ID: 56

Submitted: 2013-03-23

Last Updated: 2013-03-23

Title: Exploratory Trading

Author 1:
First Name: Adam
Last Name: Clark-Joseph
Organization: Harvard University
Country: United States
Email: adjoseph@fas.harvard.edu

Exploratory Trading

Author 90901221*

November 14, 2012

Abstract

Using comprehensive, account-labeled message records from the E-mini S&P 500 futures market, I investigate the mechanisms underlying high-frequency traders' capacity to profitably anticipate price movements. Of the 30 high-frequency traders (HFTs) that I identify in my sample, eight earn positive overall profits on their aggressive orders. I find that all eight of these HFTs consistently lose money on their smallest aggressive orders, and these losses are not explained by inventory management. These losses on small orders, as well as the more-than-offsetting gains on larger orders, could be rationalized if the small orders provided some informational value, and I model how a trader could gather valuable private information by using her own orders in an exploratory manner to learn about market conditions. This "exploratory trading" model predicts that the market response to the trader's "exploratory" order would help to explain her earnings on her next order, but would not explain any other traders' subsequent performance. In direct confirmation of the model's predictions, I find that a simple measure of changes in the orderbook immediately following small aggressive orders placed by the eight HFTs explains a significant additional component of those HFTs' earnings on subsequent, larger orders, but this information offers little or no additional power to explain *other* traders' earnings on subsequent orders. These findings help to clarify nature of the information on which HFTs trade and offer a starting point to address the open questions about social welfare implications of high-frequency trading.

*The views expressed in this paper are my own and do not constitute an official position of the Commodity Futures Trading Commission, its Commissioners, or staff.

1 Introduction

Over the past three decades, information technology has reshaped major financial exchanges worldwide. Physical trading venues have increasingly given way to electronic ones, and trading responsibilities that once fell on human agents have increasingly been delegated to computer algorithms. Automation now pervades financial markets; for example, Hendershott and Riordan (2009) and Hendershott *et al.* (2011) respectively document the dramatic levels of algorithmic trading on the Deutsche Boerse and the New York Stock Exchange. Much of the algorithmic activity in major markets emanates from so-called "high-frequency traders" ("HFTs"). Although it dominates modern financial exchanges, HFTs' activity remains largely mysterious and opaque—it is the "dark matter" of the trading universe.

HFTs are distinguished not only by the large number of trades they generate (i.e., their literal high trading frequency), but also by the speed with which they can react to market events. HFTs achieve these remarkable reaction times, typically measured in milliseconds, by using co-location services, individual data feeds, and high-speed computer algorithms. Two further hallmarks of HFTs are their extremely short time-frames for maintaining positions, and their propensity for "ending the trading day in as close to a flat position as possible (that is, not carrying significant, unhedged positions over-night [when markets are closed])."[1]

Empirical study of high-frequency trading has proven challenging, but not impossible. For example, Brogaard *et al.* (2012) obtain and analyze a NASDAQ dataset that flags messages from an aggregated group of 26 HFT firms, and Hasbrouck and Saar (2011) conduct a complementary analysis by statistically reconstructing "strategic runs" of linked messages in NASDAQ order data. Both of these analyses suggest beneficial effects from HFTs' activity, but inherent limitations of the underlying data restrict these studies' scope to explain how and why such effects arise.

Understanding and explaining the impacts of high-frequency trading requires some understanding of what HFTs are actually doing, and of how their strategies work. Even in a market for a single asset, HFTs exhibit considerable heterogeneity, so aggregate HFT activity reveals little about what individual HFTs really do. Data suitable for the study of individual HFTs' activity are difficult to obtain. Whereas publicly available 13-F forms reveal the behavior of institutional investors at a quarterly frequency, there is no comparable public data that can be used to track and analyze the behavior of individual traders at a second- or millisecond-frequency.

The only fully adequate data currently available for academic research on high-frequency trading

[1] U.S. Securities and Exchange Commission Concept Release No. 34-61358, pages 45-46.

come from regulatory records that the Chicago Mercantile Exchange provides to the U.S. Commodity Futures Trading Commission. Kirilenko *et al.* (2010) pioneered the use of transaction data from these records to investigate high-frequency trading in their analysis of the so-called "Flash Crash" of May 6, 2010 in the market for E-mini Standard & Poors 500 stock index futures contracts (henceforth, "E-mini"). This work introduced a scheme to classify trading accounts using simple measures of overall trading activity, intraday variation in net inventory position, and inter-day changes in net inventory position. Of the accounts with sufficiently small intra- and inter-day variation in net position, Kirilenko *et al.* classify those with the highest levels of trading activity as HFTs, and these accounts are archetypes of high-frequency traders.

Kirilenko *et al.* find that HFTs participate in over one-third of the trading volume in the E-mini market, and subsequent research by Baron *et al.* (2012) documents the large and stable profits that HFTs in the E-mini market earn. This work provides empirical confirmation of HFTs' importance, and it offers some crisp descriptions of HFTs' activity. However, it does not attempt to explain why HFTs act as they do, or how HFTs earn profits. Indeed, no extant empirical research attempts such explanations. In this paper, I address a central aspect of this open problem. HFTs in the E-mini market earn roughly 40% of their profits from the transactions that they initiate—that is, from their so-called "aggressive" orders—and I examine the mechanism underlying HFTs' capacity to earn these profits.

How do HFTs in the E-mini market make money from their aggressive orders? One possibility is that HFTs merely react to public information faster than everyone else; this premise underlies the models of Biais *et al.* (2010), Jarrow and Protter (2011), and Cespa and Foucault (2008). A second possibility is that HFTs simply front-run coming demand when they can predict future aggressive orders. However, I find neither of these hypotheses to be consistent with the data.

I identify the HFTs who profit from their aggressive orders, then I investigate how these HFTs manage to do so. I show that the HFTs who profit from their aggressive trading use small aggressive orders to obtain private information that helps to forecast the price-impact of predictable demand innovations. Demand innovations in the E-mini market are easy to predict, but the price-elasticity of supply is not, and price-impact is usually too small for indiscriminate front-running of predictable demand to be profitable.[2] However, the private information about price-impact generated by an

[2] To be more precise, it is extremely easy to predict whether future aggressive orders will be buy orders or sell orders. The dynamic behavior of passive orders resting in the orderbook—analogous to supply elasticity—is considerably harder to forecast.

HFT's small aggressive orders enables that HFT to trade ahead of predictable demand at only those times when it is profitable to do so (i.e., when price-impact is large). To elucidate how this works, I develop a theoretical model of what I term "exploratory trading."

Fundamentally, the model of exploratory trading rests on the notion that an HFT's aggressive orders generate valuable private information, specifically, information about the price-impact of the aggressive orders that follow. When an HFT places an exploratory order and observes a large price-impact, he learns that supply is temporarily inelastic. If the HFT knows that there is going to be more demand soon thereafter, he can place a larger order (even with a big price-impact) knowing that the price-impact from the coming demand will drive prices up further and ultimately enable him to sell at a premium that exceeds the price-impact of his unwinding order. When an HFT knows that supply is temporarily inelastic, he follows a routine demand-anticipation strategy. The purpose of exploratory trading is not to learn about future demand, but rather to identify the times at which trading in front of future demand will be profitable. Active learning in financial markets is a relatively old idea, dating back at least to Leach and Madhavan's papers in 1992 and 1993, but exploratory trading is an active-learning mechanism that is new to the academic literature. In section 2, I present a model to formalize the concept of exploratory trading, and I derive the model's central testable predictions.

Using novel electronic message data at the Commodity Futures Trading Commission, I examine the profitability of individual HFTs' aggressive orders. I find that eight of the 30 HFTs in my sample profit from their aggressive trading overall and significantly outperform non-HFTs. However, these same eight HFTs all lose money on their smallest aggressive orders. (For brevity, I refer to these eight HFTs as "A-HFTs," and to the remaining 22 as "B-HFTs.") Exploratory trading would produce just such a pattern of incurring small losses on exploratory orders then realizing large gains; these descriptive results both motivate further tests and suggest the A-HFTs' small aggressive orders as natural candidates for potential exploratory orders.

To explicitly test the predictions of the exploratory trading model for the eight A-HFTs, I examine the extent to which information about the changes in the orderbook following small aggressive orders explains the profits that various traders earn on subsequent aggressive orders. The exploratory trading model predicts that information about the changes following an A-HFT's small aggressive order will explain a significant additional component of the A-HFT's subsequent performance, but that this information will not explain any additional component of other traders' subsequent performance. Consistent with these predictions, I find that the orderbook changes immediately following A-HFTs' small aggressive orders provide significant additional explanatory power for the respective A-HFTs' performance on their larger aggressive orders, but not for other traders' performance.

The remainder of this paper is organized as follows: Section 2 presents a simple model of exploratory trading, along with the model's central predictions, and establishes the empirical agenda. Section 3 describes the data and precisely defines HFTs. Section 4 addresses the overall profitability of HFTs' aggressive orders and precisely characterizes the A-HFTs, then examines the A-HFTs' losses on small aggressive orders. Section 5 presents direct empirical tests of the exploratory trading model's key predictions, section 6 examines the practical significance of exploratory information, and section 7 discusses extensions and implications of the empirical results. Section 8 concludes.

2 Exploratory Trading: Theory

The ultimate objective of this paper is to explain the mechanism underlying HFTs' capacity to profit from their aggressive orders in the E-mini market, and this section establishes the theoretical framework for my empirical investigation.

As noted in the introduction, demand innovations in the E-mini market are easy to predict from public market data, but the price-elasticity of supply is not. Although there are times when supply is unaccommodating and high future demand forecasts price changes that are large enough to profit from, such times are difficult or impossible to identify by merely observing public market data. In this type of setting, a trader can obtain additional information about supply conditions by placing an "exploratory" aggressive order and observing how prices and supply respond. The additional exploratory information enables the trader to determine whether supply is accommodating (and expected price-impact small) or unaccommodating (and expected price-impact large), and this helps the trader to decide whether he can profit by trading ahead of an imminent demand innovation.

The basic model I examine is a simple representation of a market in which demand is easy to predict, but supply elasticity is not. I consider a two-period model with two possible states for supply conditions (accommodating or unaccommodating), and three possible demand innovations in the second period (positive, negative, or zero). The demand innovation is automatically revealed before it arrives in the second period, but the state of supply conditions is only revealed if a trader places an aggressive order in the first period.

In this context, I consider the problem facing a single trader, the "HFT." In the first period, the

HFT has the opportunity to place an aggressive order and thereby learn about supply conditions. In the second period, regardless of what happened in the first period, the HFT observes a signal about future demand, after which he again has an opportunity to place an aggressive order. The signal of future demand forecasts price innovations much more accurately when combined with information about supply conditions than it does when used on its own. If the HFT places an aggressive order in the first period, he effectively "buys" supply information that he can use in the second period to better decide whether he should place another aggressive order. Consequently, the HFT may find it optimal in the first period to place an order that he expects to be unprofitable, since the information that the order generates will be valuable in the second period.

The rest of this section is devoted to formally developing a model of exploratory trading and deriving the model's testable predictions. In addition to the basic result about the value of exploratory information sketched above, I address the key issue of why an order generates more information for the trader who submitted it than it does for everyone else. Appendix A contains full mathematical details.

2.1 Baseline Model

In an order-driven market, every regular transaction is initiated by one of the two executing transactors. The transactor who initiates is referred to as the "aggressor," while the opposite transactor is referred to as the "passor." The passor's order was resting in the orderbook, and the aggressor entered a new order that executed against the passor's preexisting resting order. Assuming that prices are discrete, the lowest price of any resting sell order in the orderbook ("best ask") always exceeds the highest price of any resting buy order in the book ("best bid") by at least one increment (the minimal price increments are called "ticks"). A transaction initiated by the seller executes at the best bid, while a transaction initiated by the buyer executes at the best ask; the resulting variation in transaction prices between aggressive buys and aggressive sells is known as "bid-ask bounce." Hereafter, except where otherwise noted, I will restrict attention to price changes distinct from bid-ask bounce. Empirically, the best ask for the most actively traded E-mini contract almost always exceeds the best bid by exactly one tick during regular trading hours, so movements of the best bid, best ask, and mid-point prices are essentially interchangeable.

If the best bid and best ask were held fixed, a trader who aggressively entered then aggressively exited a position would lose the bid-ask spread on each contract, whereas a trader who passively entered then passively exited a position would earn the bid-ask spread on each contract. Intuitively, aggressors pay for the privilege of trading precisely when they wish to do so, and passors are compensated for the costs of supplying this "immediacy," cf. Grossman and Miller (1988). These costs include fixed operational costs and costs arising from adverse selection. Cf. Glosten and Milgrom (1985), Stoll (1989).

An aggressive order will execute against all passive orders at the best available price level before executing against any passive orders at the next price, so an aggressive order will only have a literal price-impact if it eats through all of the resting orders at the best price. In the E-mini market, it is rare for an aggressive order to have a literal price-impact, not only because there are enormous numbers of contracts at the best bid and best ask, but also because aggressive orders overwhelmingly take the form of limit orders priced at the opposite best (which cannot execute at the next price level).

2.1.1 Market Structure

Let time be discrete, consisting of two periods, $t = 1, 2$. This model should be interpreted as a single instance of the hundreds or thousands of similar scenarios that arise throughout the trading day.

Consider an order-driven market with discrete prices, and assume that both the orderbook and order-flow are observable. Conceptually, the flow of aggressive orders is analogous to demand, while the set of passive orders in the orderbook ("resting depth") is analogous to supply.

2.1.2 Passive Orders

There are two possible states for the behavior of passive orders: accommodating and unaccommodating. Let the variable Λ represent this state, which I call the "liquidity state." The liquidity state is the same in both periods of the model. Denote the accommodating liquidity state by $\Lambda = A$, and the unaccommodating state by $\Lambda = U$. Assume that $\Lambda = U$ with *ex-ante* probability u, and $\Lambda = A$ with complementary *ex-ante* probability $1 - u$.

The liquidity state characterizes the behavior of resting depth in the orderbook after an aggressive order executes—a generalization of price-impact appropriate for an order-driven market. When an aggressive buy (sell) order executes, it mechanically depletes resting depth on the sell (buy) side of the orderbook. Following this mechanical depletion, traders may enter, modify, and/or cancel passive orders, so resting depth at the best ask (bid) can either replenish, stay the same, or deplete further. The aggressive order's impact is offset to some extent—or even reversed—if resting depth replenishes, whereas the aggressive order's impact is amplified if resting depth depletes further. In the accommodating state ($\Lambda = A$) resting depth weakly replenishes, while in the unaccommodating state ($\Lambda = U$) resting depth further depletes. Intuitively, aggressive orders have a small price-impact in the accommodating state, and a large price-impact in the unaccommodating state.

Although the orderbook is always observable, static features of passive orders in the orderbook do not directly reveal the liquidity state. Because the liquidity state relates to the dynamic behavior of resting depth after an aggressive order executes, this state can only be observed through the changes in the orderbook that follow the execution of an aggressive order.

2.1.3 Aggressive Order-Flow

At the end of period 2, traders other than the HFT exogenously place aggressive orders. Let the variable $\varphi \in \{-1, 0, +1\}$ describe this exogenous aggressive order-flow. The variable φ is just a coarse summary of the order-flow—It does not represent the actual number of contracts. Intuitively, $\varphi = -1$ represents predictable selling pressure and $\varphi = +1$ represents predictable buying pressure, while $\varphi = 0$ represents an absence of predictable pressure in either direction.

Assume that $\varphi = +1$ and $\varphi = -1$ with equal probabilities $\mathbb{P}\{\varphi = +1\} = \mathbb{P}\{\varphi = -1\} = v/2$, and $\varphi = 0$ with complementary probability $1 - v$. The value of φ does not depend on the liquidity state, Λ, nor does it depend on the HFT's actions.

The price-change at the end of period 2, which I denote by y, is jointly determined by the exogenous aggressive order-flow and the liquidity state. In the notation of the model,

$$y = \begin{cases} \varphi & if\ \Lambda = U \\ 0 & if\ \Lambda = A \end{cases} \tag{1}$$

In other words, if the liquidity state is unaccommodating ($\Lambda = U$), aggressive order-flow affects the price, and $y = \varphi$. However, if the liquidity state is accommodating ($\Lambda = A$), aggressive order-flow does not affect the price, and $y = 0$ regardless of the value of φ.

2.1.4 The HFT

The HFT submits only aggressive orders, and these aggressive orders are limited in size to N contracts or fewer. Let $q_t \in \{-N, \ldots, -1, 0, 1, \ldots, N\}$ denote the signed quantity of the aggressive order that the HFT places in period t, where a negative quantity represents a sale, and a positive quantity represents a purchase.

Assume that the HFT pays constant trading costs of $\alpha \in (0.5, 1)$ per contract. The lower bound of 0.5 on α corresponds to half of the minimum possible bid-ask spread, while the upper bound of 1 merely excludes trivial cases of the model in which aggressive orders are always unprofitable. When $\alpha > u$, the HFT will never place an order in period 2 if he doesn't know the liquidity state, and I focus on this case to simplify the exposition; results are qualitatively unchanged for $u \geq \alpha$ (see Appendix A).

I assume that the HFT's aggressive orders have no literal price-impact. Intuitively, the HFT only trades contracts at the initial best bid/ask. For example, in period 2, if the HFT has learned that the liquidity state is unaccommodating and $\varphi = +1$, he will buy all of the contracts available at the best ask. This is one way to interpret the size limitation on the HFT's orders.

The HFT's profit from the aggressive order he places in period t is given by

$$\pi_t = yq_t - \alpha |q_t| \tag{2}$$

where y denotes the price-change at the end of period 2. Let

$$\pi_{total} := \pi_1 + \pi_2$$

denote the HFT's total combined profits from periods 1 and 2. Assume that the HFT is risk-neutral and seeks to maximize the expectation of his total profits, π_{total}.

2.1.5 Model Timeline

Period 1 In period 1, the HFT has the opportunity to submit an aggressive order and then observe any subsequent change in resting depth. The HFT cannot observe the liquidity state directly, but he can infer the value of Λ from changes in resting depth if he places an aggressive order; the HFT can conclude that $\Lambda = u$ if resting depth further depletes following his order, and $\Lambda = A$ otherwise. If the HFT does not place an aggressive order in period 1, he does not learn Λ.

Period 2 At the start of period 2, the HFT observes the signal of future aggressive order-flow, φ. The HFT observes φ regardless of whether he placed an aggressive order in period 1. After the HFT observes φ, he once again has an opportunity to place an aggressive order. Finally, after the HFT has the chance to trade, aggressive order-flow characterized by φ arrives, and prices change as determined by φ and Λ in equation (1).

Conceptually, the HFT's automatic observation of φ corresponds to the notion that aggressive order-flow is easy to predict on the basis of public market data. The HFT can always condition his period-2 trading strategy on φ, but he can condition this strategy on Λ only if he placed an aggressive order in period 1.

2.2 Exploratory Information is Valuable

The baseline model of exploratory trading illustrates why exploratory information can be valuable, and it highlights the trade-off between the direct costs of placing an exploratory order and the informational gains from exploration.

2.2.1 Solving the Baseline Model

Period 2 If the HFT learned the liquidity state during period 1, his optimal aggressive order in period 2 will depend on the values of both φ and Λ. The HFT's optimal strategy when he knows Λ is to set $q_2 = \varphi N$ if $\Lambda = U$, and to set $q_2 = 0$ if $\Lambda = A$. Taking expectations with respect to φ and then Λ, we find

$$\begin{aligned} \mathbb{E}\left[\pi_2 | \Lambda \ known\right] &= Nv\left(1-\alpha\right)^* u + 0^* \left(1-u\right) \\ &= Nvu\left(1-\alpha\right) \end{aligned} \tag{3}$$

If the HFT did not learn the liquidity state during period 1, his (constrained) optimal aggressive order in period 2 will still depend on the value of φ, but it will only depend on the *distribution* of Λ, rather that the actual value of Λ. The HFT's optimal strategy when he does not know Λ is to set $q_2 = \varphi N$ when $u \geq \alpha$, and to set $q_2 = 0$ when $\alpha > u$. I assumed for simplicity that $\alpha > u$, so

$$\mathbb{E}\left[\pi_2 | \Lambda \ unknown\right] = 0 \tag{4}$$

Period 1 At the start of period 1, the HFT knows neither φ nor Λ, but he faces the same trading costs (α per contract) as in period 2. Consequently, the HFT's expected direct trading profits from a period-1 aggressive order are negative, and given by

$$\mathbb{E}[\pi_1] = -\alpha |q_1|$$

Since there is no noise in this baseline model, and the HFT learns Λ perfectly from any aggressive order that he places in the first period, we can restrict attention to the cases of $q_1 = 0$ and $|q_1| = 1$.

We obtain the following expression for the difference in the HFT's total expected profits if he sets $|q_1| = 1$ instead of $q_1 = 0$:

$$\mathbb{E}[\pi_{total} | |q_1| = 1] - \mathbb{E}[\pi_{total} | q_1 = 0] = Nvu(1-\alpha) - \alpha \tag{5}$$

The HFT engages in exploratory trading if he sets $|q_1| = 1$, and he does not engage in exploratory trading if he sets $q_1 = 0$, so equation (5) represents the expected net gain from exploration. Exploratory trading is optimal for the HFT when this expected net gain is positive.

2.2.2 Conditions for Exploratory Trading

The results in section 2.2.1 demonstrate the trade-off between direct trading costs and informational gains at the heart of exploratory trading. By placing a (costly) aggressive order in period 1, the HFT "buys" the perturbation needed to elicit a response in resting depth that reveals the liquidity state. Knowing the liquidity state enables the HFT, in period 2, to better determine whether placing an aggressive order will be profitable. Parameters of the model determine the relative costs and payoffs of exploration.

Recall that when the exogenous aggressive order-flow is described by $\varphi = 0$, the HFT does not have any profitable period-2 trading opportunities in either liquidity state. The probability that $\varphi \neq 0$, given by the parameter v, represents the extent to which the exogenous aggressive order-flow is predictable. To characterize how various parameters affect the viability of exploratory trading, I consider the minimal value of v for which the HFT finds it optimal to engage in period-1 (i.e., exploratory) trading. Denoting this minimal value by $\underline{v}$, we have

$$\underline{v} = \left(\frac{\alpha}{u}\right) \frac{1}{(1-\alpha) N} \tag{6}$$

The closer is $\underline{v}$ to 0, the more conducive are conditions to exploratory trading.

The implications of equation (6) are intuitive. First, higher trading costs (α) tend to discourage exploratory trading. Second, when the HFT can use exploratory information to guide larger orders, the gains from exploration are magnified, so larger values of N tend to promote exploratory trading. Finally exploratory trading becomes less viable when u is smaller. The HFT will take the same action in period 2 when he knows that $\Lambda = A$ as when he doesn't know Λ, so when u is small, knowledge of the liquidity state is less valuable because it is less likely to change the HFT's period-2 actions.[3]

Given the dearth of exogenous variation in the real-world analogues of α, N and u, the comparative statics above do not readily translate into empirically testable predictions. However, the model generates a much more fundamental prediction that can be tested empirically: *if an agent is engaging in exploratory trading, then the market response following his exploratory orders should help to explain his performance on subsequent aggressive orders.* The market response after a trader's exploratory orders should help to forecast price movements, and the trader will tend to follow up by placing further aggressive orders in the appropriate direction when the expected price movement is sufficiently large. Note that because the follow-up orders will tend to be larger than the exploratory orders, the market response after an agent's exploratory orders should help to explain not only the performance, but also the *incidence* of his larger aggressive orders.

2.3 Private Gains from Exploratory Trading

The baseline model of exploratory trading presented above abstracted away from the details of the HFT's inference about Λ. This simplifying assumption does not qualitatively affect the central result about the value of exploratory information, but it obscures why the HFT learns more from placing an aggressive order himself than he does from merely observing an aggressive order placed by someone else.

Factors other than aggressive order arrivals can affect the behavior of resting depth. In particular, a trader may adjust her passive orders in response to new information. Just as a trader might place an aggressive buy order if he believes that prices are too low, so might another trader who shared this belief cancel some of her passive sell orders. As a result, changes in resting depth are typically correlated with aggressive order-flow, even when the aggressive orders do not actually cause those changes. However, changes in resting depth not caused by aggressive orders do not help to forecast

[3]When $u > \alpha$, the HFT will take the same action in period 2 when he knows that $\Lambda = U$ as when he doesn't know Λ, so knowledge of the liquidity state is less likely to change the HFT's period-2 actions when u is large. In the case of $u > \alpha$, equation (6) becomes $\underline{v} = \frac{1}{(1-u)N}$, and exploratory trading indeed becomes less viable as u approaches 1.

the price impact of future aggressive order-flow. The HFT learns more from aggressive orders that he places himself than he learns from those placed by other traders because he can better infer causal effects from his own orders.

2.3.1 Intuition

An analogy to street traffic illustrates the main intuition for why the HFT obtains additional information from an aggressive order that he himself places. Consider a stoplight that tends to turn green shortly before a car arrives at it. This could arise for two reasons. First, the stoplight could operate on a timer, and cars might tend to approach the stoplight just before it turns green, due (e.g.) to the timing pattern of other traffic signals in the area. Alternatively, the stoplight might operate on a sensor that *causes* it to typically turn green when a car approaches.

A driver who knows why she arrived at the stoplight at a certain time has a greater capacity to distinguish between the two explanations than does a pedestrian standing at the stoplight. In particular, if a driver knows that the moment of her arrival at the stoplight was not determined by the timing pattern of nearby traffic signals (e.g., if she had been parked, and the stoplight was the first traffic signal that she encountered), she will learn considerably more from her observation of the stoplight than will the pedestrian. Both pedestrian and driver can update their beliefs, but the pedestrian only weights the new observation by the average probability that the driver's arrival did not depend on the timing pattern of nearby signals.

Much as the driver's private knowledge about why she approaches the stoplight at a certain moment enables her to learn more than the pedestrian, the HFT's private knowledge of why he places an aggressive order enables him to learn more from the subsequent market response than he could learn from the response to an aggressive order placed by someone else.

2.3.2 Formalizing the Intuition

To make the preceding intuition more rigorous, consider a variant of the baseline model from section 2.1 in which some trader other than the HFT places an aggressive order at the beginning of period 1. With probability ρ, this aggressive order is the result of an unobservable informational shock, and resting depth further depletes following the order, regardless of the liquidity state Λ. Otherwise (with probability $1-\rho$) resting depth further depletes after the order if and only if the liquidity state is unaccommodating. Aside from this new aggressive order, all other aspects of the baseline model remain unchanged.

If the HFT places an aggressive order in period 1, his expected total profits are the same as they were in the baseline model, i.e.,

$$\mathbb{E}\left[\pi_{total} \middle| |q_1| = 1\right] \quad = \quad Nvu\left(1-\alpha\right) - \alpha$$

However, the HFT's expected profits if he does not place an order in period 1 are higher than in the baseline model, because the HFT now learns something from the depth changes following the other trader's aggressive order. If resting depth weakly replenishes after that order, the HFT learns with certainty that the liquidity state is accommodating (i.e., $\Lambda = A$), so the HFT will not submit an aggressive order in period 2, and his total profits will be zero. Alternatively, if resting depth further depletes following the other trader's aggressive order, we have

$$\mathbb{P}\left\{\Lambda = U | resting\ depth\ further\ depletes\right\} \quad = \quad \frac{u}{u + \rho\left(1-u\right)}$$

The HFT's optimal strategy when he does not know Λ is to set $q_2 = \varphi N$ when $\frac{u}{u+\rho(1-u)} \geq \alpha$, and to set $q_2 = 0$ otherwise. Taking expectations with respect to Λ and φ, we find that the HFT's *ex-ante* expected total profits in this case are given by

$$\mathbb{E}\left[\pi_{total} | AO\ by\ someone\ else\right] = \max\left\{Nv\left(\frac{u}{u+\rho\left(1-u\right)} - \alpha\right), 0\right\} \tag{7}$$

2.3.3 Analysis

The features of the baseline model discussed in section 2.2.2 are qualitatively unchanged in the modified version, but now the "privacy" parameter ρ also exerts an influence. In the limiting case where the depth change following an aggressive order placed by someone else is completely uninformative to the HFT (i.e., $\rho = 1$), equation (7) collapses down to equation (4) from the baseline model. At the opposite extreme, when the HFT learns the liquidity state perfectly from observing another trader's aggressive order (i.e., $\rho = 0$), the HFT's expected total profits are unambiguously lower if he places an aggressive order in period 1 himself.

When the HFT can learn more about the liquidity state through mere observation, as he can when ρ is smaller, he has less incentive to incur the direct costs of exploratory trading. Viewed differently, if the HFT *does* find it optimal to engage in exploratory trading, it must be the case that he obtains more useful information from the market response to his aggressive orders than he does from the market response to other traders' aggressive orders. By symmetry, it must also be the case that each other trader obtains no more useful information from the market response to the HFT's aggressive orders than they do from the market response to another arbitrary trader's aggressive orders.

2.4 Testable Predictions

Before attempting any empirical evaluation of the exploratory trading model's predictions, two basic issues must be addressed. First, it must be determined which HFTs, if any, actually earn positive and abnormal profits from their aggressive trading. I address this matter in section 4.2, and I identify eight such HFTs, to whom I refer as "A-HFTs." Next, among the A-HFTs' aggressive orders, suitable candidates for putative exploratory orders must be identified in some manner. The results from section 2.2.1 suggest that small, unprofitable aggressive orders are prime candidates. In section 4.4, I find that all of the A-HFTs, indeed, tend to lose money on their smallest aggressive orders, consistent with the theory that these orders are placed for exploratory ends.

With these two preliminary matters resolved, I turn to direct empirical tests of the model's key predictions. As a benchmark, I consider the market response following the last small aggressive order placed by anyone, which is public information. The empirical implications discussed earlier in this section can then be condensed into two central predictions, namely that relative to the public-information benchmark, the market response following an A-HFT's small aggressive order:

PREDICT.1. EXPLAINS A SIGNIFICANT ADDITIONAL COMPONENT OF THAT A-HFT'S EARNINGS ON SUBSEQUENT AGGRESSIVE ORDERS, BUT

PREDICT.2. DOES NOT EXPLAIN ANY ADDITIONAL COMPONENT OF OTHER TRADERS' EARNINGS ON SUBSEQUENT AGGRESSIVE ORDERS

In section 4.3, I make rigorous the notion of "explaining earnings on subsequent aggressive orders," then in section 5, I introduce an explicit numeric measure of "market response" and formally test the predictions above.

3 High-Frequency Trading in the E-mini Market

The E-mini S&P 500 futures contract is a cash-settled instrument with a notional value equal to $50.00 times the S&P 500 index. Prices are quoted in terms of the S&P 500 index, at minimum increments, "ticks", of 0.25 index points, equivalent to \$12.50 per contract. E-mini contracts are created directly by buyers and sellers, so the quantity of outstanding contracts is potentially unlimited.

All E-mini contracts trade exclusively on the CME Globex electronic trading platform, in an order-driven market. Transaction prices/quantities and changes in aggregate depth at individual price levels in the orderbook are observable through a public market-data feed, but the E-mini market provides full anonymity, so the identities of the traders responsible for these events are not released. Limit orders in the E-mini market are matched according to strict price and time priority; a buy (sell) limit order at a given price executes ahead of all buy (sell) limit orders at lower (higher) prices, and buy (sell) limit orders at the same price execute in the sequence that they arrived. Certain modifications to a limit order, most notably size increases, reset the time-stamp by which time-priority is determined.

E-mini contracts with expiration dates in the five nearest months of the March quarterly cycle (March, June, September, December) are listed for trading, but activity typically concentrates in the contract with the nearest expiration. Aside from brief maintenance periods, the E-mini market is open 24 hours a day, though most activity occurs during "regular trading hours," namely, weekdays between 8:30 a.m. and 3:15 p.m. CT.

3.1 Description of the Data

I examine account-labeled, millisecond-timestamped records at the Commodity Futures Trading Commission of the so-called "business messages" entered into the Globex system between September 17, 2010 and November 1, 2010 for all E-mini S&P 500 futures contracts. This message data captures not only transactions, but also events that do not directly result in a trade, such as the entry, cancellation, or modification of a resting limit order. Essentially, business messages include any action by a market participant that could potentially result in or affect a transaction immediately, or at any point in the future.[4] I restrict attention to the December-expiring E-mini contract (ticker ESZ0). During my sample period, ESZ0 activity accounted for roughly 98% of the message volume across all E-mini contracts, and more than 99.9% of the trading volume.

The price of an ESZ0 contract during this period was around \$55,000 to \$60,000, and (one-sided) trading volume averaged 1,991,252 contracts or approximately \$115 billion per day. Message volume averaged approximately 5 million business messages per day.

[4]Excluded from these data are purely administrative messages, such as log-on and log-out messages. The good-'til-cancel orders in the orderbook at the start of September 2, and a small number of modification messages (around $2-4\%$) are also missing from these records. Because I restrict attention to aggressive orders, and I only look at *changes* in resting depth (rather than its actual level), my results are not sensitive to these omitted messages.

3.2 Defining "High-Frequency Traders"

Kirilenko *et al.* identify as HFTs those traders who exhibit minimal accumulation of directional positions, high inventory turnover, and high levels of trading activity. I, too, use these three characteristics to define and identify HFTs. To quantify an account's accumulation of directional positions, I consider the magnitude of changes in end-of-day net position as a percentage of the account's daily trading volume. Similarly, I use an account's maximal intraday change in net position, relative to daily volume, to measure inventory turnover. Finally, I use an account's total trading volume as a measure of trading activity.

I select each account whose end-of-day net position changes by less than 6% of its daily volume, and whose maximal intraday net position changes are less than 20% of its daily volume. I rank the selected accounts by total trading volume, and classify the top 30 accounts as HFTs. The original classifications of Kirilenko *et al.* and Baron *et al.* guided the rough threshold choices for inter-day and intraday variation. Thereafter, since confidentiality protocols prohibit disclosing results for groups smaller than eight trading accounts, the precise cutoff values of 6%, 20%, and 30 accounts were chosen to ensure that all groups of interest would have at least eight members. My central results are not sensitive to values of these parameters.

The set of HFTs corresponds closely to the set of accounts with the greatest trading volume in my sample, so the set of HFTs is largely invariant both to the exact characterizations of inter-day and intraday variation in net position relative to volume, and to the exact cutoff values for these quantities. Similarly, changing the 30-account cutoff to (e.g.) 15 accounts or 60 accounts does not substantially alter my results, because activity heavily concentrates among the largest HFTs. For example, the combined total trading volume of the 8 largest HFTs exceeds that of HFTs 9-30 by roughly three-quarters, and the combined aggressive volume of the 8 largest HFTs exceeds that of HFTs 9-30 by a factor of almost 2.5.

3.3 HFTs' Prominence and Profitability

Although HFTs constitute less than 0.1% of the $41,778$ accounts that traded the ESZ0 contract between September 17, 2010 and November 1, 2010, they participate in 46.7% of the total trading volume during this period. In addition to trading volume, HFTs are responsible for a large fraction of message volume. During the sample period, HFTs account for 31.9% of all order entry, order modification and order cancellation messages. The HFTs also appear to earn large and stable profits.

Gross of trading fees, the 30 HFTs earned a combined average of $1.51 million per trading day during the sample period. Individual HFTs' annualized Sharpe ratios are in the neighborhood of 10 to 11.

The Chicago Mercantile Exchange reduces E-mini trading fees on a tiered basis for traders whose average monthly volume exceeds various thresholds. Trading and clearing fees were either $0.095 per contract or $0.12 per contract for the 20 largest HFTs, and were at most $0.16 per contract for the remaining HFTs. Initial and maintenance margins were both $4,500 for all of the HFTs.

Hereafter, unless otherwise noted, I restrict attention to activity that occurred during regular trading hours. HFTs' aggressive trading occurs almost exclusively during regular trading hours (approximately 95.6%, by volume), and market conditions during these times differ substantially from those during the complementary off-hours.

4 HFTs' Profits from Aggressive Orders

Aggressive trading is a tremendously important component of HFTs' activity. In aggregate, approximately 48.5% of HFTs' volume is aggressive, and this figure rises to 54.2% among the 12 largest HFTs. Furthermore, many HFTs consistently profit from their aggressive trading. Since the bid-ask spread in the E-mini market rarely exceeds the minimum imposed upon it by the granularity of prices, there is little mystery about how a trader's passive trades could consistently earn money.[5] By contrast, explaining how a trader who uses only market data could consistently profit on aggressive trades is somewhat difficult.

4.1 Measuring Aggressive Order Profitability

Because all E-mini contracts of a given expiration date are identical, it is neither meaningful nor possible to distinguish among the individual contracts in a trader's inventory, so there is generally no way to determine the exact prices at which a trader bought and sold a particular contract. As a result, it is typically impossible to measure directly the profits that a trader earns on an individual aggressive order. However, the cumulative price change following an aggressive order, normalized by the order's direction ($+1$ for a buy, or -1 for a sell), can be used to construct a meaningful proxy for the order's profitability. Intuitively, the average expected profit from an aggressive order equals

[5] Explaining the profitability of individual passive trades does not resolve the question of how various HFTs manage to participate in so many passive trades. In equilibrium, we would expect new entrants to reduce the average passive volume of an individual trader until her total profits from passive trades equaled her fixed costs.

the expected favorable price movement, minus trading/clearing fees and half the bid-ask spread. See Appendix B for rigorous justification.

Estimating the cumulative favorable price movement after an aggressive order is straight-forward. Consider a trader who can forecast price movements up to j time periods in the future, but no further. If the trader places an aggressive order in period t, any price changes that she could have anticipated at the time she placed the order will have occurred by period $t+j+1$. Provided that price is a martingale with respect to its natural filtration, the expected change in price from period $t+j+1$ onward is zero, both from the period$-t$ perspective of the trader and from an unconditional perspective. Thus the change in price between period t and any period after $t+j$, normalized by the direction of the trader's order ($+1$ for a buy, or -1 for a sell), will provide an unbiased estimate of the favorable price movement following the trader's order.

The remarks above imply that we can derive a proxy for the profitability of an HFT's aggressive order using the (direction-normalized) accumulated price-changes following that aggressive order out to some time past the HFT's maximum forecasting horizon. If we choose too short an accumulation window, the resulting estimates of the long-run direction-normalized average price changes following the HFT's aggressive orders will be biased downward. As a result, we can empirically determine an adequate accumulation period by calculating cumulative direction-normalized price changes over longer and longer windows until their mean ceases to significantly increase. Using too long an accumulation period introduces extra noise, but it will not bias the estimates. I find that an accumulation period, measured in event-time, of 30 aggressive order arrivals is sufficient to obtain unbiased estimates; for all of the empirical work in this paper, I use an accumulation period of 50 aggressive order arrivals to allow a wide margin for error. See Appendix B for further details.

As noted earlier, the bid-ask spread for the E-mini is almost constantly \$12.50 (one tick) during regular trading hours, and the HFTs in my sample face trading/clearing fees of \$0.095 to \$0.16 per contract, so the average favorable price movement necessary for an HFT's aggressive order to be profitable is between \$6.345 and \$6.41 per contract. Since trading/clearing fees vary across traders, I report aggressive order performance in terms of favorable price movement, that is, earnings gross of fees and the half-spread.

4.2 HFTs' Overall Profits from Aggressive Orders

To measure the overall mean profitability of a given account's aggressive trading, I compute the average cumulative price change following each aggressive order placed by that account, weighted by executed quantity and normalized by the direction of the aggressive order. As a group, the 30 HFTs in my sample achieve size-weighted average aggressive order performance of $7.01 per contract. On an individual basis, nine HFT accounts exceed the relevant $\$6.25 + fees$ profitability hurdle, and each of these nine accounts exceeds this hurdle by a margin that is statistically significant at the 0.05 level. One of these nine accounts is linked with another HFT account, and their combined average performance also significantly exceeds the profitability hurdle.

Overall, the HFTs vastly outperform non-HFTs, who earn a gross average of $3.19 per aggressively-traded contract. However, these overall averages potentially confound effects of very coarse differences in the times at which traders place aggressive orders with effects of the finer differences more directly related to strategic choices. For example, if all aggressive orders were more profitable between 1 p.m. and 2 p.m. than at other times, and HFTs only placed aggressive orders during this window, the HFTs' outperformance would not depend on anything characteristically high-frequency.

To control for potential low-frequency confounds, I divide each trading day in my sample into 90-second segments and regress the profitability of non-HFTs' aggressive orders during each segment on both a constant and the executed quantities of the aggressive orders. Using these local coefficients, I compute the profitability of each aggressive order by an HFT in excess of the expected profitability of a non-HFT aggressive order of the same size during the relevant 90-second segment. With these additional controls, only 27 HFT accounts continue to exhibit significant outperformance of non-HFTs, and only eight of the 27 accounts are among those whose absolute performance exceeded the profitability hurdle.

4.2.1 A-HFTs and B-HFTs

For expositional ease, I will refer to the eight HFT accounts that make money on their aggressive trades *and* outperform the time-varying non-HFT benchmark as "A-HFTs," and to the complementary set of HFTs as "B-HFTs." The eight A-HFTs have a combined average daily trading volume of 982,988 contracts, and on average, 59.2% of this volume is aggressive. The 22 B-HFTs have a combined average daily trading volume of 828,924 contracts, of which 35.9% is aggressive. Gross of fees, the A-HFTs earn a combined average of $793,342 per day, or an individual average of $99,168 per day, while the

B-HFTs earn a combined average of $\$715,167$ per day, or an individual average of $\$32,508$ per day.[6] The highest profitability hurdle among the A-HFTs is \$6.37 per aggressively traded contract.

4.3 Relative Aggressive Order Profitability: HFT vs. Econometrician

To gain some insight into the factors that affect aggressive trading profits, I examine the extent to which econometric price forecasts explain the realized performance of aggressive orders placed by A-HFTs, B-HFTs, and non-HFTs. The methodology that I develop in this subsection also provides the starting point for my direct tests of the exploratory trading model's predictions in section 5.

4.3.1 Variables that Forecast Price Movements

Bid-ask bounce notwithstanding, the price at which aggressive orders execute changes rather infrequently in the E-mini market. On average, only about $1-3\%$ of aggressive buy (sell) orders execute at a final price different from the last price at which the previous aggressive buy (sell) order executed, and the price changes that do occur are almost completely unpredictable on the basis of past price changes. However, several other variables forecast price innovations surprisingly well.

In contrast to price innovations, the direction of aggressive order flow in the E-mini market is extremely persistent. On average, the probability that the next aggressive order will be a buy (sell) given that the previous aggressive order was a buy (sell) is around 75%. In addition to forecasting the direction of future aggressive order flow, the direction of past aggressive order flow also forecasts future price innovations to statistically and economically significant extent, and forecasts based on past aggressive order signs alone are modestly improved by information about the (signed) quantities of past aggressive orders. Price forecasts can be further improved using simple measures of recent changes in the orderbook.

4.3.2 Econometric Benchmark

For each trading day in my sample, I regress the cumulative price-change (in dollars) between the aggressive orders k and $k+50$, denoted y_k, on lagged market variables suggested by the remarks above. Specifically, I regress y_k on the changes in resting depth between aggressive orders $k-1$ and k at each of the six price levels within two ticks of the best bid or best ask, the signs of aggressive orders $k-1$ through $k-4$, and the signed executed quantities of aggressive orders $k-1$ through $k-4$.

[6] All of the preceding descriptive statistics include the small amount of trading activity that occurred outside regular trading hours.

For symmetry, I adopt the convention that sell depth is negative, and buy depth is positive, so that an increase in buy depth has the same sign as a decrease in sell depth. Denoting the row vector of the 14 regressors by z_{k-1}, and a column vector of 14 coefficients by Γ, I estimate the equation

$$y_k \quad = \quad z_{k-1}\Gamma + \epsilon_k \tag{8}$$

Appendix C presents coefficient estimates and direct discussion of the regression results.

To compute the excess performance of aggressive order k, denoted ξ_k, I normalize the kth regression residual by $sign_k$, the sign of the kth aggressive order:

$$\xi_k = sign_k \left(y_k - z_{k-1}\hat{\Gamma} \right)$$

As discussed in section 4.1, normalizing the cumulative price-change y_k by the sign of the kth aggressive order yields a measure of the kth aggressive order's profitability. Likewise, the quantity ξ_k provides a measure of kth aggressive order's profitability in excess of that expected on the basis of the benchmark econometric specification. I compute the vectors of direction-normalized residuals separately for each of the 32 trading days in my sample, then combine all of them into a single vector for the entire sample period.

4.3.3 Explained Performance

The price movements predicted by (8) explain a substantial component of the performance of aggressive orders placed by A-HFTs, B-HFTs, and non-HFTs alike.[7] Looking ahead, this explanatory power validates the use of specification (8) as a basis for the more sophisticated analyses in section 5. Figure 1 and Table 1, below, summarize the overall size-weighted average performance of aggressive orders placed by various trader groups, both in absolute terms, and in excess of the econometric benchmark. Confidence intervals are computed via bootstrap.

[7] Although the variables in z_{k-1} are all observable before the kth aggressive order arrives, the fitted value $z_{k-1}\hat{\Gamma}$ is not literally a forecast of y_k in the strictest sense, as $\hat{\Gamma}$ is estimated from data for the entire day. However, the coefficient estimates are extremely stable throughout the sample period, so thinking of $z_{k-1}\hat{\Gamma}$ as a forecast of y_k is innocuous in the present setting. See section 6.1.

Figure 1. Aggressive Order Performance Relative to Econometric Benchmark

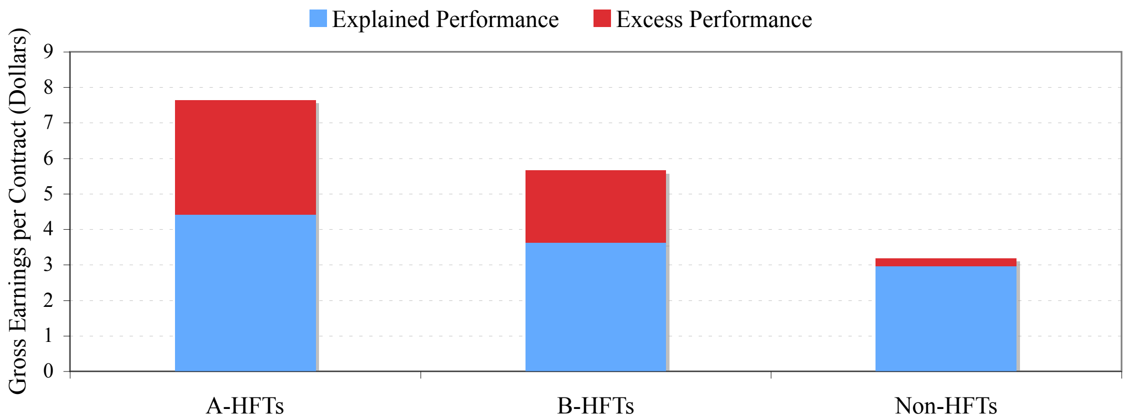


Table 1. Aggressive Order Performance vs. Econometric Benchmark

	Mean Absolute	Mean Absolute 99% CI	Mean Excess	Mean Excess 99% CI	Explained Performance
A-HFTs	7.65	(7.54, 7.73)	3.22	(3.09, 3.36)	57.8%
B-HFTs	5.67	(5.57, 5.77)	2.04	(1.90, 2.17)	64.0%
Non-HFTs	3.19	(3.12, 3.26)	0.22	(0.14, 0.29)	93.06%

The exploratory trading model developed earlier assumed that A-HFTs ultimately traded ahead of easily predictable demand innovations (when liquidity conditions were suitably unaccommodating), and the explanatory power of equation (8) for the A-HFTs' performance substantiates this assumption. At the same time, although the econometric controls explain over half of the A-HFTs' performance on their aggressive orders, the remaining unexplained component of performance is massive. The A-HFTs' average excess performance is over 50% greater than that of the B-HFTs, and over 10 times greater than that of the non-HFTs. Price forecasts more sophisticated that those from (8) may better explain A-HFTs' performance; I return to this matter in section 5.

4.4 A-HFTs' Losses on Small Aggressive Orders

A-HFTs' aggressive orders tend to become more profitable as order size increases.[8] In fact, despite earning money from their aggressive orders on average, the A-HFTs all tend to *lose* money on the

[8]This effect appears whether price-changes are measured between the respective last prices at which successive aggressive orders execute (correcting for bid-ask bounce), or between the respective first prices at which they execute, so the positive relationship between executed quantity and subsequent favorable price movements is not simply an artifact of large orders that eat through one or more levels of the orderbook.

smallest aggressive orders that they place. Note also that I refer here to the size of the aggressive orders A-HFTs submit, not the quantity that executes, so the small orders were intentionally chosen to be small, and the large orders intentionally chosen to be large.

The baseline model of exploratory trading in section 2 produces exactly the sort of losses on small aggressive orders and profits on large aggressive orders that the A-HFTs exhibit. The A-HFTs' differing performance on small and large aggressive orders is consistent with the pattern that we would expect to see if the small orders were generating valuable information that enabled the A-HFTs to earn greater profits from their large orders.

To make precise both the meaning of "small" aggressive orders, and A-HFTs' losses on them, I specify cutoffs for order size and compute the average performance of A-HFTs' aggressive orders below and above those size cutoffs. Figure 2 and Table 2, below, display bootstrap confidence intervals for the executed-quantity-weighted average performance of A-HFTs' aggressive orders weakly below and strictly above various order-size cutoffs.

Figure 2. A-HFT Performance on Small and Larger Aggressive Orders (95% Conf. Intervals)

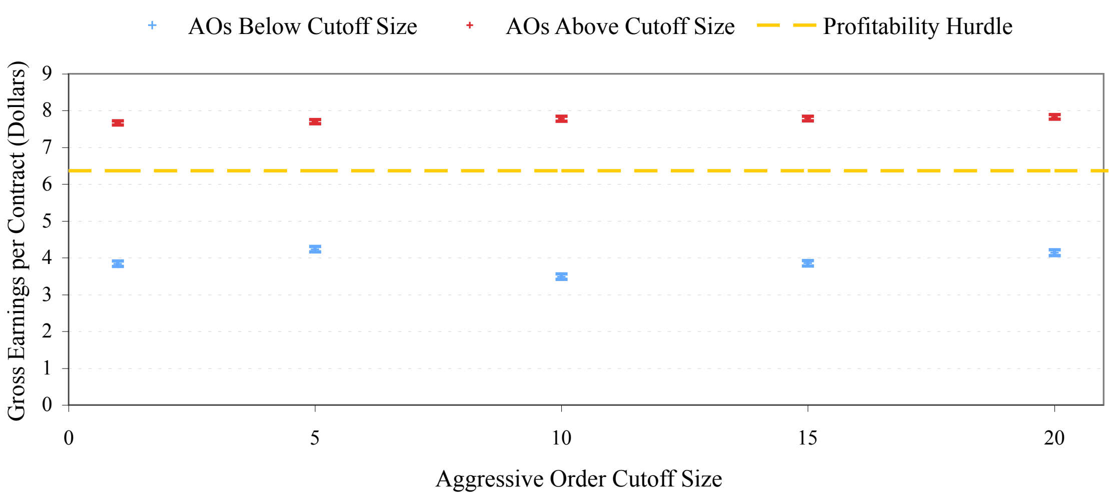


Table 2. Performance of A-HFTs' Aggressive Orders (Dollars per Contract)

Cutoff	Below Cutoff 95% CI	Above Cutoff 95% CI	AOs Below Cutoff % of All AOs	AOs Below Cutoff % of Aggr. Volume
1	$(3.78, 3.89)$	$(7.59, 7.74)$	24.31%	0.40%
5	$(4.17, 4.29)$	$(7.62, 7.78)$	43.74%	1.44%
10	$(3.42, 3.55)$	$(7.71, 7.85)$	54.64%	3.09%
15	$(3.79, 3.92)$	$(7.71, 7.86)$	56.75%	3.54%
20	$(4.08, 4.20)$	$(7.75, 7.90)$	60.82%	4.80%

As shown in Table 2, small aggressive orders represent a substantial fraction of the aggressive orders that A-HFTs place, but these small orders make up very little of the A-HFTs' total aggressive volume. Nevertheless, A-HFTs' losses on these small orders are non-negligible. On average, each A-HFT loses roughly $7,150 per trading day ($1.8 million, annualized) on aggressive orders of size 20 or less; this loss represents approximately 7.2% of an average A-HFT's daily profits.

Although HFTs may tolerate only limited levels of inventory, inventory-management does not adequately explain the A-HFTs' losses on small aggressive orders. We can control for A-HFTs' respective net positions at the times they submit aggressive orders, and restrict attention to only those aggressive orders that move an A-HFT *away* from a zero net position. Such "non-rebalancing" orders account for over half of the small aggressive orders that A-HFTs place, and they cannot possibly be motivated by inventory management. Nevertheless, A-HFTs still lose money on the smallest of these orders, yet make money on the larger ones. See Table D.1 in Appendix D.

The A-HFTs' qualitative pattern of losses on small aggressive orders and more-than-offsetting gains on larger aggressive orders suggests that the small orders are reasonable candidates for exploratory orders. This finding provides a foundation for direct tests of the exploratory trading model's sharper empirical predictions.

5 Explicitly Isolating Exploratory Information

If the exploratory trading model is correct, and if the A-HFTs' small aggressive orders are indeed exploratory in nature, the two key model predictions presented in section 2.4 must hold. For convenience, I summarize these predictions below.

Relative to a benchmark that incorporates the public information about the market response following small aggressive orders placed by anyone, the market response following small aggressive orders placed by an A-HFT:

Predict.1. Explains a significant additional component of that A-HFT's earnings on subsequent aggressive orders, but

Predict.2. Does not explain any additional component of other traders' earnings on subsequent aggressive orders

In this section, I consider a simple numeric characterization of the market response following an aggressive order, and I directly test whether the above predictions of the exploratory trading model hold. I estimate results for the A-HFTs individually, but for compliance with confidentiality protocols, I present cross-sectional averages of these estimates. Empirically, these average results are representative of the results for individual A-HFTs.[9]

5.1 Empirical Strategy: Overview

Though the implementation is slightly involved, my basic empirical strategy is straight-forward. First, I augment the benchmark regression from section 4.3 using

1. Market response information from the last small aggressive order placed by anyone, and
2. Both market response information from the last small aggressive order placed by anyone, AND market response information from the last small aggressive order *placed by a specified A-HFT*

As I discuss in more detail in the next subsection, the market-response variable that I consider essentially amounts to a measure of the change in orderbook depth that follows an aggressive order.

After estimating both of the specifications above, I find the additional component of performance on larger aggressive orders explained by (2) relative to (1). The market response following an arbitrary small aggressive order is publicly observable. However, because the E-mini market operates anonymously, the distinction between a small aggressive order placed by a particular A-HFT and an arbitrary small aggressive order is private information, available only to the A-HFT who placed the

[9] Throughout the E-mini market, there exist assorted linkages between various trading accounts (as, for example, in the simple case where single firm trades with multiple accounts), so the trading-account divisions do not necessarily deliver appropriate atomic A-HFT units. Though the specifics are confidential, the appropriate partition of the A-HFTs is entirely obvious. For brevity, I use "individual A-HFT" as shorthand to "individual atomic A-HFT unit," as applicable.

order. Comparing the second specification above to the first isolates the effects attributable to this private information from effects attributable to public information.

Finally, I compare the additional explained performance for the specified A-HFT to the additional explained performance for all other traders. Intuitively, we want to verify that the A-HFT's exploratory information provides extra explanatory power for the subsequent performance of trader privy to that information (the A-HFT), but not for the performance of traders who aren't privy to it (everyone else). Note that "everyone else" includes the A-HFTs other than the specified A-HFT.

Some A-HFT accounts and B-HFT/non-HFT accounts belong to the same firms, and various B-HFTs/non-HFTs may be either directly informed or able to make educated inferences about what one or more A-HFTs do. As a result, we should not necessarily expect exploratory information generated by an A-HFT's small orders to provide no explanatory power whatsoever for all other traders' performance. However, we should still expect the additional explanatory power for the A-HFT's performance to significantly exceed that for the other traders' performance.

5.2 Empirical Implementation

Define an aggressive order to be "small" if that order's submitted size is less than or equal to a specified size parameter, which I denote by $\bar{q}$.

5.2.1 A Simple Measure of Market Response

I characterize the market response to a small aggressive order using subsequent changes in orderbook depth. I examine the interval starting immediately after the arrival of a given small aggressive order and ending immediately before the arrival of the next aggressive order (which may or may not be small), and I sum the changes in depth at the best bid and best ask that occur during this interval. As in section 4.3, I treat sell depth as negative and buy depth as positive. I also normalize these depth changes by the sign of the preceding small aggressive order to standardize across buy orders and sell orders.

To simplify the analysis and stack the deck against finding significant results, I initially focus only on the sign of the direction-normalized depth changes. Note the direct analogy to the two-liquidity-state setting of the exploratory trading model in section 2.

For a given value of $\bar{q}$, I construct the indicator variable Ω, with kth element Ω_k defined by

$$\Omega_k = \begin{cases} 1 & if\ DC\left(k; any, \bar{q}\right) > 0 \\ 0 & otherwise \end{cases}$$

where $DC\left(k; any, \bar{q}\right)$ denotes the direction-normalized depth change following the last small aggressive order (submitted by anyone) that arrived before the kth aggressive order. Similarly, I construct the indicator variable Ω^A, with kth element Ω_k^A defined by

$$\Omega_k^A = \begin{cases} 1 & if\ DC\left(k; AHFT, \bar{q}\right) > 0 \\ 0 & otherwise \end{cases}$$

where $DC\left(k; AHFT, \bar{q}\right)$ denotes the direction-normalized depth change following the last small aggressive order *submitted by a specified A-HFT* that arrived before the kth aggressive order.

5.2.2 Estimation Procedure

In the model of exploratory trading presented earlier, exploratory information was valuable only in conjunction with information about future aggressive order flow. Following this result, I incorporate market-response information by using the indicators Ω and Ω^A to partition the benchmark regression from section 4.3.

Recall that in section 4.3, I estimated the equation

$$y_k = z_{k-1}\Gamma + \epsilon_k$$

where y_k denoted the cumulative price-change between the aggressive orders k and $k+50$, and the vector z_{k-1} consisted of changes in resting depth between aggressive orders $k-1$ and k, in addition to the signs and signed executed quantities of aggressive orders $k-1$ through $k-4$. Using the indicator Ω, I now partition the equation above into two pieces and estimate the equation

$$y_k = \Omega_k z_{k-1}\Gamma^a + \left(1 - \Omega_k\right) z_{k-1}\Gamma^b + \epsilon_k \tag{9}$$

Next, I use the indicator Ω^A to further partition (9), and I estimate the equation

$$\begin{aligned} y_k &= \Omega_k^A(k)\left(\Omega_k z_{k-1}\Gamma^c + (1-\Omega_k) z_{k-1}\Gamma^d\right) + \\ &\quad \left(1-\Omega_k^A\right)\left(\Omega_k z_{k-1}\Gamma^e + (1-\Omega_k) z_{k-1}\Gamma^f\right) + \epsilon_k \end{aligned} \tag{10}$$

The variables y_k and z_{k-1} denote the same quantities as before, and the Γ^j terms each represent vectors of 14 coefficients.

I estimate (9) and (10) for $\bar{q} = 1, 5, 10, 15, 20$, and for each specification I calculate the relative excess performance of the specified A-HFT, and of all other trading accounts on aggressive orders of size strictly greater than $\bar{q}$. As in section 4.3, I compute the performance of aggressive order k in excess of that explained by each regression by normalizing the kth residual from the regression by the sign of the kth aggressive order. I now also control for order-size effects directly by regressing the direction-normalized residuals (for the orders of size strictly greater than $\bar{q}$) on the (unsigned) executed quantities and a constant, then subtracting off the executed quantity multiplied by its estimated regression coefficient. Controlling for size effects in this manner makes results more comparable for different choices of $\bar{q}$. Size effects can be addressed by other means with negligible impact on the final results.

For each aggressive order larger than $\bar{q}$ placed by the A-HFT under consideration, I compute the additional component of performance explained by (10) relative to (9) by subtracting the order's excess performance over (10) from its excess performance over (9); I stack these additional explained components in a vector that I denote by Ξ_A. I repeat this procedure to obtain the analogous vector for everyone else, Ξ_{ee}.

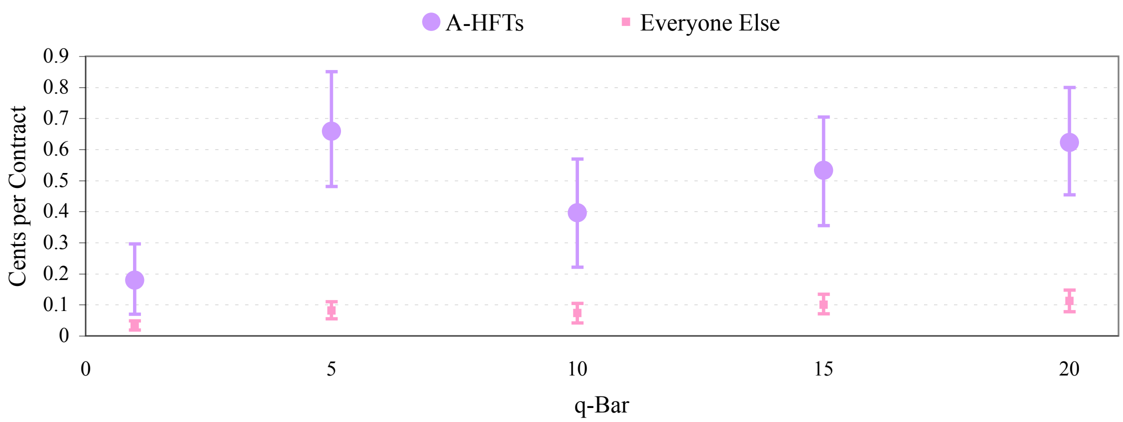
5.3 Results

I evaluate the empirical predictions of the exploratory trading model by comparing the additional explained component of performance for each A-HFT to the additional explained component of performance for all other traders. Specification (10) has more free parameters than (9), but additional explanatory power of (10) due exclusively to the extra degrees of freedom will manifest equally, in expectation, for all traders, so the extra degrees of freedom alone should not cause Ξ_A and Ξ_{ee} to differ significantly.

Figure 3, below, displays the cross-sectional means of Ξ_A and Ξ_{ee} for different values of $\bar{q}$ (see

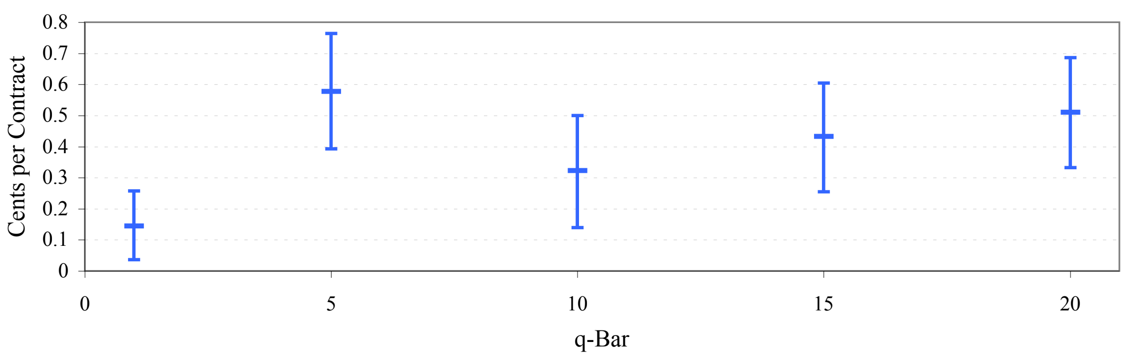
Appendix Table D.2 for numeric values). Both predictions of the exploratory trading model are borne out in these results. Using information about the market activity immediately following an A-HFT's smallest aggressive orders (in the form of Ω^A) improves our ability to explain that A-HFT's performance on larger aggressive orders by a highly significant margin, relative to using only information about the activity following any small aggressive order (in the form of Ω). By contrast, relative to using Ω alone, incorporating the information in Ω^A provides little or no significant additional explanatory power for other traders' performance on larger aggressive orders.

Figure 3. Additional Performance Explained (95% Confidence Intervals)



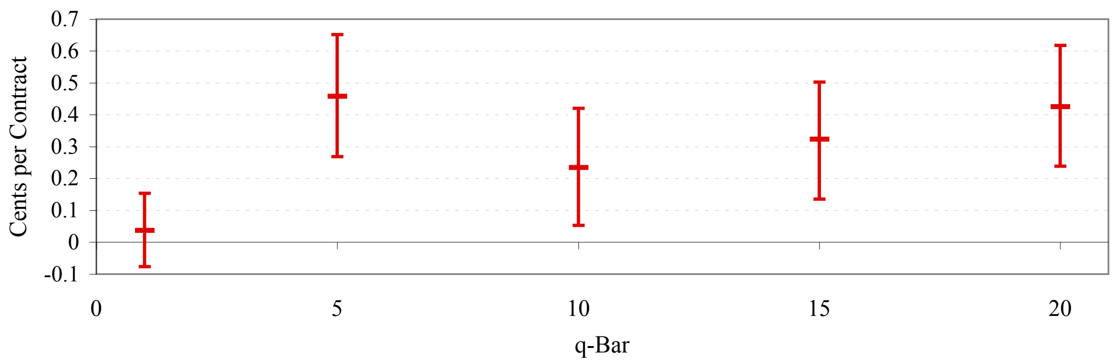
As a more formal comparison of the gain in explanatory power for the A-HFTs relative to the gain for everyone else, I construct 95% bootstrap confidence intervals for the difference of the pooled means $Mean\left(\Xi_A\right) - Mean\left(\Xi_{ee}\right)$, for $\bar{q} = 1, 5, 10, 15, 20$. Figure 4 summarizes these results, which confirm what the preceding results suggested: the extra component of A-HFTs' performance on large aggressive orders explained by using Ω^A in addition to Ω is significantly greater than the extra component explained for other traders. (See Appendix Table D.3 for numeric values)

Figure 4. [A-HFT Addt'l Explained] - [Everyone Else Addt'l Explained] (95% Conf. Intervals)



Although the extra explanatory power for an average individual A-HFT is significantly greater than that for all other traders, the amount of performance to be explained is also somewhat greater. Comparing extra explanatory power for an individual A-HFT to extra explanatory power for the complementary set of HFTs mitigates this difference. Consistent with the notion that certain B-HFTs may know something about what various A-HFTs are doing, the extra component of performance explained by using Ω^A in addition to Ω is larger for the complementary set of HFTs than it is for the broader "everyone except the A-HFT of interest" group. Nevertheless, aside from the case of $\bar{q} = 1$, the average extra explanatory power for an individual A-HFT is still significantly greater than is that for the complementary set of HFTs, as shown in Figure 5, below. (See Appendix Table D.3 for numeric values.)

Figure 5. [A-HFT Addt'l Explained] - [Other HFTs Addt'l Explained] (95% Conf. Intervals)



5.4 Incidence of A-HFTs' Larger Aggressive Orders

As suggested by the remarks at the end of section 2.2.2, the prediction that exploratory information explains a significant additional component of the A-HFTs' performance tacitly requires exploratory information to help explain the incidence of the A-HFTs' larger aggressive orders. In particular, all else being equal, the exploratory trading model predicts that an A-HFT will have a greater tendency to place large aggressive orders when $\Omega^A = 1$ than when $\Omega^A = 0$. A direct test of this prediction about the incidence of the A-HFTs' larger aggressive orders offers a robustness check on the results in subsection 5.3.

Much as the HFT in the model from section 2 considered the signal of future aggressive order-flow as well as the liquidity state, A-HFTs consider public market data as well as exploratory information to decide when to place large aggressive orders. The size and direction of A-HFTs' aggressive orders depend on the same variables that forecast price movements, or equivalently on the forecasts of price movements themselves. On average, the signed quantity of an A-HFT's aggressive order should be an increasing function of the future price-change expected on the basis of public information. In this context, the exploratory trading model predicts that the expected future price-change will have a larger effect on the signed quantity of an A-HFT's aggressive orders when $\Omega^A = 1$ than it will when $\Omega^A = 0$.

To test the exploratory trading model's prediction about the incidence of A-HFTs' larger aggressive orders, I regress the signed quantities of a given A-HFT's aggressive orders on the associated fitted values of y from equation (9), partitioned by Ω^A. In other words, for a specified A-HFT and a given value of $\bar{q}$, I estimate the equation

$$q_k = \beta_0 \left(1 - \Omega_k^A\right) \hat{y}_k + \beta_1 \Omega_k^A \hat{y}_k + \epsilon_k \tag{11}$$

where q_k denotes the signed submitted quantity of the A-HFT's kth aggressive order, $\hat{y}_k$ denotes the relevant fitted value of y_k from the public-information regression (9), and Ω^A is the usual indicator function. I restrict the β coefficients to be the same across all A-HFTs.

Table 3, below, displays the coefficient estimates from (11) for various values of $\bar{q}$. A Wald test rejects the null hypothesis $\beta_0 = \beta_1$ at the 10^{-15} level for all values of $\bar{q}$. As the exploratory trading model predicts, holding fixed the price-change expected on the basis of public information, the average A-HFT places significantly larger aggressive orders when $\Omega^A = 1$ than when $\Omega^A = 0$.

Table 3. Differential Effects of Predicted Price-Changes on A-HFT Signed Order Size

	Point Estimates		Standard Errors	
	β_0	β_1	β_0	β_1
$\bar{q} = 1$	13.35	15.26	0.094	0.162
$\bar{q} = 5$	13.41	15.11	0.093	0.169
$\bar{q} = 10$	13.42	14.97	0.095	0.160
$\bar{q} = 15$	13.34	15.10	0.095	0.159
$\bar{q} = 20$	13.23	15.30	0.094	0.160

6 Practical Significance of Exploratory Information

The losses that A-HFTs incur on their small aggressive orders offer a natural point of comparison for the gains that can be explained from the information generated by those orders. Table 4, below, displays the additional component of A-HFTs' profits on large aggressive orders directly explained using exploratory information (in the form of Ω^A) as a percentage of A-HFTs' losses on small aggressive orders.

In one sense, given the extreme simplicity and coarseness of the $\Omega-$operators as representations of exploratory information, the results in Table 4 suggest gains that are surprisingly large in practical terms. At the same time, the gains from exploration should at least weakly exceed the costs, and the additional gains directly explained using Ω^A fall short of this mark.

Table 4. Extra Explained Gains on Large AOs vs. Losses on Small AOs

	$(Extra\ Explained\ Gains)\ /\ \lvert Losses\rvert$
$\bar{q} = 1$	33.05%
$\bar{q} = 5$	11.27%
$\bar{q} = 10$	4.48%
$\bar{q} = 15$	5.39%
$\bar{q} = 20$	4.79%

Representations of exploratory information richer than Ω^A are extremely easy to construct. For example, an obvious extension would be to consider the not only the sign, but also the *magnitude* of the direction-normalized depth change following an exploratory order. Regardless of the particular representation of exploratory information used, though, the additional explained component of A-HFTs' profits on the aggressive orders they place is likely to understate the true gains from exploration. As the simple model in section 2 illustrates, exploratory information is valuable in large part because it enables a trader to avoid placing unprofitable aggressive orders. However, estimates of the additional explained component of profits on A-HFTs' aggressive orders necessarily omit the effects of such avoided losses. While this bias, if anything, makes the preceding findings of statistical significance all the more compelling, it also complicates the task of properly determining the practical importance of exploratory information.

6.1 Simulated Trading Strategies

To investigate the gains from exploratory information, including the gains from avoiding unprofitable aggressive orders, I examine the effects of incorporating market-response information from small aggressive orders into simulated trading strategies. The key advantage of working with these simulated trading strategies is that avoided unprofitable aggressive orders can be observed directly.

The basic trading strategy that I consider is a simple adaptation of the benchmark regression from section 4.3. I specify a threshold value, and the strategy entails nothing more than placing an aggressive order with the same sign as $\hat{y}_k$ whenever $|\hat{y}_k|$ exceeds that threshold. To make this strategy feasible (in the sense of using only information available before time t to determine the time-t action) I compute the forecast of future price movement, $\hat{y}_k$, using the regression coefficients estimated from the previous day's data. I incorporate market-response information into this strategy by modifying the rule for placing aggressive orders to, "place an aggressive order (with the same sign as $\hat{y}_k$) if and only if all three of the following conditions hold:

- $|\hat{y}_k|$ exceeds its specified threshold,
- The direction-normalized depth-change following the last small aggressive order (placed by anyone) exceeds a specified threshold, and
- The direction-normalized depth-change following the last small aggressive order *placed by an A-HFT* exceeds a (possibly different) specified threshold."

Choosing a threshold of $-\infty$ will effectively remove any of these conditions.

Each strategy yields a set of times to place aggressive orders, and the associated direction for each order. To measure the performance of a given strategy, I compute the average profitability of the indicated orders in the usual manner, with the assumption that these aggressive orders are all of a uniform size.

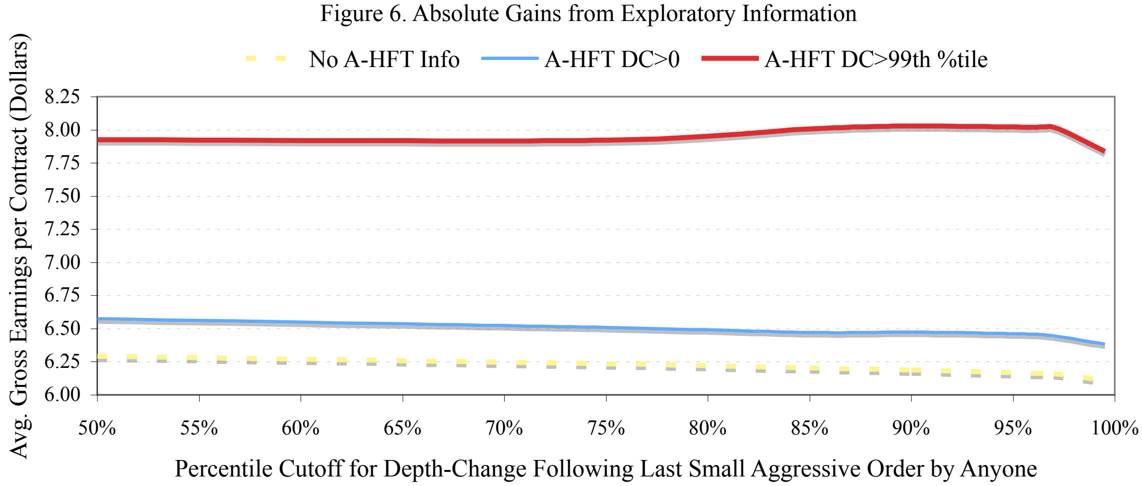
Relative to A-HFTs' losses on small aggressive orders, the additional component of A-HFTs' profits directly explained using Ω^A is smallest when $\bar{q} = 10$, and I present results for $\bar{q} = 10$ to highlight the impact of accounting for avoided losses on estimates of the gains from exploratory information. Results for other values of $\bar{q}$ are similar.

6.1.1 Three Specific Strategies

All three threshold parameters affect strategy performance, so to emphasize the role of market-response information, I present results with the threshold for $|\hat{y}_k|$ held fixed. Varying the threshold for $|\hat{y}_k|$ does not alter the qualitative results. In particular, it is not possible to achieve the same gains in performance that result from incorporating exploratory information by merely raising the threshold for $|\hat{y}_k|$. The forecast $\hat{y}_k$ uses coefficients estimated from the previous day's data, and these forecasts exhibit increasing bias as the z_{k-1} observations assume more extreme values.

I consider a range of threshold values for the direction-normalized depth-change following the last small aggressive order placed by anyone, but, for expositional clarity, I present results for three illustrative threshold choices for the direction-normalized depth-change following the last small aggressive order placed by an A-HFT. Specifically, I consider thresholds of $-\infty$ (no A-HFT market-response information), 0 (the same information contained in Ω^A), and 417 (the 99th percentile value). Figure 6 displays the performance of these three strategies over a range of threshold values for the market response following arbitrary small aggressive orders.

Figure 6. Absolute Gains from Exploratory Information



While the performance gains from incorporating A-HFT exploratory information are obvious, an equally important feature of the results above is more subtle. The A-HFTs' average gross earnings on aggressive orders over size 10 of \$7.78 per contract are well above the peak performance of the strategy that uses only public information, but substantially below the performance of the strategy that incorporates the A-HFTs' exploratory information with the higher threshold. This is exactly the pattern that we should expect, given that the former strategy excludes information that is available to the A-HFTs and the latter strategy includes information that is not available to any individual A-HFT, so these results help to confirm the relevance and validity of this simulation methodology.

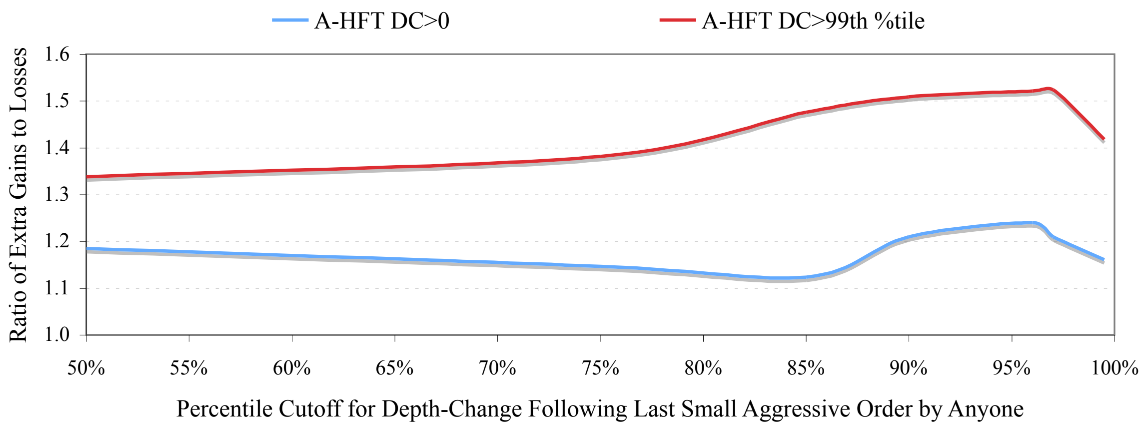
6.1.2 Gains from Exploration Relative to Losses on Exploratory Orders

Although the two strategies that incorporate exploratory information from the A-HFTs' small aggressive orders outperform the strategy that does not, the orders that generated the exploratory information were costly. To compare the gains from this exploratory information to the costs of acquiring it, I first multiply the increases in per-contract earnings for the two exploratory strategies (scaled by the respective number of orders relative to the public-information strategy) by the A-HFTs' combined aggressive volume on orders over size 10.[10] I then divide these calibrated gains by the A-HFTs' actual losses on aggressive orders size 10 and under. The resulting ratio is the direct analogue of the percentages in Table 4.

Figure 7 displays the calibrated ratio of additional gains to losses for each exploratory simulated strategy over a range of threshold values for the market response following arbitrary small aggressive orders. Using information from the A-HFTs' exploratory orders analogous to that in Ω^A, the additional gains are roughly 15% larger than the losses on exploratory orders. Whereas the extra component of the A-HFTs' performance directly explained using Ω^A represented less than 5% of A-HFTs' losses on exploratory orders, the analogous estimated performance increases more than offset the costs of exploration once we include the gains from avoiding unprofitable aggressive orders. In the case of the strategy that employs information from the A-HFTs' exploratory orders with the higher threshold, the estimated gains from exploration exceed the costs by more than one-third.

[10] The two strategies that incorporate exploratory information select subsets of the aggressive order placement times generated by the public-information-only strategy. Although the selected orders tend to be more profitable, they are also fewer in number.

Figure 7. Gains from A-HFT Exploratory Info Relative to Losses on Exploratory Orders



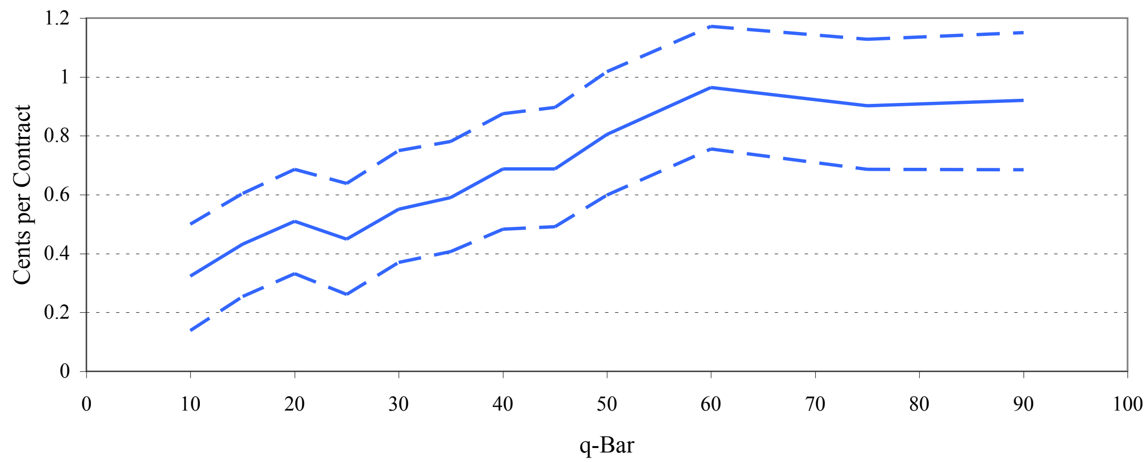
7 Discussion

7.1 Broader Opportunities for Exploratory Gains from Aggressive Orders

The empirical results in the preceding sections focused on the information generated by the A-HFTs' smallest aggressive orders. While their otherwise-perplexing unprofitability made these orders the most obvious starting point for an empirical study of exploratory trading, there is no theoretical reason why these small orders should be the sole source of exploratory information. In the baseline exploratory trading model, it was only to highlight the key aspects of the model that I assumed the HFT's period-1 order was expected to lose money and served no purpose other than exploration.

In principle, even aggressive orders that an A-HFT expects to be directly profitable could produce valuable, private, exploratory information. To investigate this possibility, I repeat the analysis of section 5.2.2 setting $\bar{q} = 25, 30, 35, 40, 45, 50, 60, 75, 90$. The A-HFTs' incremental aggressive orders included with each increase of $\bar{q}$ beyond $\bar{q} = 20$ are directly profitable on average, and yet the market response following these orders still provides significantly more additional explanatory power for the A-HFTs' performance on larger aggressive orders than it provides for that of other traders. Indeed, the additional explained components of the A-HFTs' performance are markedly larger than those for $\bar{q} = 1, 5, ..., 20$; see Figure 8, below, and see Appendix Table D.4 for numeric results.

Figure 8. [A-HFT Addt'l Explained] - [Everyone Else Addt'l Explained] (95% Conf. Bands)



7.2 Exploratory Trading and Speed

Further analysis of the exploratory trading model reveals natural connections between exploration and two important concepts of speed.

7.2.1 Low Latency

One common measure of trading speed is *latency*—the amount of time required for messages to pass back and forth between a trader and the market. While low-latency operation and high-frequency trading are not equivalent, minimal latency is nonetheless a hallmark of high-frequency traders. For a trader who can identify profitable trading opportunities, there is obvious value to possessing latency low enough to take advantage of these opportunities before they disappear. The new insight from the exploratory trading model concerns the more subtle matter of how low latency connects to the identification of such opportunities.

In the model of exploratory trading developed in section 2, the HFT's inference about Λ on the basis of market activity following his aggressive order in period 1 implicitly depends on a notion related to latency. If we suppose that random noise perturbs the orderbook, say according to a Poisson arrival process, then the amount of noise present in the HFT's observation of the market response in some interval following his aggressive order will depend on the duration of that interval. The duration of this interval will depend in large part upon the rate at which market data is collected and disseminated to the HFT, that is, the "temporal resolution" of the HFT's data. Although this temporal resolution does not directly depend on the HFT's latency, the HFT's latency *is* implicitly constrained by the temporal resolution of his market information.

The finer temporal resolution required for low-latency operation enables low-latency traders to obtain meaningful—and empirically valuable—information about the market activity immediately following their aggressive orders, and this information degrades at coarser temporal resolutions. The empirical results from section 5.3 provide a concrete illustration of this effect. The changes in resting inside depth immediately following an arbitrary aggressive order are less useful for forecasting price movements than are the analogous changes following an A-HFT's aggressive order, but the two can only be distinguished (by the A-HFT) in data with a sufficient level of temporal disaggregation.

7.2.2 High Frequency

Exploratory trading bears a natural relationship to the practice of placing large numbers of aggressive orders—what might be considered "high-frequency trading" in the most literal sense.

Exploratory information generated by a given aggressive order is only valuable to the extent that it can be used to improve subsequent trading performance. Because exploratory information remains relevant for only some finite period, the value of exploratory information diminishes as the average interval between a trader's orders lengthens. The exploratory trading model readily captures this effect if we relax the simplifying assumption that the liquidity state Λ remains the same between periods 1 and 2. Suppose that Λ evolves according to a Markov process, such that with probability τ, a second Λ is drawn in period 2 (from the same distribution as in period 1), and with probability $1 - \tau$, the original value from period 1 persists in period 2. Intuitively, τ parametrizes the length of period 1, and this length increases from zero to infinity as τ increases from zero to unity. As τ tends towards unity—i.e., as the length of period 1 increases to infinity—the liquidity state in period 1 becomes progressively less informative about the liquidity state in period 2.

As discussed in section 7.1, both theory and empirical evidence suggest that almost any aggressive order that a trader places generates some amount of exploratory information. Consequently, as a trader places aggressive orders in greater numbers, he will gain access to greater amounts of exploratory information. Furthermore, the average time interval between a trader's aggressive orders necessarily shrinks as the number of those orders grows, so the exploratory information produced by each order tends to become more valuable to the trader. These synergistic effects dramatically magnify the potential gains from exploratory information for traders who place large numbers of aggressive orders.

7.2.3 Latency Détente

There has been much speculation about HFTs engaging in an "arms race" for ever-faster processing and ever-lower latency. If high-frequency trading entailed nothing more than reacting to publicly observable trading opportunities before anyone else, HFTs would indeed face nearly unbounded incentives to be faster than their competitors. While reaction speed is certainly one dimension along which HFTs compete, the empirical evidence of exploratory trading suggests that the A-HFTs, at least, can also compete along another dimension—exploration. Since exploratory trading provides the A-HFTs with private information, a trader who uses only public information will not necessarily be able to dominate the A-HFTs, even if that trader is faster than every A-HFT. Similarly, an A-HFT could potentially compensate for having (slightly) slower reactions than the other A-HFTs by engaging in greater levels of exploration.

7.3 Beyond A-HFTs: Other HFTs and Other Markets

Exploratory trading is not universally relevant to all HFT activity in all markets. Equities markets, for instance, may not exhibit the predictability in demand that makes exploratory trading viable in the E-mini market, so HFTs in these markets might primarily concern themselves with obtaining superior forecasts of demand, or they might employ some completely different technique. However, exploratory trading in the E-mini market depends only on the market's structure and aggregate dynamics; it does not depend directly on any specific features of the E-mini contract. The prevalence of exploratory trading in other markets is ultimately an empirical matter, but markets similar to the E-mini in size and structure could easily support exploratory trading.

Even in the E-mini market, an important component of HFT activity lies outside the immediate province of the exploratory trading model. Nevertheless, the scope for exploratory trading extends well beyond the aggressive activity of A-HFTs considered thus far. Though I have focused on the A-HFTs up to this point, the B-HFTs could also reap exploratory rewards from their aggressive orders, as could potentially any trader with similar capabilities. The B-HFTs' overall performance on aggressive orders does not present the same ostensible affront to market efficiency as does that of the A-HFTs, but the B-HFTs' aggressive orders nonetheless outperform both those of non-HFTs, and the baseline econometric benchmark, by a wide margin. If inventory management or risk-control considerations force B-HFTs to place unprofitable aggressive orders, exploratory trading could help to explain how the B-HFTs mitigate the associated losses.

Alternatively, if nothing forces the B-HFTs to place aggressive orders, then the B-HFTs' consistent losses from aggressive trading are puzzling in their own right, much as the A-HFTs' losses on small aggressive orders were. Although the B-HFTs do not recoup their losses on other aggressive orders as do the A-HFTs, they make enough from their passive trading to earn positive profits overall. Passive trading strategies, just like aggressive ones, would benefit from the superior price forecasting that exploratory information makes possible, so exploratory trading could help to explain the activity of B-HFTs in this scenario as well.[11]

8 Conclusion

Empirical evidence strongly suggests that the concept of exploratory trading developed in this paper helps to explain the mechanism underlying certain HFTs' superior capacity to profitably anticipate price movements in the E-mini market. The exploratory trading model also illuminates the manner in which these HFTs benefit from low-latency capabilities and from their submission of large numbers of aggressive orders.

Exploratory trading is a form of costly information acquisition, albeit an unfamiliar one. HFTs who engage in exploratory trading are doing something more than merely reacting to public information sooner other market participants. This raises the possibility that HFTs, through exploratory trading, uniquely contribute to the process of efficient price discovery. However, exploratory trading differs from traditional costly information acquisition in several important respects. First, the information that exploratory trading generates does not relate directly to the traded asset's fundamental value, but rather pertains to unobservable aspects of market conditions that could eventually become public, *ex-post*, through ordinary market interactions. Also, because exploratory trading operates through the market mechanism itself, exploration exerts direct effects on the market, distinct from the subsequent effects of the information that it generates.

Finally, since HFTs appear to trade ahead of predictable demand innovations—albeit in a sophisticatedly selective manner—the research of De Long *et al.* (1990) potentially suggests that HFTs could have a destabilizing influence on prices if suitable positive-feedback mechanisms exist.

[11] Total trading profits from any transaction net to zero, so if a trader earns money on an aggressive order, his passive counter-party loses money. Since exploratory information is valuable to an aggressor, it follows immediately that it is also valuable to a passor.

Comprehensive analysis of the theoretical and empirical aspects of these myriad issues lies beyond the scope of this paper, but the theory and evidence presented herein provide a starting point from which to rigorously address the market-quality implications of high-frequency trading going forward.

References

[1] Matthew Baron, Jonathan Brogaard, and Andrei Kirilenko. The trading profits of high frequency traders. September 2012.

[2] Bruno Biais, Thierry Foucault, and Sophie Moinas. Equilibrium algorithmic trading. Working Paper, October 2010.

[3] Jonathan Brogaard, Terrence Hendershott, and Ryan Riordan. High frequency trading and price discovery. 2012.

[4] Giovanni Cespa and Thierry Foucault. Insiders-outsiders, transparency and the value of the ticker. Queen Mary University Dept. of Economics Working Paper No. 628, April 2008.

[5] J.Bradford de Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45(2):379–395, June 1990.

[6] Lawrence R. Glosten and Paul R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71–100, 1985.

[7] Sanford J. Grossman and Merton H. Miller. Liquidity and market structure. *The Journal of Finance*, 43(3), July 1988.

[8] Joel Hasbrouck and Gideon Saar. Technology and liquidity provision: The blurring of traditional definitions. *Journal of Financial Markets*, 12:143–172, 2009.

[9] Joel Hasbrouck and Gideon Saar. Low-latency trading. NYU Stern/ Cornell GSM Working Paper, May 2011.

[10] Terrence Hendershott, Charles M. Jones, and Albert J. Menkveld. Does algorithmic trading improve liquidity? *The Journal of Finance*, 66(1), February 2011.

[11] Terrence Hendershott and Ryan Riordan. Algorithmic trading and information. NET Institute Working Paper No. 09-08, September 2009.

[12] Robert Jarrow and Philip Protter. A dysfunctional role of high frequency trading in electronic markets. Johnson School Research Paper Series No. 08-2011, March 2011.

[13] Andrei Kirilenko, Mehrdad Samadi, Albert S. Kyle, and Tugkan Tuzun. The flash crash: The impact of high frequency trading on an electronic market. October 2010.

[14] Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, November 1985.

[15] J.Chris Leach and Ananth N Madhavan. Intertemporal price discovery by market makers: Active versus passive learning. *Journal of Financial Intermediation*, 2(2):Pages 207–235, June 1992.

[16] J.Chris Leach and Ananth N. Madhavan. Price experimentation and security market structure. *The Review of Financial Studies*, 6(2):pp. 375–404, 1993.

[17] U.S. SEC. Concept release on equity market structure, concept release no. 34-61358. FileNo. 17 CFR Part 242 [Release No. 34-61358; File No. S7-02-10] RIN 3235-AK47, January 2010.

[18] U.S. SEC. Findings regarding the market events of may 6, 2010, September 2010.

[19] Hans R. Stoll. Inferring the components of the bid-ask spread: Theory and empirical tests. *The Journal of Finance*, 44(1):115–134, March 1989.

A Exploratory Trading Model Details

A.1 Solving the Baseline Exploratory Trading Model

Let s_t denote the sign of q_t.

A.1.1 Solving the Model: Period 2

If $\varphi = 0$, the HFT's optimal choice is to not submit an aggressive order in period 2, or equivalently, to set $|q_2| = 0$. If $\varphi \neq 0$, then it is optimal for the HFT to set $s_2 = \varphi$ (unless the optimal $|q_2|$ is zero), so we only need to determine the optimal magnitude, $|q_2|$. Because π_2 is linear in $|q_2|$ when s_2 is held fixed, we can restrict attention to corner solutions (0 or N) for the optimal choice of $|q_2|$ without loss of generality. Note that if $q_2 = 0$, then $\pi_2 = 0$, regardless of the values of φ and Λ.

Suppose that the HFT sets $|q_2| = N$. Without loss of generality, assume that $s_2 = \varphi \neq 0$. The HFT's period-2 profits are given by

$$\tilde{\pi_2} = \begin{cases} N(1-\alpha) & if \ \Lambda = U \\ -N\alpha & if \ \Lambda = A \end{cases}$$

where the tilde on $\tilde{\pi_2}$ denotes the fact that the HFT's choice of q_2 does not condition on the value of Λ.

HFT Does Not Know Λ If the HFT does not know the value of Λ, then in the case where $\varphi \neq 0$, the HFT's expected period-2 profit if he sets $|q_2| = N$ is

$$\begin{aligned} \mathbb{E}\left[\tilde{\pi_2} | \varphi \neq 0, |q_2| = N\right] &= uN(1-\alpha) - (1-u)N\alpha \\ &= (u-\alpha)N \end{aligned}$$

Taking expectations with respect to φ, we find that the *ex-ante* expectation of $\tilde{\pi_2}$ when the HFT sets $|q_2| = N$ (and $s_2 = \varphi$) is given by

$$\mathbb{E}\left[\tilde{\pi_2} | \, |q_2| = N\right] = v(u-\alpha)N$$

When $u - \alpha < 0$, if the HFT did not know Λ, he would set $q_2 = 0$ rather than $|q_2| = N$. Hence the *ex-ante* expectation of $\tilde{\pi}_2$ is

$$\mathbb{E}[\tilde{\pi}_2] = \max\{v(u-\alpha)N, 0\}$$

HFT Knows Λ Next, if the HFT *does* know the value of Λ, then he will set $|q_2| = N$ (and $s_2 = \varphi$) only when $\Lambda = U$ and $\varphi \neq 0$. Denoting the HFT's period-2 profits from this strategy by $\hat{\pi}_2$, we find

$$\begin{aligned}
\mathbb{E}[\hat{\pi}_2|\varphi \neq 0] &= u(1-\alpha)N \\
&= (u-\alpha)N + \alpha(1-u)N \\
\mathbb{E}[\hat{\pi}_2] &= vu(1-\alpha)N \\
&= v(u-\alpha)N + v\alpha(1-u)N
\end{aligned}$$

Note that

$$\mathbb{E}[\hat{\pi}_2] > \max\{v(u-\alpha)N, 0\}$$

so the HFT's expected period-2 profits are strictly greater when he knows Λ than when he doesn't know Λ.

A.1.2 Solving the Model: Period 1

At the start of period 1, the HFT knows neither φ nor Λ, but he faces the same trading costs, α, as he does in period 2. Consequently, the HFT's expected direct trading profits from a period-1 aggressive order are negative:

$$\begin{aligned}
\mathbb{E}[\pi_1|q_1] &= \mathbb{E}[|q_1|(s_1 y - \alpha)|s_1, q_1] \\
&= |q_1| s_1 \mathbb{E}[y] - \alpha |q_1| \\
&= -\alpha |q_1|
\end{aligned}$$

The second equality relies on the assumptions that φ and Λ (and hence y) are independent of s_1 and q_1, while the final equality uses the fact that $\mathbb{E}[y] = 0$.

Since there is no noise in this baseline model, the HFT learns Λ perfectly from any aggressive order that he places in the first period with $|q_1| \geq 1$. An aggressive order of size greater than one yields no more information about Λ than a one-contract aggressive order in this setting, but the larger aggressive order incurs additional expected losses. Thus without loss of generality, we can restrict attention to the case of $q_1 = 0$ and the case of $|q_1| = 1$.

If the HFT sets $q_1 = 0$, he neither learns Λ nor incurs any direct losses in period 1, so his total expected profits are simply

$$\begin{aligned} \mathbb{E}\left[\pi_{total}|q_1 = 0\right] &= \mathbb{E}\left[\tilde{\pi_2}\right] \\ &= \max\left\{v\left(u-\alpha\right)N, 0\right\} \end{aligned}$$

Alternatively, if the HFT sets $|q_1| = 1$, his total expected profits are given by

$$\begin{aligned} \mathbb{E}\left[\pi_{total}|\,|q_1| = 1\right] &= -\alpha\,|q_1| + \mathbb{E}\left[\hat{\pi_2}\right] \\ &= vu\left(1-\alpha\right)N - \alpha \end{aligned}$$

A.1.3 Remark on the Sequence of Events

The central results of the model would not change if the HFT observed the signal of future aggressive order-flow *before* deciding whether to engage in exploratory trading, rather than observing it after deciding. However, the sequence of events outlined in section 2, in which the HFT must choose whether or not to explore before he observes φ, is more appropriate from an empirical perspective. For the HFT to learn about the liquidity state after he submits an aggressive order, he must wait for 1) his order to reach the market and execute, 2) information about that execution to reach other traders, 3) other traders to decide what to do, 4) other traders' decisions to reach the market, and 5) information about the market response to get back to him. Of these five steps, (1), (2), (4) and (5) each take approximately $3-4$ milliseconds, and (3) takes considerably longer, perhaps $3-20$ milliseconds, for an overall total of $15-40$ milliseconds. An HFT who has already done his exploration will be able to take advantage of predictable aggressive order-flow long before an HFT who only engages in exploratory trading after seeing an order-flow signal.

A.2 Solving the Model of Section 2.3

If the HFT places an order in the first period, it follows immediately from the baseline model results that his expected total profits are given by

$$\mathbb{E}\left[\pi_{total}\,|\,|q_1| = 1\right] \quad = \quad Nvu\,(1-\alpha) - \alpha$$

However, the HFT's expected profits if he does not place an order in period 1 are higher than in the baseline model, because the HFT now learns something from the depth changes following the other trader's aggressive order. If resting depth weakly replenishes after that order, the HFT learns with certainty that the liquidity state is accommodating (i.e., $\Lambda = A$), so the HFT will not submit an aggressive order in period 2, and his total profits will be zero. Alternatively, if resting depth further depletes following the aggressive order in period 1 (denote this event by g_1), we have

$$\begin{aligned}
& \mathbb{P}\left\{\Lambda = U | g_1\right\} \\
= \; & \frac{\mathbb{P}\left\{\Lambda = U,\ and\ g_1\right\}}{\mathbb{P}\left\{g_1\right\}} \\
= \; & \frac{\mathbb{P}\left\{g_1 | \Lambda = U\right\}\mathbb{P}\left\{\Lambda = U\right\}}{\mathbb{P}\left\{g_1 | \Lambda = U\right\}\mathbb{P}\left\{\Lambda = U\right\} + \mathbb{P}\left\{g_1 | \Lambda = A\right\}\mathbb{P}\left\{\Lambda = A\right\}} \\
= \; & \frac{1 * \mathbb{P}\left\{\Lambda = U\right\}}{1 * \mathbb{P}\left\{\Lambda = U\right\} + \rho * \mathbb{P}\left\{\Lambda = A\right\}} \\
= \; & \frac{u}{u + \rho\,(1-u)}
\end{aligned}$$

It follows immediately from the analogous result in the baseline model that the HFT's expected period-2 profits are given by

$$\mathbb{E}\left[\pi_2 | AO\ by\ someone\ else\right] = \max\left\{Nv\left(\frac{u}{u+\rho\,(1-u)} - \alpha\right), 0\right\}$$

Since the HFT does not place an order in the first period, his expected total profits equal his expected period-2 profits. Overall, then, the HFT's expected total profit if he observes an aggressive order placed by someone else in period 1 but does not place an aggressive order himself, is given by

$$\mathbb{E}\left[\pi_{total} | AO\ by\ someone\ else\right] = \max\left\{Nv\left(\frac{u}{u+\rho\,(1-u)} - \alpha\right), 0\right\}$$

B Measuring Aggressive Order Profitability

Calculating round-trip profits using a FIFO or LIFO approach is not a useful way to measure the profitability of individual aggressive orders. Even the most aggressive HFTs engage in some passive trading, so a FIFO/LIFO-round-trip measure would either confound aggressive trades with passive trades, or require some arbitrary assumption to distinguish between inventory acquired passively and inventory acquired aggressively (on top of the already-arbitrary assumption of FIFO or LIFO). A second, more general problem is that a measurement scheme based on inventory round-trips will always combine at least two orders (an entry and an exit), so such measurement schemes do not actually measure the profitability of *individual* aggressive orders.

In this appendix, I provide rigorous justification for the claim that the average expected profit from an aggressive order in the E-mini market equals the expected favorable price movement, minus trading/clearing fees and half the bid-ask spread. After presenting the formal proof, I discuss details of empirically estimating expected favorable price movement.

B.1 Preliminaries

Trading/clearing fees apply equally to both passively and aggressively traded E-mini contracts, so to simplify the exposition, I will initially ignore these fees. Similarly, I make the simplifying assumption that the bid-ask spread is constant, and identically equal to one tick; for the E-mini market, this assumption entails minimal loss of generality.

In the E-mini market, the profitability of individual aggressive orders can be considered in isolation from passive orders. Because E-mini contracts can be created directly by buyers and sellers, a trader's net inventory position does not constrain his ability to participate in a given trade[12]. As long as he can find a buyer, a trader who wishes to sell an E-mini contract can always do so, regardless of whether he has a preexisting long position. More generally, if a trader enters a position aggressively then exits it passively, he could have conducted the passive transaction even if he hadn't engaged in the preceding aggressive transaction. While a desire to dispose of passively-acquired inventory might *motivate* a trader to submit an aggressive order, the question of underlying motivation is distinct from the question of whether the aggressive order was directly profitable.

[12]The one exception would arise in the extremely rare event that a trader who did not qualify for a position-limit exemption held so many contracts (either long or short) that his inventory after the trade would exceed the position limit of 100,000 E-mini contracts. For HFTs, this minor exception can safely be ignored.

B.2 Formal Argument

With these preliminaries established, I turn to the rigorous argument. Consider a trader who executes J aggressive sell orders of size one, and J aggressive buy orders of size one, for some large J. Following the remarks above, the trader's passive transactions can be ignored. Let the average direction-normalized price change after these aggressive orders be $\tilde{\vartheta} \equiv \vartheta\left(\frac{2J}{2J-1}\right)$ ticks for some ϑ that does not depend on J.

First, suppose that the trader always submits an aggressive sell after an aggressive buy, and always submits an aggressive buy after an aggressive sell. Without loss of generality, assume that the trader's first aggressive order is a buy. The trader's combined profit from all $2J$ aggressive orders is

$$
\begin{aligned}
\pi_{2J} &= -a_1 + b_2 - a_3 + b_4 - \ldots - a_{2J-1} + b_{2J} \\
&= -a_1 + (a_2 - 1) - a_3 + (a_4 - 1) - \ldots - a_{2J-1} + (a_{2J} - 1) \\
&= -a_1 + a_2 - a_3 + a_4 - \ldots - a_{2J-1} + a_{2J} - J \\
&= -a_1 + (a_1 + \zeta_{b,1}) - (a_2 + \zeta_{s,2}) + (a_3 + \zeta_{b,2}) - \ldots \\
&\quad \ldots - (a_{2J-2} + \zeta_{s,J}) + (a_{2J-1} + \zeta_{b,J}) - J \\
&= \sum_{i=1}^{J} (a_{2i-1} + \zeta_{b,i}) - \sum_{j=2}^{J} (a_{2j-2} + \zeta_{s,j}) - a_1 - J \\
&= \sum_{i=1}^{J} a_{2i-1} - \left(a_1 + \sum_{j=1}^{J-1} a_{2j}\right) + \sum_{i=1}^{J} \zeta_{b,i} - \sum_{j=2}^{J} \zeta_{s,j} - J
\end{aligned}
$$

where a_k and b_k respectively denote the prevailing best ask and best bid at the time the kth aggressive order executes, $\zeta_{b,r}$ denotes the change in midpoint price following the rth aggressive buy order, and $\zeta_{s,r}$ denotes the change in midpoint price following the rth aggressive sell order. Note that $\vartheta \equiv \frac{1}{2J}\left(\sum_{r=1}^{J} \zeta_{b,r} + \sum_{r=1}^{J} (-\zeta_{s,r})\right)$.

Next, taking expectations, we find

$$
\begin{aligned}
\mathbb{E}\left[\pi_{2J}\right] &= \sum_{i=1}^{J}\mathbb{E}\left[a_{2i-1}\right] - \left(\mathbb{E}\left[a_1\right] + \sum_{j=1}^{J-1}\mathbb{E}\left[a_{2j}\right]\right) \\
&\quad + \sum_{i=1}^{J}\mathbb{E}\left[\zeta_{b,i}\right] - \sum_{j=2}^{J}\mathbb{E}\left[\zeta_{s,j}\right] - J \\
&= J\mathbb{E}\left[a_1\right] - \mathbb{E}\left[a_1\right] - (J-1)\,\mathbb{E}\left[a_1\right] + J\mathbb{E}\left[\tilde{\vartheta}\right] - (J-1)\left(-\mathbb{E}\left[\tilde{\vartheta}\right]\right) - J \\
&= (2J-1)\,\mathbb{E}\left[\tilde{\vartheta}\right] - J \\
&= (2J-1)\left(\mathbb{E}\left[\vartheta\right]\frac{2J}{2J-1}\right) - J \\
&= J\left(2\mathbb{E}\left[\vartheta\right] - 1\right)
\end{aligned}
$$

where the second equality uses the assumption that midpoint prices follow a martingale with respect to their natural filtration, together with the assumption of a constant bid-ask spread. From the final equality above, it follows immediately that the trader's average expected profit on an individual aggressive order is given by

$$
\frac{1}{2J}\mathbb{E}\left[\pi_{2J}\right] = \mathbb{E}\left[\vartheta\right] - \frac{1}{2}
$$

Finally, note that none of the calculations above relied on the assumption that the aggressive orders alternated between buys and sells (this only simplified the notation). It follows immediately from grouping together multiple aggressive orders of the same sign that the result would hold for orders of varying sizes, provided that the overall aggressive buy and aggressive sell volumes were equal.

Under the usual regularity conditions, as $J \to \infty$, $\tilde{\vartheta} \to_{A.S.} \lim_{J\to\infty}\mathbb{E}\left[\tilde{\vartheta}\right] = \mathbb{E}\left[\vartheta\right]$. □

B.2.1 Independent Importance of the Result

Establishing a meaningful technique to estimate the performance of individual aggressive orders was merely a necessary stepping stone for a detailed analysis of HFTs' performance on aggressive orders, but to the best of my knowledge, I am the first to propose and rigorously justify this technique. Anecdotal evidence from the CFTC suggests that the technique I develop in this paper may have broad applications for academics, regulators and practitioners alike.

B.3 Obtaining Unbiased Estimates

Recall that the discussion in section 4.1 implied that we can estimate the profitability of an HFT's aggressive order using the (direction-normalized) accumulated price-changes following that aggressive order out to some time past the HFT's maximum forecasting horizon. If we choose too short an accumulation window, the resulting estimates of the long-run direction-normalized average price changes following an HFT's aggressive orders will be biased downward. This enables us to empirically determine an adequate accumulation period by calculating cumulative direction-normalized price changes over longer and longer windows until their mean ceases to significantly increase

Market activity varies considerably in its intensity throughout a trading day, so event-time, which I measure in terms of aggressive order arrivals, provides a more uniform standard for temporal measurements than does clock-time. Empirically, an accumulation period of about 30 aggressive orders suffices to obtain unbiased estimates of the price movement following an HFT's aggressive order, but I consider results for an accumulation period of 50 aggressive orders to allow a wide margin for error. The mean direction-normalized price changes following individual HFTs' aggressive orders does not differ significantly for accumulation periods of 50, 200, or 500 aggressive orders, even if we distinguish aggressive orders by size. The same holds true for aggressive orders placed by non-HFTs. Using too long an accumulation period will not bias the estimates, but it will introduce unnecessary noise, so I opt for an accumulation period of 50 aggressive orders.

As I discuss at greater length in section 4.3, future price movements are moderately predictable from past aggressive order flow and orderbook activity, but only at very short horizons. Of the variables that meaningfully forecast future price changes, the direction of aggressive order flow is by far the most persistent, but even its forecasting power diminishes to nonexistence for price movements more than either about 12 aggressive orders or 200 milliseconds in the future. The adequacy of a 30+ aggressive order accumulation period is entirely consistent with these results.

As a simple empirical check on the validity of direction-normalized cumulative price changes as a proxy for the profitability of aggressive orders, I use each HFT's explicit overall profits and passive trading volume, together with the profits on aggressive orders as measured by the proxy, to back out the HFT's implicit profit on each passively traded contract. The resulting estimates of HFTs' respective profits from passive transactions are all plausible from a theoretical perspective, and are comparable to non-HFTs' implicit performance on passive trades.

C Benchmark Regression Results

In this appendix, I present and discuss results from regression (8) of section 4.3.

Recall that for each trading day in my sample, I regress the cumulative price-change (in dollars) between the aggressive orders k and $k+50$, denoted y_k, on the following variables: changes in resting depth between aggressive orders $k-1$ and k at each of the six price levels within two ticks of the best bid or best ask, the signs of aggressive orders $k-1$ through $k-4$, and the signed executed quantities of aggressive orders $k-1$ through $k-4$. For symmetry, I adopt the convention that sell depth is negative, and buy depth is positive, so that an increase in buy depth has the same sign as a decrease in sell depth. I estimate the equation

$$\begin{aligned} y_k &= z_{k-1}\Gamma + \epsilon_k \\ &:= \gamma_1 d^1_{k-1} + \ldots + \gamma_6 d^6_{k-1} + \\ &\quad \gamma_7 sign_{k-1} + \ldots + \gamma_{10} sign_{k-4} + \\ &\quad \gamma_{11} q_{k-1} + \ldots + \gamma_{14} q_{k-4} + \epsilon_k \end{aligned}$$

where d^r_{k-1} denotes the change in resting depth at price level r ($r = 3$ corresponds to the best bid, $r = 4$ corresponds to the best ask), $sign_l$ denotes the sign of aggressive order l, and q_l denotes the signed executed quantity of aggressive order l.

Table C.1, below, summarizes the estimates from the regression above, computed over my entire sample. All of the variables are antisymmetrical for buys and sells, and so have means extremely close to zero, but the mean magnitudes in the rightmost column of Table C.1 provide some context for scale.

Table C.1. Estimates from Benchmark Regression

	Coefficient $\times 10^3$	Robust t-Statistic	Variable Avg. Magnitude
$d_{k-1}^{best\ bid-2}$	-0.90	-1.02	4.13
$d_{k-1}^{best\ bid-1}$	-2.08	-4.29	10.8
$d_{k-1}^{best\ bid}$	1.13	4.94	23.1
$d_{k-1}^{best\ ask}$	1.11	4.97	23.4
$d_{k-1}^{best\ ask+1}$	-2.03	-4.24	11.2
$d_{k-1}^{best\ ask+2}$	-1.60	-1.90	4.44
$sign_{k-1}$	1186	33.3	1
$sign_{k-2}$	753	20.2	1
$sign_{k-3}$	544	14.6	1
$sign_{k-4}$	472	13.4	1
q_{k-1}	4.09	9.29	12.6
q_{k-2}	2.66	6.59	12.6
q_{k-3}	1.85	4.66	12.6
q_{k-4}	1.16	2.98	12.6

Comparable results obtain using as few as two lags of aggressive order sign and signed quantity. Linear forecasts of y_k do not benefit appreciably from the inclusion of data on aggressive orders before $k-4$, or on changes in resting depth prior to aggressive order $k-1$. Because the price-change y_k is not normalized by the sign of the kth aggressive order, it has an expected value of zero, so I do not include a constant term in the regression. Including a constant term in the regression has negligible effect on the results.

Although the last several aggressive order signs do offer rather remarkable explanatory power, the respective distributions of resting depth changes and executed aggressive order quantities have much heavier tails than the distribution of order sign, so price forecasts are substantially improved by the inclusion of these variables.

The positive coefficients on the lagged aggressive order variables and on the depth changes at the best bid and best ask are consistent with the general intuition that buy orders portend price increases, and sell orders portend price decreases. The negative coefficients on depth changes at the outside price levels require slightly more explanation.

Because the E-mini market operates according to strict price and time priority, a trader who seeks priority execution of his passive order will generally place that order at the best bid (or best ask); however, if the trader believes that an adverse price movement is imminent, he will place his order at the price level that he expects to be the best bid (ask) following the price change. It is relatively uncommon for prices to change immediately after an aggressive order in the E-mini market, but when prices do change, it is extremely rare during regular trading hours for the change to exceed one tick. As a result, the expected best bid (ask) following a price change is typically one tick away from the previous best, so it is not surprising that (e.g.) an increase in resting depth one tick below the best bid tends to precede a downward price change. These features of the E-mini market also shed some light on why changes in depth more than one tick away from the best (i.e., $d_{k-1}^{best\ bid-2}$ and $d_{k-1}^{best\ ask+2}$) are not significant predictors of future price movements.[13]

[13]Similarly, this line of reasoning helps to explain why changes in resting depth prior to aggressive order $k-1$ do not help to forecast price changes after the kth aggressive order.

D Supplemental Tables of Empirical Results

D.1 Performance of A-HFTs' Non-Rebalancing Aggressive Orders

Table D.1, below, summarizes the performance of aggressive orders that move the submitting account *away* from a neutral inventory position. A net position of zero is the most likely "neutral" inventory target, but I allow for the possibility that a given HFT has an arbitrary constant target inventory position, and I restrict attention to aggressive orders that move the HFT away from that target inventory level.

Table D.1. Performance of A-HFTs' Non-Rebalancing Aggressive Orders

(Dollars per Contract)

Cutoff	Below Cutoff 95% CI	Above Cutoff 95% CI	AOs Below Cutoff % of All AOs	AOs Below Cutoff % of Aggr. Volume
1	$(3.24, 3.42)$	$(6.90, 7.12)$	15.00%	0.25%
5	$(3.64, 3.79)$	$(6.95, 7.17)$	27.25%	0.89%
10	$(2.48, 2.63)$	$(7.13, 7.35)$	34.06%	1.94%
15	$(2.79, 2.96)$	$(7.13, 7.35)$	35.01%	2.14%
20	$(3.01, 3.19)$	$(7.21, 7.44)$	37.16%	2.81%

D.2 Additional Performance Explained by Exploratory Information

Table D.2, below, presents cross-sectional averages of the estimated additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9). The extra explanatory power of (10) reflects the contribution from the private component of information (available to the A-HFT under consideration) manifested in Ω^A. Numbers reported for the A-HFTs are averages over the estimates for individual A-HFTs. The membership of the "all others" depends upon the particular A-HFT being excluded, and the numbers reported for "All Others" are averages over these slightly different groups. Units are cents per contract, and confidence intervals are constructed by bootstrap.

Table D.2. Additional Explained Performance (Cents per Contract)

	Point Estimate	95% Confidence Interval
$\bar{q} = 1$		
A-HFTs	0.179	$(0.069, 0.296)$
All Others	0.034	$(0.018, 0.048)$
$\bar{q} = 5$		
A-HFTs	0.659	$(0.480, 0.850)$
All Others	0.082	$(0.055, 0.109)$
$\bar{q} = 10$		
A-HFTs	0.397	$(0.221, 0.570)$
All Others	0.074	$(0.041, 0.104)$
$\bar{q} = 15$		
A-HFTs	0.533	$(0.355, 0.705)$
All Others	0.101	$(0.071, 0.133)$
$\bar{q} = 20$		
A-HFTs	0.623	$(0.453, 0.799)$
All Others	0.113	$(0.077, 0.147)$

D.3 Differences in Additional Performance Explained by Exploratory Information

Table D.3, below, presents cross-sectional averages of the difference in mean additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9), between the indicated groups. Units are cents per contract, and confidence intervals are constructed by bootstrap.

Table D.3. Additional Explained Performance of A-HFTs vs. Everyone Else and vs. Other HFTs (Cents per Contract)

	Point Estimate	95% Confidence Interval
$\bar{q} = 1$		
A-HFTs vs. Everyone Else	0.145	$(0.036, 0.257)$
A-HFTs vs. Other HFTs	0.038	$(-0.076, 0.154)$
$\bar{q} = 5$		
A-HFTs vs. Everyone Else	0.577	$(0.393, 0.764)$
A-HFTs vs. Other HFTs	0.458	$(0.269, 0.651)$
$\bar{q} = 10$		
A-HFTs vs. Everyone Else	0.323	$(0.139, 0.500)$
A-HFTs vs. Other HFTs	0.235	$(0.053, 0.420)$
$\bar{q} = 15$		
A-HFTs vs. Everyone Else	0.432	$(0.254, 0.604)$
A-HFTs vs. Other HFTs	0.323	$(0.135, 0.503)$
$\bar{q} = 20$		
A-HFTs vs. Everyone Else	0.510	$(0.332, 0.686)$
A-HFTs vs. Other HFTs	0.426	$(0.238, 0.618)$

D.4 Results for Extended Values of $\bar{q}$

Table D.4, below, presents results for extended values of $\bar{q}$. Each large row corresponds to a single value of $\bar{q}$, and within a given row, the topmost two sub-rows present averages of the estimated additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9). Numbers reported for the A-HFTs are averages over the estimates for individual A-HFTs. The membership of the "all others" depends upon the particular A-HFT being excluded, and the numbers reported for "All Others" are averages over these slightly different groups. The bottommost two sub-rows for each value of $\bar{q}$ present cross-sectional averages of the difference in mean additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9), between the indicated groups. Units are cents per contract, and confidence intervals are constructed by bootstrap. The abbreviation "EE" stands for "Everyone Else."

Table D.4. Additional Explained Performance for Extended Values of $\bar{q}$

$\bar{q}$		Point Est.	95% CI	A-HFTs vs.	Point Est.	95% CI
25	A-HFTs	0.624	$(0.437, 0.806)$	EE	0.450	$(0.262, 0.639)$
	EE	0.175	$(0.135, 0.213)$	Other HFTs	0.360	$(0.171, 0.554)$
30	A-HFTs	0.697	$(0.520, 0.894)$	EE	0.551	$(0.370, 0.750)$
	EE	0.147	$(0.106, 0.182)$	Other HFTs	0.454	$(0.272, 0.660)$
35	A-HFTs	0.733	$(0.551, 0.920)$	EE	0.590	$(0.406, 0.781)$
	EE	0.143	$(0.106, 0.184)$	Other HFTs	0.489	$(0.299, 0.684)$
40	A-HFTs	0.850	$(0.646, 1.037)$	EE	0.688	$(0.483, 0.876)$
	EE	0.162	$(0.120, 0.205)$	Other HFTs	0.580	$(0.362, 0.779)$
45	A-HFTs	0.850	$(0.659, 1.053)$	EE	0.688	$(0.492, 0.896)$
	EE	0.162	$(0.119, 0.205)$	Other HFTs	0.560	$(0.354, 0.776)$
50	A-HFTs	1.003	$(0.810, 1.219)$	EE	0.804	$(0.599, 1.019)$
	EE	0.199	$(0.145, 0.253)$	Other HFTs	0.643	$(0.435, 0.871)$
60	A-HFTs	1.181	$(0.985, 1.381)$	EE	0.964	$(0.755, 1.172)$
	EE	0.218	$(0.162, 0.272)$	Other HFTs	0.786	$(0.562, 0.992)$
75	A-HFTs	1.073	$(0.860, 1.293)$	EE	0.902	$(0.687, 1.128)$
	EE	0.171	$(0.112, 0.230)$	Other HFTs	0.757	$(0.535, 1.003)$
90	A-HFTs	1.040	$(0.821, 1.242)$	EE	0.920	$(0.685, 1.151)$
	EE	0.120	$(0.053, 0.187)$	Other HFTs	0.746	$(0.498, 0.991)$

D.5 Aggressive Order Performance by Size and Trader Type

Table D.5. Size-Weighted Average Performance of Aggressive Orders Below and Between Size Thresholds (Dollars per Contract)

	A-HFTs		B-HFTs		Non-HFTs	
$\bar{q}$	All AOs $\leq \bar{q}$	Incremental AOs	All AOs $\leq \bar{q}$	Incremental AOs	All AOs $\leq \bar{q}$	Incremental AOs
1	3.84	-	4.37	-	1.68	-
5	4.23	4.38	4.56	4.64	1.72	1.75
10	3.49	2.84	4.66	4.85	1.77	1.89
15	3.85	6.39	4.71	4.95	1.79	1.97
20	4.14	4.95	4.77	5.08	1.83	2.06
25	4.41	6.65	4.81	5.14	1.86	2.20
30	4.79	6.79	4.87	5.24	1.90	2.53
35	4.99	7.00	4.88	5.10	1.91	2.40
40	5.28	6.69	4.91	5.14	1.95	2.55
45	5.42	7.04	4.92	5.49	1.95	2.14
50	5.61	6.90	4.95	5.14	1.96	2.00
60	5.87	7.00	4.98	5.30	1.99	2.87
75	6.12	7.20	5.00	5.40	2.03	2.71
90	6.38	7.20	5.01	5.52	2.03	2.16

D.6 Aggressive Order Size vs. Frequency and Volume, by Trader Type

Table D.6. Aggressive Orders Below Size Thresholds: Fractions of All Aggressive Orders and Aggressive Volume

	A-HFTs		B-HFTs		Non-HFTs	
$\bar{q}$	% of All AOs	% of Aggr. Volume	% of All AOs	% of Aggr. Volume	% of All AOs	% of Aggr. Volume
1	24.31%	0.40%	39.48%	4.77%	53.79%	5.68%
5	43.74%	1.44%	76.09%	16.13%	83.26%	14.88%
10	54.64%	3.09%	84.10%	23.95%	89.87%	20.68%
15	56.75%	3.54%	88.01%	30.03%	91.57%	23.05%
20	60.82%	4.80%	90.70%	35.82%	93.27%	26.43%
25	62.38%	5.37%	92.36%	40.29%	94.41%	29.32%
30	64.62%	6.39%	94.14%	46.30%	94.98%	31.07%
35	65.82%	7.02%	94.97%	49.56%	95.23%	31.94%
40	68.27%	8.51%	96.03%	54.31%	95.66%	33.69%
45	69.29%	9.20%	96.32%	55.76%	95.80%	34.33%
50	71.07%	10.55%	97.66%	63.29%	97.09%	41.08%
60	73.81%	12.97%	98.54%	69.12%	97.36%	42.66%
75	76.65%	16.01%	99.02%	73.14%	97.64%	44.76%
90	80.68%	21.11%	99.20%	74.86%	97.83%	46.37%