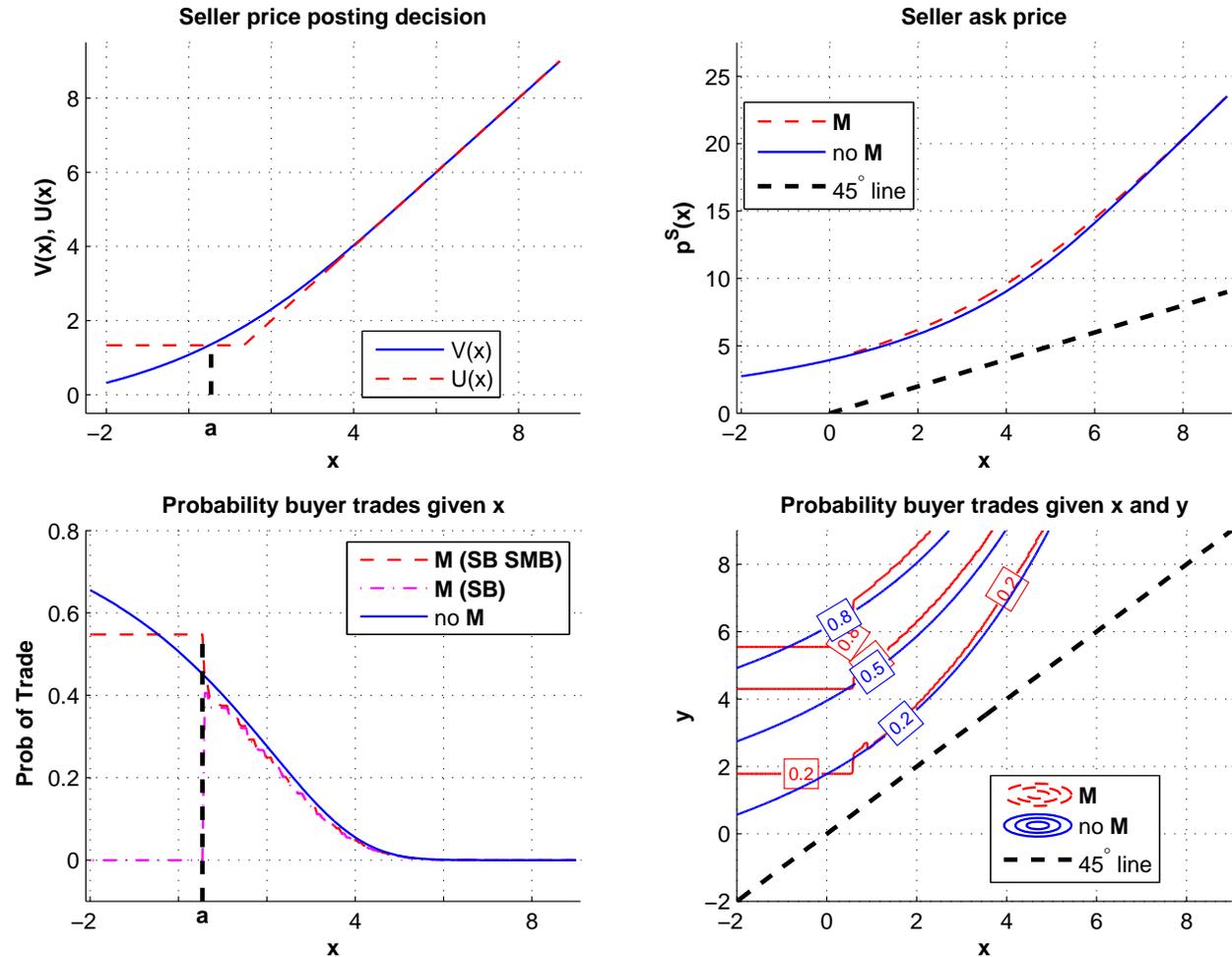
6   Conclusion

We have proposed a model where high-frequency traders (HFTs), equipped with hard information about common values, interact with investors in a limit order market. The model shows that if

43

HFTs are the only ones with such information, they can reduce welfare. If, on the other hand, late-arriving investors see the hard information with some probability then HFT entry can raise welfare. When calibrated to an HFT-entry event, the model implies that entry realized one percentage point more of the maximum gains from trade.

We further show that the introduction of a double auction could raise welfare more than HFT entry does. The design of the auction is critical to realizing further welfare improvement. Double auctions only beat HFT entry if the transaction price is set sufficiently close to the limit order price of the more informed agent, the buyer in this case. This prompts the follow-up question: How might exchange officials determine each trader's (average) order informativeness? Should it be the profitability of recent trades done by the trader? If so, how would it affect traders' incentives to search for information? These questions are left for future research.

Appendix

A   Notation (general notation and model parameters)

notation	meaning
M	middleman/machine, a representation of high-frequency traders
S	seller, an investor who owns the security and arrives early
B	buyer, an investor like **S** except that he does not own the security and arrives late (no overlap with **S**)
a, b	ask and bid price, respectively, in double auction (section 4)
F, G	CDF of private value (x or y) and common value (z), respectively
x, y	private value of security ownership for **S** and **B**, respectively
z_h, z_s	z_h is hard information, i.e., "quantitative, easy to store and transmit in impersonal ways;" z_s is soft information; z_h is known by **M** and potentially by **B**; z_s is known by **B** only
η	buyer valuation of the asset: $z+y$

parameter	description	modeling device for...
α	relative size of hard information in z	relative size on how informed **M** is on z
η	value of security to **S**: $\eta = y + z$	
k	determines transaction price in double auction, i.e., $p=(1-k)a+kb$ where b is the bid and a is the ask and $b > a$	
λ	probability that **B** observes hard information z_h	extent to which there is an adverse selection friction on hard information to begin with
θ	relative size of common value variance (σ_z^2) relative to private value variance (σ_x^2)	relative size of the informational friction
$\sigma_x^2=\sigma_y^2=\sigma^2=\pi$ $\mu_x=\mu_y=\mu$	mean and variance of independent normally distributed private values of ownership for **S** and **B**, respectively	**M**'s private value of ownership is set to zero; variance of private value is set to π so that expected social value of optimal reallocation is standardized to one ($E\max(x,y)-\mu = \frac{\sigma}{\sqrt{\pi}}$); μ captures expected social loss of **M** ownership relative to investor ownership
σ_z^2	variance of normally distributed common value innovation z	size of adverse selection friction in the absence of **M**

B Proofs

Proof of Proposition 1. Before proving the proposition, it is useful to first establish the following result.

LEMMA 2. ———————————————————————— *(Probability of buyer accepting)*
Under the normality assumptions in (6), the probability that the buyer accepts the offer of seller x can be expressed as

$$\tau(x) = \lambda \frac{\alpha(\sigma_h + \sigma_s)}{\sqrt{2\pi}(\sigma_h + \sigma_s)} \left(\frac{1}{1+(\sigma_h+\sigma_s)^2} p + \alpha(\sigma_h + \sigma_s)^2 - x \right) \exp\left[\frac{-(1-p)^2}{2\left(1+[\sigma_h+\sigma_s]^2\right)} \right]$$
$$+ (1-\lambda) \frac{\alpha(\sigma_s)}{\sqrt{2\pi}\sigma_s} \left(\frac{1}{1+\sigma_s^2} p + \alpha(\sigma_s)^2 - x \right) \exp\left[\frac{-(1-p)^2}{2\left(1+\sigma_s^2\right)} \right], \tag{51}$$

where $p = p(x)$ and

$$\alpha(\sigma) = \sqrt{\frac{\sigma^2}{1+\sigma^2}}.$$

Proof of lemma 2. For any arbitrary $\sigma \in \{\sigma_s, \sigma_h + \sigma_s\}$, in the expression

$$\phi\left(\frac{p-y}{\sigma}\right)\phi(y-1) = \frac{1}{2\pi} \exp\left(-\frac{1}{2\sigma^2}\left[(p-y)^2 + \sigma^2(y-1)^2\right] \right)$$

we need to complete the square:

$$(p-y)^2 + \sigma^2(y-1)^2 = p^2 + y^2 - 2py + \sigma^2\left(y^2 + 1 - 2y\right)$$
$$= \left(1+\sigma^2\right)y^2 - 2\left(p+\sigma^2\right)y + p^2 + \sigma^2$$
$$= \left(1+\sigma^2\right)\left[y^2 - 2\frac{p+\sigma^2}{1+\sigma^2}y + \frac{p^2+\sigma^2}{1+\sigma^2}\right].$$

Now for any constant A, we have $y^2 - 2Ay = (y-A)^2 - A^2$, and therefore:

$$(p-y)^2 + \sigma^2(y-1)^2 = \left(1+\sigma^2\right)\left[\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 - \left(\frac{p+\sigma^2}{1+\sigma^2}\right)^2 + \frac{p^2+\sigma^2}{1+\sigma^2}\right]$$
$$= \left(1+\sigma^2\right)\left[\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 + \left(\frac{p+\sigma^2}{1+\sigma^2}\right)\left(\frac{p^2+\sigma^2}{p+\sigma^2} - \frac{p+\sigma^2}{1+\sigma^2}\right)\right]$$
$$= \left(1+\sigma^2\right)\left[\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 + \frac{(1-p)^2\sigma^2}{(1+\sigma^2)^2}\right].$$

46

Therefore:

$$-\frac{1}{2\sigma^2}\left[(p-y)^2 + \sigma^2(y-1)^2\right] = -\frac{1+\sigma^2}{2\sigma^2}\left[\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 + \frac{(1-p)^2\sigma^2}{(1+\sigma^2)^2}\right]$$

$$= -\frac{1+\sigma^2}{2\sigma^2}\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 - \frac{1+\sigma^2}{2\sigma^2}\frac{(1-p)^2\sigma^2}{(1+\sigma^2)^2}$$

$$= -\frac{1+\sigma^2}{2\sigma^2}\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 + C,$$

where

$$C(\sigma) = \frac{-(1-p)^2}{2(1+\sigma^2)}. \tag{52}$$

Therefore:

$$\phi\left(\frac{p-y}{\sigma}\right)\phi(y-1) = \frac{1}{2\pi}\exp\left(-\frac{1+\sigma^2}{2\sigma^2}\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2 + C\right)$$

$$= \frac{1}{2\pi}e^C\exp\left(-\frac{1}{2\frac{\sigma^2}{1+\sigma^2}}\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2\right)$$

$$= \frac{\sqrt{\frac{\sigma^2}{1+\sigma^2}}}{\sqrt{2\pi}}e^C\frac{1}{\sqrt{2\pi\frac{\sigma^2}{1+\sigma^2}}}\exp\left(-\frac{1}{2\frac{\sigma^2}{1+\sigma^2}}\left(y - \frac{p+\sigma^2}{1+\sigma^2}\right)^2\right)$$

$$= \frac{\alpha}{\sqrt{2\pi}}e^C\frac{1}{\alpha\sqrt{2\pi}}\exp\left(-\frac{1}{2\alpha^2}\left(y - \alpha^2 - \frac{p}{1+\sigma^2}\right)^2\right).$$

Therefore:

$$\phi\left(\frac{p-y}{\sigma}\right)\phi(y-1) = \frac{\alpha}{\sqrt{2\pi}}e^C\frac{1}{\alpha}\phi\left(\frac{y - \alpha^2 - \frac{p}{1+\sigma^2}}{\alpha}\right).$$

Let

$$B(\sigma) = \frac{\alpha}{\sqrt{2\pi}}\frac{e^C}{\sigma}. \tag{53}$$

Then $B\sigma$ multiplies a normal density of y that has mean $\alpha^2 + \frac{p}{1+\sigma^2}$ and variance α^2.

Now, in **S**'s problem in (7) the FOC w.r.t. p which $p(x)$ solves then is

$$0 = \int_{-\infty}^{\infty}\left\{\begin{array}{l}\left[1 - \lambda\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) - (1-\lambda)\Phi\left(\frac{p-y}{\sigma_s}\right)\right] \\ + [x-y]\left(\frac{\lambda}{\sigma_h+\sigma_s}\phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) + \frac{1-\lambda}{\sigma_s}\phi\left(\frac{p-y}{\sigma_s}\right)\right)\end{array}\right\}\phi(y-1)\,dy. \tag{54}$$

Therefore in (54):

$$\int_{-\infty}^{\infty}\frac{x-y}{\sigma}\phi\left(\frac{p-y}{\sigma}\right)\phi(y-1)\,dy = B\left(x - \alpha^2 - \frac{1}{1+\sigma^2}p\right).$$

47

Now (54) becomes

$$
\begin{aligned}
0 &= \int_{-\infty}^{\infty} \left\{ \begin{array}{l} \left[1 - \lambda\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) - (1-\lambda)\Phi\left(\frac{p-y}{\sigma_s}\right)\right] \\ + [x-y]\left(\frac{\lambda}{\sigma_h+\sigma_s}\phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) + \frac{1-\lambda}{\sigma_s}\phi\left(\frac{p-y}{\sigma_s}\right)\right) \end{array} \right\} \phi(y-1)\,dy \\
&= \int_{-\infty}^{\infty} \left[1 - \lambda\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) - (1-\lambda)\Phi\left(\frac{p-y}{\sigma_s}\right)\right]\phi(y-1)\,dy \\
&\quad + \lambda B(\sigma_h+\sigma_s)\left(x - \alpha(\sigma_h+\sigma_s)^2 - \frac{1}{1+(\sigma_h+\sigma_s)^2}p\right) \\
&\quad + (1-\lambda)B(\sigma_s)\left(x - \alpha(\sigma_s)^2 - \frac{1}{1+\sigma_s^2}p\right).
\end{aligned} \tag{55}
$$

Rearranging and using (5), we get the probability of trade being

$$
\begin{aligned}
\tau(x) &= \lambda B(\sigma_h+\sigma_s)\left(\frac{1}{1+(\sigma_h+\sigma_s)^2}p + \alpha(\sigma_h+\sigma_s)^2 - x\right) \\
&\quad + (1-\lambda)B(\sigma_s)\left(\frac{1}{1+\sigma_s^2}p + \alpha(\sigma_s)^2 - x\right).
\end{aligned} \tag{56}
$$

Using (53), we get (51).

Proof of part (i) of Proposition 1 with result from Lemma 2. There are several cases to consider.

Case 1: $\sigma_s \to \infty$.

Case 1.1: If for given x, $p/(\sigma_h+\sigma_s) \to \infty$, then

$$
\tau(x) = \int_{-\infty}^{\infty} \left[1 - \lambda\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) - (1-\lambda)\Phi\left(\frac{p-y}{\sigma_s}\right)\right]\phi(y-1)\,dy \to 0
$$

as $\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) \to 1$ and $\Phi\left(\frac{p-y}{\sigma_s}\right) \to 1$.

Case 1.2: If p/σ is bounded for either $\sigma \in \{\sigma_s, \sigma_h+\sigma_s\}$ then so is $C(\sigma) = \frac{-(1-p)^2}{2(1+\sigma^2)}$, and thus $B(\sigma) = \frac{\alpha(\sigma)}{\sqrt{2\pi}}\frac{e^{C(\sigma)}}{\sigma} \to 0$, and so is τ given (56).

Case 2: $\sigma_h \to \infty$.

Case 2.1: If for a given x, $p/(\sigma_h+\sigma_s) \to \infty$, then $p/\sigma_s \to \infty$. Therefore:
$\tau(x) = \int_{-\infty}^{\infty} \left[1 - \lambda\Phi\left(\frac{p-y}{\sigma_h+\sigma_s}\right) - (1-\lambda)\Phi\left(\frac{p-y}{\sigma_s}\right)\right]\phi(y-1)\,dy \to 0$.

Case 2.2: If $p/(\sigma_h + \sigma_s)$ is bounded, then so is $C(\sigma) = \frac{-(1-p)^2}{2(1+\sigma^2)}$ $\forall \sigma \in \{\sigma_s, \sigma_h + \sigma_s\}$. Thus $B(\sigma_h + \sigma_s) = \frac{\alpha(\sigma_h+\sigma_s)}{\sqrt{2\pi}} \frac{e^{C(\sigma_h+\sigma_s)}}{\sigma_h+\sigma_s} \to 0$. Then using (56):

$$\tau(x) \to (1 - \lambda) B(\sigma_s)\left(\frac{1}{1+\sigma_s^2}p + \alpha(\sigma_s)^2 - x\right).$$

If the RHS is not zero, this implies that

$$\lim_{\sigma_h \to \infty} p(x) < \infty.$$

Call the limit $\bar{p}$. Evaluate **S**'s utility and show that it converges to $-\infty$. In the second line of (7), it is enough to look at the term $\int_{-\infty}^{\bar{p}-y}(x+z)\frac{\lambda}{\sigma_h+\sigma_s}\phi\left(\frac{z}{\sigma_h+\sigma_s}\right)dz$ for each y. For given y,

$$\int_{-\infty}^{\bar{p}-y}(x+z)\frac{\lambda}{\sigma_h+\sigma_s}\phi\left(\frac{z}{\sigma_h+\sigma_s}\right)dz = \Phi\left(\frac{\bar{p}-y}{\sigma_h+\sigma_s}\right)\left[x\lambda + \lambda\lim_{\sigma_h\to\infty}E(z\mid z<\bar{p}-y)\right],$$

which, as long as $\lambda > 0$, converges to $-\infty$ if $\lim_{\sigma_h\to\infty}E(z\mid z<\bar{p}-y) = -\infty$. But the latter is the mean of a truncated normal distribution. Here the inverse Mill's ratio formula (with a mean of zero for z) applies so that

$$E(z\mid z<\bar{p}-y) = -\frac{\phi\left(\frac{\bar{p}-y}{\sigma_h+\sigma_s}\right)}{\Phi\left(\frac{\bar{p}-y}{\sigma_h+\sigma_s}\right)}(\sigma_h+\sigma_s).$$

But $\bar{p}$ is finite, and therefore:

$$E(z\mid z<\bar{p}-y) \to -\frac{\phi(0)}{\Phi(0)}\left(\sigma_s + \lim_{\sigma_h\to\infty}\sigma_h\right) = -\infty.$$

Proof of part (ii) of Proposition 1. The proof of these two claims is identical to the proof that $\partial V/\partial\sigma_h < 0$ that leads up to (58).

Proof of Proposition 2. (*i*) follows from (26). (*ii*) v^M does not depend on σ_h^2 and therefore by (17), neither does $U(x)$. Next, we show that $V(x)$ is strictly decreasing in σ_h^2. At any p, (20) can be written as:

$$V(x) = \max_p \int \left[p\left[1 - \Phi\left(\frac{p-\xi}{\sigma_h}\right)\right] + \int_{-\infty}^{(p-\xi)/\sigma_h}(x+z^s+\sigma_h\varepsilon)\phi(\varepsilon)\,d\varepsilon\right]dG^s dF, \qquad (57)$$

where $\xi = \max(v^M, y + z^s)$. Differentiating w.r.t. σ_h, using the envelope theorem, and canceling p yields:

$$\frac{\partial V}{\partial\sigma_h} = \int\left[-\frac{p-\xi}{\sigma_h^2}(x+z^s-\xi)\phi\left(\frac{p-\xi}{\sigma_h}\right) + \int_{-\infty}^{(p-\xi)/\sigma_h}\varepsilon\phi(\varepsilon)\,d\varepsilon\right]dG^s dF,$$

where ϕ and Φ are the standard normal density and CDF. The FOC w.r.t. p now says:

$$0 = \int \left(1 - \Phi\left(\frac{p - \xi}{\sigma_h}\right) + (x + z^s - \xi)\frac{1}{\sigma_h}\phi\left(\frac{p - \xi}{\sigma_h}\right)\right)dG^s dF.$$

Therefore:

$$\frac{\partial V}{\partial \sigma_h} = \int \left(\frac{p - \xi}{\sigma_h}\left(1 - \Phi\left(\frac{p - \xi}{\sigma_h}\right)\right) + \int_{-\infty}^{(p-\xi)/\sigma_h} \varepsilon\phi(\varepsilon)\,d\varepsilon\right)dG^s dF$$

$$= \int \left(\int_{-\infty}^{\infty} \min\left(\varepsilon, \frac{p - \xi}{\sigma_h}\right)\phi(\varepsilon)\,d\varepsilon\right)dG^s dF < 0, \tag{58}$$

because $\int_{-\infty}^{\infty} \varepsilon\phi(\varepsilon)\,d\varepsilon = 0$ and because $\phi(\varepsilon) > 0$ for all $\varepsilon \in R$.

Proof of Lemma 1. Suppose that for some x (27) fails. Then (57) implies that $V(x) \to -\infty$. But **S** can guarantee himself x, hence $V(x) \geq x$, a contradiction that proves (27). Now, a well-known inequality on the Mill's ratio (see Sampford (1953, eqn. (2))) states that for any $s > 0$, $s\int_x^{\infty} e^{-\frac{1}{2}u^2}du < e^{-\frac{1}{2}s^2}$ and therefore for any ξ, $p\left(1 - \Phi\left(\frac{p-\xi}{\sigma_h}\right)\right) \to 0$. Moreover, $\int_{-\infty}^{(p-\xi)/\sigma_h} \varepsilon\phi(\varepsilon)\,d\varepsilon \leq 0$, which proves (28) for all x.

Proof of Proposition 3. By (25), the limit is v^M, and by Proposition 3 it is reached monotonically. This proves the first claim. For the second, it suffices to note that **S**'s probability of a sale and, hence, **M**'s participation is bounded below by $F\left(v^M\right)$ whereas **B** is in the limit shut out of the market due to (27).

C Numerical procedure to determine optimal double auction strategies

The algorithm estimated (45) and (46) using numerical integral approximation. Using a grid of $100 + 50 \times 50$ for (x, y, z), the steps are as follows:

1. For each grid point of (y, z), guess $b(y + z)$ and label this guess $b^i(y + z)$ where i is set to zero;

2. For the current value of i, calculate $a^i(x)$ for each x from (45) based on $b^i(y + z)$, using numerical integration;

3. Given $a^i(x)$, solve for $b^{i+1}(y + z)$ from (46), using numerical integration;

4. The algorithm terminates if either

(a)

$$\sup \left\| b^{i+1}(y+z) - b^i(y+z) \right\| < 10^{-6}$$

or

(b)

$$\inf \left(\sqrt{\frac{1}{N_z N_y} \left(b^{i+1}(y+z) - b^i(y+z) \right)' \left(b^{i+1}(y+z) - b^i(y+z) \right)} \right) < 10^{-2},$$

where N_z and N_y are number of grid points for z and y;

5. Increment i by one and repeat the procedure from step 2.

References

Aït-Sahalia, Yacine and Mehmet Saglam. 2014. "High Frequency Traders: Taking Advantage of Speed." Manuscript, Princeton University.

Amihud, Yakov and Haim Mendelson. 1980. "Dealership Market: Market-Making with Inventory." *Journal of Financial Economics* 8:31–53.

Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko. 2012. "The Trading Profits of High Frequency Traders." Manuscript, University of Washington.

Biais, B., P. Hillion, and C. Spatt. 1995. "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse." *Journal of Finance* 50:1655–1689.

Biais, Bruno, Thierry Foucault, and Sophie Moinas. 2015. "Equilibrium Fast Trading." *Journal of Financial Economics* 73:3–36.

Black, F. 1995. "Equilibrium Exchanges." *Financial Analysts Journal* 51:23–29.

Boehmer, Ekkehart, Kingsley Fong, and Juan Wu. 2014. "International Evidence on Algorithmic Trading." Manuscript, EDHEC Business School.

Boehmer, Ekkehart, Dan Li, and Gideon Saar. 2015. "Correlated High-Frequency Trading." Manuscript, Cornell University.

Breckenfelder, Johannes. 2013. "Competition between High-Frequency Traders, and Market Quality." Manuscript, European Central Bank.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. "High Frequency Trading and Price Discovery." *Review of Financial Studies* 27:2267–2306.

Budish, Eric, Peter Cramton, and John Shim. 2015. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." *Quarterly Journal of Economics* 130:1547–1621.

Cartea, Álvaro and Jose Penalva. 2012. "Where is the Value in High Frequency Trading?" *Quarterly Journal of Finance* 2:1–46.

Chaboud, Alain, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega. 2014. "Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market." *Journal of Finance* 69:2045–2084.

Chatterjee, Kalyan and William Samuelson. 1983. "Bargaining under Incomplete Information." *Operations Research* 31:835–851.

Clark, Charles E. 1961. "The Greatest of a Finite Set of Random Variables." *Operations Research* 9:145–162.

Duffie, Darrell, Semyon Malamud, and Gustavo Manso. 2009. "Information Percolation with Equilibrium Search Dynamics." *Econometrica* 77:1513–1574.

Easley, David, Marcos M. López de Prado, and Maureen O'Hara. 2013. *High-Frequency Trading*. London: Risk Books.

Fama, Eugene. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25:383–417.

Farmer, Doyne J. and Spyros Skouras. 2012. "Review of the Benefits of a Continuous Market vs. Randomized Stop Auctions and of Alternative Priority Rules (Policy Options 7 and 12)." Manuscript, Foresight, Government Office for Science, UK.

Foucault, Thierry, Johan Hombert, and Ioanid Roşu. 2016. "News Trading and Speed." *Journal of Finance* 71:335–382.

Foucault, Thierry and Albert J. Menkveld. 2008. "Competition for Order Flow and Smart Order Routing Systems." *Journal of Finance* 63:119–158.

Glode, Vincent and Christian Opp. 2016. "Asymmetric Information and Intermediation Chains." *American Economic Review (forthcoming)* .

Glosten, Lawrence and Paul Milgrom. 1985. "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents." *Journal of Financial Economics* 14:71–100.

Glosten, L.R. 1987. "Components of the Bid Ask Spread and the Statistical Properties of Transaction Prices." *Journal of Finance* 42:1293–1307.

Golosov, Michael, Guido Lorenzoni, and Aleh Tsyvinki. 2014. "Decentralized Trading with Private Information." *Econometrica* 82:1055–1091.

Grossman, Sanford J. and Merton H. Miller. 1988. "Liquidity and Market Structure." *Journal of Finance* 43:617–633.

Grossman, S.J. and J.E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70(3):393–408.

Hagströmer, Björn and Lars Nordén. 2013. "The Diversity of High Frequency Traders." *Journal of Financial Markets* 16:741–770.

Hasbrouck, Joel. 1995. "One Security, Many Markets: Determining the Contributions to Price Discovery." *Journal of Finance* 50:1175–1199.

———. 2007. *Empirical Market Microstructure*. New York: Oxford University Press.

Hasbrouck, Joel and Gideon Saar. 2013. "Low-Latency Trading." *Journal of Financial Markets* 16:646–679.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld. 2011. "Does Algorithmic Trading Improve Liquidity?" *Journal of Finance* 66:1–33.

Hendershott, Terrence and Albert J. Menkveld. 2014. "Price Pressures." *Journal of Financial Economics* (forthcoming).

Ho, Thomas and Hans R. Stoll. 1981. "Optimal Dealer Pricing under Transaction Cost and Return Uncertainty." *Journal of Financial Economics* 9:47–73.

Hoffmann, Peter. 2014. "A Dynamic Limit Order Market with Fast and Slow Traders." *Journal of Financial Economics* 113:156–169.

Hollifield, Burton, Robert A. Miller, Patrik Sandås, and Joshua Slive. 2006. "Estimating the Gains from Trade in Limit Order Markets." *Journal of Finance* 61:2753–2804.

Jones, Charles M. 2013. "What Do We Know About High-Frequency Trading?" Manuscript, Columbia University.

Kirilenko, Andrei, Albert Kyle, Mehrdad Samadi, and Tugkan Tuzun. 2014. "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market." Manuscript, University of Maryland.

Kyle, A. 1985. "Continuous Auctions and Insider Trading." *Econometrica* 53:1315–1335.

Li, Yiting. 1996. "Middlemen and Private Information." *Journal of Monetary Economics* 42:131–159.

Masters, Adrian. 2007. "Middlemen in Search Equilibrium." *International Economic Review* 48:343–362.

Menkveld, Albert J. 2013. "High Frequency Trading and the New Market Makers." *Journal of Financial Markets* 16:712–740.

———. 2016. "The Economics of High-Frequency Trading: Taking Stock." *Annual Review of Financial Economics (forthcoming)* .

Menkveld, Albert J. and Marius A. Zoican. 2016. "Need for Speed? Low Latency Trading and Adverse Selection." Manuscript, Vrije Universiteit Amsterdam.

Myerson, Roger and Mark A. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 29:265–281.

Parlour, Christine A. and Duane J. Seppi. 2008. "Limit Order Markets: A Survey." In *Handbook of Financial Intermediation and Banking*, edited by A.W.A. Boot and A.V. Thakor. Amsterdam, Netherlands: Elsevier Publishing.

Petersen, Mitchell A. 2004. "Information: Hard and Soft." Manuscript, Northwestern University.

Roşu, Ioanid. 2016. "Fast and Slow Informed Trading." Manuscript, HEC Paris.

Roll, Richard. 1988. "R^2." *Journal of Finance* 43:541–566.

Rubinstein, Ariel and Asher Wolinsky. 1987. "Middlemen." *Quarterly Journal of Economics* 102:581–593.

Sampford, M.R. 1953. "Some Inequalities on Mill's Ratio and Related Functions." *Annals of Mathematical Statistics* 24:130–132.

SEC. 2010. "Concept Release on Equity Market Structure." Release No. 34–61358. File No. S7–02-10.

———. 2014. "Equity Market Structure Literature Review. Part II: High Frequency Trading." Literature review.

Wah, Elaine and Michael Wellman. 2013. "Latency Arbitrage, Market Fragmentation, and Efficiency: A Two-Market Model." Manuscript, 14th ACM Conference on Electronic Commerce, June.