New Order Latency in the E-Mini Futures Market

Raymond P. H. Fishe
Richard Haynes
and
Esen Onur*

This version: November 2018

ABSTRACT

We estimate a proportional hazard model to study the latency between market exits and new orders. Although there is significant heterogeneity across participants, the delay to re-enter the market is a function of state covariates, such as manual- and algorithmic-entry characteristics, state conditions at an exit, such as inventory levels or whether the exit was by a market order, and importantly time-varying covariates, such as the volume of executions or cancellations during the latency gap. The results for the E-mini futures contract support the view that market participants place significant weight on publicly available signals to determine latency decisions.

Keywords: Algorithmic traders, Manual traders, Hazard rates, New Order Latency

JEL classification: G10, G13

***Fishe**: Patricia A. and George W. Wellde, Jr. Distinguished Professor of Finance, Department of Finance, Robins School of Business, University of Richmond, Richmond, VA 23173. Tel: (+1) 804-287-1269. Email: pfishe@richmond.edu. **Haynes**: U.S. Commodity Futures Trading Commission, Washington, D.C. 20581. Tel: (+1) 202-418-5000. Email: rhaynes@cftc.gov. **Onur**: U.S. Commodity Futures Trading Commission, Washington, D.C. 20581. Tel: (+1) 202-418-5000. Email: eonur@cftc.gov. Tel: (+1) 202-418-5000. We thank David Reiffen for comments on earlier versions of this research. The co-authors of this paper prepared this research under the direction of the Office of the Chief Economist at the CFTC, where economists produce original research on a broad range of topics relevant to the CFTC's mandate to regulate commodity future markets, commodity options markets, and the expanded mandate to regulate the swaps markets pursuant to the Dodd-Frank Wall Street Reform and Consumer Protection Act. This research was produced in each author's official capacity as a research economist or consultant with the Commission. The analyses and conclusions expressed in this paper are those of the author(s) and do not reflect the views of other Commission staff, the Office of the Chief Economist, or the Commission. All errors and omissions, if any, are the authors' own responsibility. Corresponding author: Raymond P. H. Fishe.

New Order Latency in the E-Mini Futures Market

1. Introduction

Market liquidity is the focus of much research in high frequency environments, particularly whether faster traders improve trading conditions and lower costs.[1] Most researchers find positive effects when faster traders participate in financial markets, such as narrower spreads and increased liquidity for smaller orders, although adverse selection costs may work against such results (Menkveld, Hendershott, and Jones, 2011; Foucault, Kozhan, and Tham, 2017; Menkveld and Zoican, 2017). A related question is how quickly participants act to provide additional liquidity; that is, how quickly depth on the order book recovers after an execution or cancellation removes some existing liquidity. Do faster traders speed up liquidity replenishment as well as overall trading speed? The evaporation of liquidity during the Flash crash in May 2010 and subsequent mini-crashes make this an important question to answer to understand the effects of fast trading environments.[2]

Fishe, Haynes, and Onur (2015) show that high frequency traders (HFTs) or algorithms are not equally fast, with speed varying significantly as measured by new order origination after an execution. Algorithmic traders are faster than manual-entry but within both groups there is high variation. Thus, individual traders decide how long to wait before they re-engage with the market, and may adjust such latency based on both market and individual metrics. Because traders may decide to speed up or delay order messages, latency is a decision variable and

[1] Menkveld (2016) and Fishe and Smith (2018) offer surveys of high frequency and algorithmic trading effects on financial and commodity markets, respectively.
[2] The causes of the Flash Crash of 2010 are examined by Aldrich, Grundfest, and Laughlin (2016), Kirilenko, et al. (2017), and Menkveld and Yueshen (2017). See also U.S. Commodity Futures Trading Commission (2018) for documentation of sharp price movements in futures markets.

liquidity a function of that decision for both algorithmic- and manual-entry traders. The goal of this paper is to identify the factors that affect new order latency decisions. We focus on factors that affect both new limit orders that add liquidity and new market orders that remove liquidity. We also examine how inventory levels affect these responses, and identify broader market characteristics that slow down or speed up a trader's response. These broader market characteristics are found to be the most important to explaining the timing of new liquidity decisions.

Earlier research on financial market latency examined execution speeds, that is, the time between trades (Engle and Russell, 1998; Engle, 2000). Angel, Harris, and Spatt (2015) find that execution latencies for securities decreased during 2001-2014, coincident with the growth of algorithmic trading. This finding suggests that the speed of liquidity provision has similarly increased to accommodate faster trade prints.

Additional latency research evaluated the timing of events after an order submission, such as the time-to-cancel, time-to-modify, or time-to-execution. An SEC (2013) study of quote lifetime distributions for corporate securities shows that 27% of quotes are executed within 500 milliseconds and 1.4% of quotes are executed within 100 microseconds, supporting the view of rapid liquidity removal. The SEC also reports that 39% (4.2%) of all quotes are canceled within 500 milliseconds (100 microseconds) thereby removing liquidity at a rapid rate. As such, order book replenishment rates must be even higher to support such removal rates.

The methods used in this study are similar to Hasbrouck and Saar (2013), who examine market latency rates conditional on quote change events, as these are events likely to trigger quick responses.[3] They estimate hazard rates after the submission of new, quote-improving orders. They find declining hazard rates for same-side submissions and for new executions at the improved quote. In each case, however, the hazard rate jumps initially suggesting that some set of participants respond within 2-3 milliseconds of a quote change, confirming that speedy traders offer and take liquidity with low latencies. Importantly, the subsequent declining hazard profiles suggest a mixture of latencies exist in the overall market. Our research provides estimates of how both trader and market characteristics affect these hazard rate profiles.

The approach here measures how long it takes participants' to respond after they exit the market—called the "latency gap". The exit could be due to an execution or an individual decision to cancel an order. Both of these actions remove liquidity from the order book. Importantly, both actions admit the possibility that the trader (or algorithm) may re-start their participation logic at this point. Thus, under certain circumstances we may be identifying the signals or factors that initiate participation in a market—that is, the trading logic. For example, if an execution brings a trader's inventory to zero, that trader is effectively out of the market, so we separately examine how quickly such traders return to the market and what factors affect the latency of this 'initiation' decision.

Using six days of order book data from the E-mini futures complex, we analyze 19,465 accounts, with approximately 14% operating as algorithmic-entry participants. Combined, these accounts executed 2,738,390 trades and cancelled 6,895,702 orders. Subsequent order submissions led to an analysis of 1,868,121 cases of new order latency after an execution-created exit, and 2,287,090 cases after a cancelation-created exit. Bootstrap methods are also used to

[3] Other studies that examine latency measures in financial markets include Riordan and Storkenmaier (2012), Kirilenko and Lamacie (2015), Menkveld and Zoican (2017), and Cartea, et al. (2018).

make inferences at the participant level, rather than the order message level. Because a small fraction of participants account for a large fraction of orders, the bootstrapping method equalizes the influence of all participants on our results.

Covariate estimates in a proportional hazard rate model provide a list of the factors affecting new order latency after either an execution or a cancellation message. Most of the covariates studied here are binary variables indicating the participant displays a particular feature, such as uses an algorithm to trade. These covariate effects work through estimated hazard rates. Technically, the hazard specifies the instantaneous "new order" rate at time *t* given that there is no "new order" between the initial time and *t*. The higher the "new order" hazard rate for a covariate, the more likely we are to observe a new order entry in the next instant of time for participants with that covariate's features relative to participants without those features. In this sense, higher hazard rates correspond to relatively faster actions to submit new order messages by participants with those covariate characteristics.

Several results arise based on the proportional hazard rate estimates. The state-level covariates at the account level affect order entry latency for both manual- and algorithmic-entry participants. Specifically, proprietary algorithmic accounts reveal higher hazards than customer-based accounts, accounts seeking to execute at the last execution price or last cancel price exhibit higher relative hazards particularly for market orders, market-order execution-exits decrease subsequent hazards, and executions that reach or cancellations that maintain a zero-inventory balance reduce these hazard rates.

These findings show that selected factors help explain the latency actions of participants. Although these results are new, they make sense in a microstructure setting. Proprietary accounts act based on the logic of the owner, but customer accounts often represent various owners, who may have very different goals, which we find makes them on average less likely or slower to provide or take liquidity. Participants who execute at the previous price must act quickly before the market moves to a new price, which also motivates participants who have cancelled at a given price. A market-order execution indicates a more immediate demand by a participant, which carries a higher transaction cost than limit orders on average, so these costs may slow subsequent actions. In effect, participants may be more circumspect about how quickly and at what price to re-enter the market.

The inventory effects tend to discriminate between liquidity providers and liquidity takers. Our results show that participants with zero inventory positions are more likely to take liquidity from the market compared to those who have added to inventory with a previous execution. The latter are more likely to submit a limit order when they re-enter the market. In addition, our results indicate that participants are likely to move more quickly towards a zero-inventory position than they are to move away from that inventory level.

A sub-sample that isolates those with zero-inventory positions helps to identify whether such cases meaningfully change how participants react to market signals. At zero inventory levels, a participant must decide to enter a buy or sell order without regard to inventory adjustment costs. This choice may cause a different reaction to the same type of signals when compared to participants with existing long or short positions. Our findings show that participants at a zero-inventory often exhibit greater hazard rates than our full sample results. For example, algorithmic proprietary participants are 1.6 times more likely to submit a market order to end a latency gap than manual customer participants, but in the zero-inventory sample they are 2.2 times more likely. In general though the covariates used for signals in the full sample data provide similar explanatory power and the same directional predictions in the zero-inventory sample. Even if the hazards of these covariates change, as a group they work well to explain participant decisions about new orders from both zero- and non-zero-inventory positions.

We also find a significant jump in explanatory power occurs when time-varying covariates that measure market events are included in the estimated models. These time-varying covariates suggest that participants are making new order latency decisions based largely on real-time information. After an execution, we find that an increase in either the volume executed or the volume of new orders in the latency gap increases the likelihood of subsequent new orders, while an increase in market-wide cancellations slows new submissions. The first two findings are consistent with participants recognizing greater execution opportunities and acting quickly before they dissipate or before the book queue greatly delays execution. The increase in market-wide cancellations is a less obvious signal. More cancellations may mean shorter wait times on the book, which would encourage order entry, but it may also mean that order flows are more informed or toxic, so participants are less inclined to enter new liquidity because of adverse selection costs. These results show that the latter effect dominates after an execution exit.

For cancellation exits the results are similar to execution exits, except that the market-wide execution covariate effect is reversed. Now, the greater are market-wide executions then the greater the delay in subsequent new orders. It would thus appear that the competing forces behind market-wide cancellations noted above are less competing when it is the participant who cancels. That is, if the participant cancels because of potential adverse selection costs, then subsequent market-wide executions may act to confirm those beliefs and thereby delay new order submission.

Lastly, we examine how hazard rates change when inferences are made at the participant level instead of the message or latency gap level. Specifically, these account-level data include multiple orders for the same participant. Those who are algorithmic, high frequency traders have many more orders than those who use manual-entry methods. Thus, when the proportional hazard regressions use the full sample, inferences are at the message level, which gives relatively greater weight to observations from a small set of active participants. To overcome this bias, we bootstrap these data to create 500 samples in which the new order latency is sampled once for each participant. We find that this approach generally alters the magnitude of the estimated hazard rates while retaining the previously noted directional effects.

The remainder of this paper proceeds as follows. Section 2 discusses the conceptual approach that underlies our methods. This approach relies on information signals to give participants incentives to bring new orders to the market. Section 3 discusses the order book data and provides summary statistics for our sample. Section 4 develops the empirical model and explains our results, and section 5 offers our conclusions.

2. Hazard Model: Signals and Order Latency

This analysis focuses on new order entry messages, which precede all other messages for an order. One may expect that both humans and algorithms rely on information signals to incite new order actions. For example, Hasbrouck and Saar (2013) posit that market participants will react after a new quote improves the best bid or offer in the market. They find that this common event results in a cascade of actions by participants, including entering new orders, cancelling existing orders, and trying to execute against the improved quote. We hypothesize that there are other such signals that cause participants to act. These signals may be grouped as external (e.g., earnings or news announcements), market related (e.g., a price or liquidity change), and/or specific to a participant (e.g., execution of or cancelling an outstanding order). This research focuses on market and participant signals because the set of external signals is too big to reasonably identify those effects and too difficult to assign timestamps for use with millisecond trade data.[4]

The delay in responding to signals makes it difficult to identify which signals matter to traders. The latency gap that results admits the possibility that many signals in the gap may be superfluous to a trader's order decision. In effect, a detailed flowchart of either the algorithm used for trading or specific instructions offered by human traders is required for exactitude. Thus, our analyses only 'approximates' the factors that affect new order latency because we do not know what combination of signals generate trader responses. We rely on signals that can be measured for each participant and those that arise during the new order latency gap.

In our context, a signal begins the new order latency gap and then the actual new order message ends the gap. Thus, there may be some ambiguity as to what signal defines the beginning of this time interval. We posit that new order latency may be analyzed by measuring how long a trader takes to act *after receiving an 'exit' signal from an existing order*. Specifically, both execution and cancellation confirmations represent exit signals in which a trader's order is removed from the matching engine. Our premise then is that these exit signals represent an approximate starting point for the trader's (or algorithm's) strategy. In the specific case when an exit signal leaves a trader with a zero inventory position, we argue that this is a natural point to re-start their trading strategy.[5] After an exit signal, the strategy then processes other signals until sufficient information is received that causes a new order message. From this view, the factors

[4] See Riordan, Storkenmaier, Wagener and Zhang (2013) for an approach to external news effects based on classifying newswire feeds into positive and negative sentiment.

[5] Another case that may admit a natural re-start of a strategy is when a participant cancels a limit order and then seeks to enter the order again with the same terms (e.g., price and quantity). However, because of anti-spoofing laws, such actions may lead to regulatory actions if repeated without intent to execute.

relevant to liquidity decisions are those observed at the time of the exit signal and those arising during the new order latency gap.

Insert Figure 1

Figure 1 illustrates new order latency based on the assumption that an execution signal starts the process that leads to a new order message. Two traders are shown, labeled “A” and “B”. Execution signals provided are marked by vertical dashed lines. These are public signals, except that the two traders know when their trades are confirmed. This confirmation indicates an exit from the market for their book or market orders. The new order latency times (t) of each trader are marked depending on when they re-enter the market with a new order. The shaded boxes capture the activity in the market during these latency gaps. We illustrate only the execution activity for simplicity and make no distinction between whether a limit or market order ends the gap. The empirical analyses accounts for differences in types of market-wide messages (i.e., order entry, cancellation, and execution) during the gap and whether limit or market orders arise after an exit.

Figure 1 shows that Trader B appears to act after receiving the next (public) trade execution signal from the market, which suggests a low threshold or simple rule for generating a new order for this trader. Alternatively, Trader B may be reacting quickly after the confirmation of a previous order, making this trader’s actions dependent on what we call ‘participant’ signals, such as whether the trader uses algorithmic- or manual-entry methods. Trader A appears less responsive to market execution signals, initiating new orders only after five and eight execution signals, respectively. In effect, Trader A’s new order decision may require greater confidence in market liquidity as shown by additional executions, or may be a function of other signals, which are not shown in this illustration.

Note that cancellation messages also create exit signals and represent a significant proportion of message traffic. Traders may use cancellations for many purposes, such as to remove stale limit orders, search for hidden liquidity when tied to more aggressive quoting, avoid interacting with informed traders, or to avoid being adversely selected in highly volatile markets. Unlike execution signals that confirm limit order matches, cancellations do not involve the actions of other traders. In effect, this exit message and the new order message are both created by a given trader's (or algorithm's) actions. Thus, new order latency derived from cancellations may show different effects at the trader level than the same latency using execution messages.

Figure 1 also suggests a modeling strategy for new order latency. Because we seek to explain how long it takes for a trader to submit a new order after an existing order has exited the market, the problem is analogous to those examined using survival analysis.[6] Cox (1992) developed a proportional risk model for survival analysis that has been shown to be flexible for many different latency applications. The Cox model assumes that covariates have a multiplicative effect on the hazard function.[7] This approach requires a specification of the hazard function (or rate). Specifically, if $t_{i,g}$ represents the time between order exit and order entry, where $g = 1,\ldots,G_i$ indexes the number of observable exit-to-reentry gaps for trader i, then a Cox-type hazard model with covariates may be specified as:

$$\lambda(t; x_{i,g}) = \lambda_0(t) exp(\boldsymbol{x}'_{i,g}\boldsymbol{\beta}), \tag{1}$$

[6] See Kalbfleisch and Prentice (2002) for a discussion of survival time models and examples.

[7] The hazard function specifies the instantaneous failure rate at time t given that there is no failure between the initial time and t. This function equals the density of failure time divided by the probability of survival beyond time t, the latter is known as the survivor function.

where $\lambda(t; x_{i,g})$ is the hazard function for the latency time between order exit and order entry, which depends on an arbitrary baseline hazard (λ_0), and covariates ($\boldsymbol{x}'_{i,g}\boldsymbol{\beta}$) that have a multiplicative effect via the exponential specification. The covariates examined here are variables that define the trader, such as manual or algorithmic, and variables that define information (or signals) arising during the gap between exit and reentry with a new order. Some covariates are time dependent, so we adjust our estimation methods to allow for such dependence.

The coefficient vector, $\boldsymbol{\beta}$, in the Cox model is estimated using partial information maximum likelihood methods. Because of proportionality, the baseline hazard is not involved with these estimates. We use the PHREG routine in SAS to estimate these coefficients.

The covariates we examine are state dependent, determined at the latency gap level, or those that would be known to the market or may be closely approximated during a latency gap between an exit signal and re-entry of a new order. Specifically, traders would know the volume of trading, and from updates to the book, they may determine flows on and off of the book. We measure these factors using the quantities adjusted by such messages. We consider such covariates as information signals received by market participants. Time dependencies arise because the longer the gap, the greater the number of messages, on average, within the gap.

3. Data

The data studied here are for the E-mini futures contract, which is typically the most actively traded U.S. futures contract based on daily volumes. We analyze order book data for participant accounts on six trading days, August 1-8, 2014. On these days there are a total of 1,868,121 execution messages and 2,287,090 cancellation messages that preceded a new order entry message as defined by our selection criteria (discussed below). These were generated by 19,465 accounts with varying levels of activity both within and between days.

As with any sample, several news events affected participants in equity and futures markets on these days. On Friday, August 1, there was a morning release by the Bureau of Labor Statistics of July employment data. Those data showed the U.S. added 209,000 non-farm payroll jobs and the unemployment rate increased to 6.2%.[8] Expectations were for 230,000 jobs and the unemployment rate unchanged at 6.1%. This release may have added some volatility to trading during this day as investors interpreted these data as reducing pressure on the Federal Reserve to raise interest rates. Overall, the week ended with poor equity performance as the S&P500 index decreased by 3%.

On Tuesday, August 5, the market received news that there was a significant buildup of Russian troops along the Ukraine border.[9] Around Noon (Eastern Time), Poland's Foreign Minister stated that Russia is poised to pressure or invade Ukraine, which negatively affected equity indices. After a see-saw week based largely on interest rate concerns and foreign news, equities rebounded in the afternoon of August 8th on news that Russia had ended its latest military exercises on the Ukraine border. Overall, these events likely added volatility to our sample relative to a randomly drawn six days during 2014, but volatility is often associated with higher volumes, so our sample may offer a larger, more comprehensive panel of participants to analyze order entry latency.

The sample data are collected as part of the regulatory oversight at the Commodity Futures Trading Commission. The sample includes only contracts traded for the September 2014

---

[8] See Myles Udland, "Jobs disappoint, unemployment rate rises to 6.2%," *Business Insider*, August 1, 2014.
[9] See http://www.valueline.com/Markets/Daily_Updates/Stock_Market_Today for equity market news and reactions during the week ending August 8, 2014.

expiration month, which was the active month at this time for the E-mini contract. For each participant, we created a message entry sequence identifying every message submitted or received in time order. We then identified all cases in which an execution message was followed by a new order entry with or without intervening cancellations or modifications, but no other executions. Next we identified all cases in which a cancellation message was following by a new order entry, with or without intervening execution messages, but no other cancellations. These gaps are the base file from which we examined new order latency observations.

Each of the identified latency gaps includes either the last execution or last cancellation prior to a new order. This approach reduces the sample size relative to the total number of executions or cancellations because many of these exits are contiguous for a participant in the message stream. The confirmation timestamp on this last message marks the beginning of signal-processing for that participant. That is, this is the time of exit from the market and by assumption the beginning of a new strategy sequence that processes signals before re-entry. Re-entry to the market occurs when a new order is submitted, and we use the time stamp for CME receipt of the new order to mark the re-entry time. The difference between the re-entry and exit times is the order entry latency or "latency gap", the dependent variable in our analyses.

The sampling process investigates latency gaps in 15-minute intervals beginning at 7:00AM CT and ending at 3:00PM CT. For a given 15 minutes, we looked forward another 10 minutes (25-minute window) to determine if the gap terminated with a new order entry. Gaps that terminated after the 25-minute window were excluded from the sample. Gaps that began in the 15-minute interval and terminated outside of the interval but before the end of the 25-minute window were included as censored observations in the proportional hazard regressions. This sample design eliminates the really long latency gaps that arise because some traders may wait hours before returning to the market, particularly those traders handling customer orders. The abundance of activity in these long-gap cases appears as noise to their decisions. The small fraction of censoring reported in our full sample results suggests that this design was effective in sampling the vast majority of latency gaps.

Table 1 provides summary statistics for these sample data and the latency gaps that we identified. This table reports both means and medians of sample covariates, which include the volume of cancellations, executions, and new order messages observed market-wide within a latency gap. These statistics are also computed for inventory at the beginning of a gap and the absolute value of beginning inventory. Additional statistics measure the percent of gaps that start with a zero inventory, use a market order to execute at the beginning of an execution-exit gap, and set the new order price equal to the price at execution or to the limit price on a cancellation message. These data include both participants who acted for their own accounts (proprietary traders) and participants who acted on behalf of customers. The fraction of proprietary traders is about 9% in the sample while the number of algorithmic traders is just over 14%.

Panel A in Table 1 reports the summary statistics for latency gaps started by an execution-created exit from the market. These data show an overall average latency gap of 40.5 seconds with algorithmic-entry cases averaging 16.9 seconds and manual-entry averaging 139.2 seconds before a new order message. The difference between algorithmic-entry and manual-entry is perhaps best seen with the median statistics. The medians show that one-half of the algorithmic-entry gaps were less than 210 milliseconds, while the midpoint for manual-entry gaps was 31.3 seconds. Not surprisingly, in the hazard rate analysis below, we find significant differences between algorithmic- and manual-entry covariates. These differences also appear in Panel B where we report statistics for the gaps in which a cancellation caused a market exit. The divergence between means and medians in both panels suggest significant skewness. The bootstrap procedure described below helps to overcome this issue.

The number of gaps analyzed is reported at the bottom of each panel. For algorithmic participants there are more cancel-to-new-order sequences for algorithmic-entry participants. Intuitively, this would imply that message origination latency for the cancel-to-new-order sequences is less than that of the execution-to-new-order sequence. The means and medians of the gap latency for algorithmic-entry participants confirm this observation, with the latency gap for cancellation exits being 68% smaller on average than the execution-exit gap latency. Interestingly, there are fewer cancellation-exit gaps than execution-exit gaps for manual-entry participants, but the latency is somewhat shorter after a cancellation-exit gap, suggesting that the decision to reenter for manual-entry traders is different than for algorithmic-entry traders.

Table 1 also shows summary statistics for covariates used in the hazard rate analysis. These covariate data may be grouped into state variables over the sample, state variables for a gap, and time varying variables. State variables for the sample are whether the account is algorithmic- or manual-entry, and if proprietary or customer based. State variables over a gap include whether the last execution was a market order, whether the new order equals the last execution price or the last cancellation price, the quantity traded or cancelled at the last execution or cancellation, how current inventory is affected by the execution or cancellation, and whether the new order is a limit or market order. The order type for the new order is partitioned in the analyses so that we may compare the latencies of liquidity providers (limit orders) and liquidity takers (market orders). Time varying variables include quantities executed, cancelled, and submitted during the gap for all other accounts in the market. We also used counts of these variables, but the results were substantially the same, so we focus on quantities.[10] These quantity variables are data that participants could observe or compute from market data feeds provided by the CME Group.

The participant count data in Table 1 show that there are 2,752 algorithmic-entry and 17,138 manual-entry participants over these six days of data. Of these 37.5% (5.2%) of algorithms (manuals) are proprietary traders. These percentages indicate that the vast majority of manual-entry actions are on behalf of customer accounts, consistent with about 50% beginning an execution-exit gap using a market order to exit the market. In comparison, only about 32% of algorithmic-entry participants use a market order to exit.

The inventory data in Table 1 reveal that a disproportionate percentage (31.6%) of cancellation gaps begin with a zero inventory compared to execution gaps (12.3%). These are state variables defined for each gap. In addition to the zero inventory measure, they also include whether an execution adds to (or subtracts from) a current long or short inventory position. For cancellation exits these measures calculate whether the cancelled order "would have" added to or subtracted from the existing inventory position. These percentages sum to 100% when the zero inventory measure is included. Thus in the hazard rate regressions, the zero inventory variable is omitted, so the hazard rate effects for inventory adjustments are relative to the zero inventory hazard rate. Note that these data show that about 64% of the execution exits begin after the participant receives a confirmation that adds to the existing inventory. The majority of these latency gaps are then for participants who are building positions. In contrast, the cancellation-exit gaps are almost equally distributed between building and reducing inventories should the cancelled order execute.

[10] Results for quantities and counts are very similar because the median order size is one contract in the E-mini futures market.

The market-wide cancel, execution, and new order quantity variables show smaller averages and medians for algorithmic- than manual-entry participants. This is consistent with the smaller average gap latency for algorithms. Finally, average order sizes for both executions and cancellations tend to be similar for algorithmic- and manual-entry participants, except the medians are two contracts for algorithms versus one contract for manuals in the cancellation-exit gaps.

In addition to the proportional hazard model, we institute a bootstrap procedure to equalize the effects that arise because some participants are more active than others. For example, Table 1 shows that algorithmic-entry participants average 548 execution-exit gaps while manual-entry participants average only 21 gaps. An even larger difference in these averages arises for cancellation-exit gaps. Thus, in a proportional hazard regression with all sample gaps, the behavior of algorithmic participants will receive greater weight in coefficient estimates. This is fine for inferences about the average latency gap, but not okay if we want to know about hazard rates for the average participant. To examine the effects at the participant level, we randomly select one latency gap for each participant forming a participant-level sample. The proportional hazard regression is then estimated across this sample, and the process is repeated for 500 random samples. The average coefficients across these samples show the hazard rate at the participant level.

4. Empirical Analysis

We begin by providing a more general analysis of how long participants in these markets wait before re-entering an order after an exit message signal. The distribution that describes this behavior is known as a survival curve. Figure 2 shows survival curves for participants grouped by order entry (manual or algorithmic) and customer type (proprietary or customer-initiated order). These data are for latency gaps measured from an execution message to a new order by these participants for the first day of our sample, August, 1, 2014. The upper panel shows survival curves for *all* observations while the lower panel bootstraps the data into 500 equal-weighted samples such that each participant is observed once in every sample. The bootstrap method shows how survival rates differ between the participant and message-gap levels.

The solid line crossing both graphs in Figure 2 highlights the 50% cutoff for each group. This line identifies the latency gap (horizontal axis) when one-half of observed new orders have been submitted after receiving a prior execution message. Both panels show that the 50% latency cutoff implies that algorithmic-entry traders are quicker to respond versus manual-entry traders after an execution signal. The 'all' observation curves in the upper panel show algorithmic proprietary traders are the quickest group with 50% of the observations responding with new orders in less than 200 milliseconds. In contrast it takes manual proprietary traders over 9 seconds on average for one-half of the observations to respond.

The lower panel in Figure 2 emphasizes the weighting problem created when all observations are collected into a single sample. The survival curves in this panel are for the average participant. There are meaningful differences between the survival curves shown here and those in the upper panel. Moreover, the groups used here (manual versus algorithmic) offer only limited controls for this heterogeneity. This point is clear from comparing the intersections of the 50% cutoff line in the lower panel to those in the upper panel. The cutoff line does not cross either manual-entry survival curve in the lower panel, so the cutoffs for those participant groups exceed the 15-second boundary on the horizontal axis. More significantly, the new cutoff for the algorithmic proprietary group increases latency time by a factor of 25 to just over 5 seconds, and by a factor of 4 to nearly 14 seconds for the customer-based algorithmic group. These results reveal how conclusions about market behavior may be meaningfully affected by disproportionate activity levels across participants. In other words, inferences drawn from all observations will be about the average observation, not the average participant.

4.1 *Execution and Cancellation Exits*

Tables 2 and 3 contain the estimates of the proportional hazard model and also the estimates using the bootstrap method. Table 2 reports results for a latency model of the time between order execution and new order entry. Table 3 shows results for the latency between order cancellation and new order entry. The results are grouped using two model specifications—one without (Model I and III) and one with (Model II and Model IV) time-varying variables. The bootstrap results are shown in Models III and IV. Each model shows estimated hazard rates and p-values for covariates in two columns. The first column shows the effects based on the new order being a limit order (i.e., liquidity providing), and the second column shows the effect when the new order is a market order (i.e., liquidity demanding). Both sets of covariates are included in the same estimated model. The bootstrap results show the average hazard rate across the 500 randomly drawn samples, with 95% confidence intervals shown below these averages. Note that the number of observations/participants, percent of censoring in the sample, and Generalized R-squareds are shown at the bottom of each model.

Because the proportional hazard model is exponential, the hazard rates are calculated as $e^{\hat{\beta}}$, where $\hat{\beta}$ is the estimated coefficient in these regressions. The multiplicative nature of these models implies that we interpret hazard rates as relative ratios for dummy variables. For example, Model I in Table 2 shows the hazard is 1.089 for manual-entry proprietary accounts who submit limit orders to end a gap. The omitted category here is manual-entry customer accounts. This estimate means that the hazard of submitting a new order for manual-entry *proprietary* accounts is 1.089 times the hazard for manual-entry *customer* accounts, which makes new orders from manual-entry proprietary accounts 8.9% more likely in the next instant of time. Similarly the hazard rate of 0.884 for the "Last Execution is a Market Order" covariate implies that those who submit limit orders to end an execution-exit gap are 11.6% less likely than those whose last execution was a limit order to end the gap.[11] With continuous variables, such as any of the quantity covariates, subtracting one from the hazard ratio and multiplying by 100 gives the percentage change in the hazard for each one-unit increase in the covariate. Our interpretation of these effects uses a one unit change in the covariate for individual quantities, but for market level quantities we examine the unit change in 1,000s.

The interpretation of hazard ratios allows us to compare covariate effects to an omitted group as well as make comparisons across covariates who condition on the same omitted group. There are two covariate groups with the same omitted variable in these hazard rate tables. The first contains sample-state covariates. These are manual proprietary, algorithmic proprietary and algorithmic customer, with the omitted group being manual customer accounts. The second group contains gap-state covariates for inventory effects. These are "last execution adds to long position," "last execution adds to short position," "last execution reduces a long position" and "last execution reduces a short positon," with the omitted variable being a zero-inventory position. For Table 3, the inventory group is the same but it is measured by what the cancelled order would have done to inventory.

[11] The multiplicative effect implies equal hazards when is the hazard ratio is 1. When it is less than one as in this example, then the effect is a fraction of the omitted factor. That is, 88.4% of the hazard when last execution was a limit order to end the gap, or equivalently interpreted as (1-0.884)*100% = 11.6% less likely.

Several results stand out for the full sample in Table 2. The hazard ratio for algorithmic proprietary accounts in Model I is 3.14 times the omitted covariate hazard (a manual customer account) and is the second largest hazard ratio in this table. An effect of this magnitude also arises for this covariate in Table 3 when the latency gap begins with a cancelled order. This hazard implies that algorithmic proprietary accounts are 1.49 (1.58) times more likely to end an execution-exit (cancel-exit) gap in the next instant of time if the new order is a limit order versus a market order. It appears that returning to the book with a new limit order after an execution (cancellation) takes time precedence over creating an automatic execution using a market order.

An exception to this limit order preference is the case when the new order is at the same price as the last execution. The participant here appears willing to chase that price as the hazard for a market order is 3.19 times the hazard when the price for the new order is different than the last execution price. In contrast, when a new limit order is submitted at the last execution price the hazard is only 1.58 times this hazard at a different price. These relatively high hazard effects based on the last execution price may arise because some participants are working a larger order. They may thus be trying to continue executing the remainder of the order at a favorable price. Note that this strong market-order effect also arises in Table 3 when the new order is at the last cancel price. In this case, the participant may or may not be working a larger order, but has decided that immediate execution is preferred. Also, these market-order hazard effects are smaller when time-varying covariates are introduced in Model II for Tables 2 and 3, but they still maintain their high relative magnitudes.

A different type of behavior arises when the execution that starts the latency gap is a market order. Here the hazard rates in Table 2 are all less than one, so this event slows down the likelihood of a new order in the next instant of time compared to an execution of a limit order starting the gap. Specifically, in Model I a new limit order is 11.6% less likely (as noted above) and a new market order is 16.8% less likely to end the gap compared to when the last execution is a limit order. In effect if participants provide liquidity by executing from the book they are more likely to continue to participate in the next instant of time than if they take liquidity with a market order. As market orders bear a cost of one-half the effective spread, it stands to reason that participants may re-think their actions if they have just incurred this cost relative to the case when they have implicitly received one-half the effective spread by executing from the book.

In Tables 2 and 3, inventory effects are measured relative to the case in which participants hold a zero inventory after an execution or before the cancellation of an order, respectively. The hazard rates after an execution-exit indicate that new limit orders from a non-zero inventory position are more likely to arise than those from a zero inventory position. Market orders from non-zero inventory positions are found less likely to arise than such orders from a zero inventory position when these orders add to a position, but are equally or slightly more likely when they reduce positions towards a zero inventory. That is, those with zero inventories are more likely to take additional liquidity from the market than those with non-zero inventories after an inventory-increasing execution. In total, the implication is that participants with a non-zero inventory are more likely to enter new limit orders versus new market orders for position adjustments. Importantly, in Model's I and II these hazard adjustments are uniformly greater when the participant's action seeks to move inventory *towards* a zero balance compared to accumulating a larger (long or short) inventory. To the extent that these instantaneous probabilities translate into speed measures, participants can be said to move more quickly towards a zero inventory than they move away from this level.

The same models in Table 3 show a different response strategy when a cancelled order would have affected an existing non-zero inventory position. For each inventory covariate, the hazard rate response is nearly the same between a new limit or market order, and greater than one. Thus, the likelihood of a new limit or market order is higher from a non-zero than a zero-inventory position after a cancellation. All together these inventory results suggest that exits which create or leave a zero inventory position make participant's less likely to supply new liquidity, and in the case of execution-exits more likely to demand new liquidity.

Model II in Tables 2 & 3 includes time-varying covariates in the regressions. These are based on measured quantities. The "quantity executed" and quantity cancelled" covariates use the quantity for the individual account, while the other quantity measures represent the volume of executions, new orders, and cancellations by the entire market (excluding this account) in the latency gap. These market-wide variables are measured in 1,000s of contracts. Subtracting one from each estimated hazard and multiplying by 100 provides the percentage change in the hazard for a one-unit change in the covariate. The results in both tables for the individual quantity covariate suggest a negligible or zero effect for changing the initial order size by one contract.

In contrast, more activity in the market has a measurable effect on these hazards. The generalized R-squared increases by 46% and 69% from Model I to Model II in Tables 2 & 3, respectively.[12] Note that we observe similar R-squared increases in all tables when we add these time-varying covariates to the model. To compute the effects for these covariates, we evaluate the hazard response using a 10% change in the mean market volume as reported in Table 1. For example, a 10% change in average execution volume is 511 contracts, which is 0.511 units of the

[12] The generalized R-squared is defined as $R^2 = 1 - exp[-\chi^2/N]$, where $\chi^2$is the chi-squared statistic for a test of whether the covariates have a zero coefficient and N is the sample size in the proportional hazard regression. Magee (1990) recommends this measure as helpful for interpreting likelihood-based inference measures.

“quantity executed during the gap” covariate. This translates to a 1.0% increase in the hazard of a new limit order (0.551*(1.018-1)*100) and a 0.1% increase in the hazard of a new market order. For the “new orders in the market” covariate, a 10% change in volume after an execution translates into 1.445 units of the covariate. This change implies a 2.6% increase in the hazard of a new limit order and a 3.3% increase in the hazard of a new market order. Similarly a 10% change in volume for the “market cancellations after an execution” covariate translates into 0.972 units of the covariate. This change implies a 6.7% decrease in the hazard of a new limit order and a 5.4% decrease in the hazard of a new market order. Thus for the same percentage change in covariate volume, there are much greater effects on execution-gap latency with higher cancel volume and higher new order volume than higher execution volume.

One interpretation of these results is that higher cancel volume implies that the rest of the market is removing liquidity, which makes an individual account pause before going against the tide to add new liquidity or to execute at current prices. For changes in market-wide new order volume, an individual account may see new orders as competitors, which creates an incentive to act as the queue in the order book lengthens. The slightly greater likelihood of a new market order message is consistent with this view. The hazard effects for market-wide executions are small, but they suggest that greater realized liquidity has a positive effect on the likelihood of a new limit order after an execution.

The effects for these time-varying covariates after a cancellation are notably different than after an execution. Following the procedures above, the quantity effects in Table 3 show that a 10% change in market-wide execution volume results in a 2.1% decrease in the hazard of a new limit order and a 2.4% decrease in the hazard of a new market order after a cancellation exit. A 10% change in new order volume results in a 5.6% increase in the hazard of a new limit order and a 6.9% increase in the hazard of a new market order. Lastly, a 10% change in cancellation volume results in a 2.6% decrease in the hazard of a new limit order and a 2.1% decrease in the hazard of a new market order.

The logic behind these cancel-exit results appears to be that more cancels by others in the market act to confirm the rationale behind the individual's cancel decision, possibly that there is now more risk to keeping an order on the book. If new order volume increases after a cancel, then concern about risk on the book is reduced, so participants are more likely to participate with either a new limit or market order.[13] These two effects mirror what is found after an execution-exit. A difference arises for execution volume during a latency gap created by a cancellation. Here the increase in market execution volume suggests that there is greater liquidity during the gap, but it appears that participants see such executions as a missed opportunity or as confirmation of adverse selection risk instead of confirmation that they may execute an order fairly quickly. We conclude that conditioning on execution-exits provides evidence of a participant's willingness to execute an order, so more observed liquidity encourages a new order response. However, conditioning on a cancellation, provides an ambiguous signal on the willingness of a participant to execute an order. Thus, participants may delay new order entry even if market liquidity increases after a cancellation.

4.2 Bootstrap Results for Execution and Cancellation Exits

Tables 2 and 3 also report average hazard rates after bootstrapping these data to give equal weight to each participant. Below each average hazard rate are the 95% confidence intervals

[13] This argument may differ by strategy. A market-making program may be more likely to enter new orders after cancelling as it tries to restore two-sided quotes, whereas directional or more aggressive accounts may wait longer.

obtained across the 500 bootstrap samples. Models III and IV replicate the covariates included in models I and II, respectively, for these tables. For the most part, the estimated hazards confirm the results found when using all of the sample data. However, it is also clear that the magnitude of these hazard estimates change at the participant level.

Specifically, the hazards for algorithmic proprietary accounts are of a smaller magnitude and no longer suggest sizable differences between limit and market order effects. In fact, for execution exits, the hazard of a new limit order is now less than the hazard of a new market order. Focusing on Model IV shows that the effects of submitting a new limit order do not now affect the covariates for when the “last execution is a market order” or “new order is at the execution price.” This insignificance also holds for cancel-exits when the covariate is “new order at the last cancel price.” In these cases, only the submission of a market order generates a hazard effect in the same direction as noted for the full sample.

The value of these bootstrap estimates is to show that the hazard rate effects for an average participant are reduced compared to those using all observations. Thus, the effects of certain high-activity participants tend to overstate the likelihood of new order activity (i.e., liquidity providing or liquidity taking) by the average participant.

4.3 Zero-Inventory Exits

The finding that exits which create or leave a zero inventory position make participants less likely to supply new liquidity suggests that participants with a zero-inventory position may be waiting for more signals before re-entering the market with new orders. This issue is investigated in Tables 4 and 5 for execution and cancellation exits, respectively. The estimated hazard rates in these tables characterize the actions of only those participants with zero inventory positions.

Thus, the finding in Table 4 that algorithmic-proprietary accounts are 5.02 times more likely to respond with new limit orders after an execution than manual customer-based accounts is only applicable to the population of accounts with zero inventories.

The main reason for examining these zero-inventory level data is that this is a natural position in which both algorithmic- and human-based strategies would restart their logic. It may then be expected that some of these covariates would have a different effect on the likelihood of new order entry. Comparing the covariates in models II and IV between Tables 2 and 4 finds only two cases in which the estimated hazard reverses its effect. In model II, the time varying covariate, “quantity executed during the gap,” has an increasing effect on both limit and market order entry, but once a zero-inventory position is reached there is a decreasing effect, meaning that the higher the level of execution volume in the gap the lower the likelihood of new order entry. This effect appears analogous to what is observed for this variable using cancel-exits. A zero-inventory position may thus give rise to the same market-based considerations as a cancel exit. The second case is that the hazard for manual proprietary accounts show an increasing likelihood of submitting a new limit order compared to manual customer accounts in the bootstrap results. In Table 2, this state-level covariate shows a decreasing likelihood of a new limit order. We offer no obvious reason for this finding.

There is only one meaningful hazard rate change for the zero-inventory cases after a cancellation exit. Model II in Table 5 shows that when the new order is at the last cancel price the likelihood of a new limit order is reduced relative to when the new order is not equal to the last cancel price. The impact here is fairly large as those who submit limit orders to end a gap are 30% more likely when the new order is at the last cancel price versus not as shown in Table 3.

Table 5 shows that this impact is now 14.7% less likely when starting at a zero inventory position.

Except for these selective reversals, what is interesting about the comparison between the zero-inventory results and the full sample findings is that the magnitudes of the hazard effects are increased for most of the estimates in the zero-inventory tables. For example, the zero-inventory bootstrap results in Table 4 indicate that the hazard for an algorithmic proprietary participant who submits a limit order to end the gap is 2.07 times the hazard of a manual customer participant. Table 2 shows this ratio to be 1.39 times. Similarly the hazard of a market order ending the gap in Table 2 when the last execution is a market order is 81% of the hazard when the last execution is not a market order. This percent changes to 63% in the bootstrap results in Table 4. Several of the cancel-exit hazard estimates also show meaningful changes in the magnitudes of these relative effects.

In summary, there are few reversals of hazard effects when using only zero-inventory data, and the generalized R-squared remains high in the time-varying models. The zero-inventory position may be said to speed up or add delays to the average responses found in the full sample results. These findings suggest that the covariates included in these models send fairly consistent signals to participants even when they fully removed from the market.

5. Conclusions

We have examined latency, defined by how long it takes a participant to place a new order after receiving an exit signal from the market, either an execution or cancellation exit. Our goal was to describe the covariates that affected this latency decision with the view that these covariates are signals that lead to participant actions. We identified several covariates that were significant in a proportional hazards model. These covariates distinguished hazard rate effects between algorithmic and manual-entry participants, and isolated the effects of inventories on new order latency. The most important of these covariates appear to be those whose signals are generated from market-wide data observed during the period of latency. Such market-based signals include the volume of executions, cancellations and new orders during the latency period. These covariates have meaningfully large explanatory power when included in the hazard rate models.

References

Aldrich, Eric M., Joseph A. Grundfest, and Gregory Laughlin, 2016, The Flash Crash: A New Deconstruction, Working paper, University of California-Santa Cruz.

Angel, James J., Harris, Lawrence E., and Chester S. Spatt, 2015, Equity trading in the 21st century: An update, *Quarterly Journal of Finance* 5(1), 1-39.

Aitken, M.J., Cummings, D.J. and F. Zhan, 2014, Trade size, high frequency trading, and colocation around the world, *European Journal of Finance*, 1-21.

Boehmer, Ekkehart, Kinglsey Fong, and Julie Wu, 2012, International evidence on algorithmic trading, Working Paper, EDHEC Business School.

Cartea, A., Payne, R., Penalva, J., and Mikel Tapia, 2018, Ultra-fast activity and intraday market quality, forthcoming in *Journal of Banking & Finance*.

Engle, Robert, 2000, The Econometrics of Ultra-High-Frequency Data, *Econometrica*, 68, 1-22.

Engle, Robert, and J. R. Russell 1998, Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, *Econometrica*, 66, 1127-1162.

Fishe, Raymond P. H., Onur, Esen, and Richard Haynes, 2015, Speed and Latency in Treasury and E-mini Futures Contracts – Part 2, White Paper, Office of the Chief Economist, Commodity Futures Trading Commission.

Foucault, Thierry, Kozhan, R., and W. W. Tham, 2017, Toxic arbitrage, *The Review of Financial Studies* 30(4), 1053–1094.

Kalbfleish, John D. and Ross L. Prentice, 2002, *The Statistical Analysis of Failure Time Data*, New Jersey: John Wiley & Sons.

Kirilenko, A. and G. Lamacie, 2015, Latency and Asset Prices, working paper, MIT Sloan School of Management.

Kirilenko, A., Kyle, A., Samadi, M., and T. Tuzun, 2017, The flash crash: High frequency trading in an electronic market, *Journal of Finance* 72(3), 967–998.

Magee, L., 1990, R^2 Measures Based on Wald and Likelihood Ratio Joint Significance Tests, *The American Statistician* 44, 250-253.

Menkveld, Albert J., 2016, The economics of high-frequency trading: Taking stock, *Annual Review of Financial Economics* 8, 1-24.

Menkveld, Albert J., Hendershott, T., and Charles M. Jones, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1-33.

Menkveld, Albert J., and Bart Z. Yueshen, 2017, The Flash Crash: A cautionary tale about highly fragmented markets, forthcoming in *Management Science*.

Menkveld, Albert J., and Marius A. Zoican, 2017, Need for speed? Exchange latency and liquidity, *The Review of Financial Studies* 30(4), 1188–1228.

Patterson, Scott, Strasburg. Jenny, and Liam Pleven, 2013, High-Speed Traders Exploit Loophole, *The Wall Street Journal*, May 1, 2013.

Riordan Ryan, and Andreas Storkenmaier, 2012, Latency, Liquidity, and Price Discovery, *Journal of Financial Markets* 15, 416-37.

Riordan Ryan, Storkenmaier, A., Wagener, M., and S. Zhang, 2013, Public information arrival: Price discovery and liquidity in electronic limit order markets, *Journal of Banking and Finance* 37(4), 1148-1159.

U.S. Commodity Futures Trading Commission, 2018, Sharp price movements in commodity futures markets, Staff Report—Market Intelligence Branch, Division of Market Oversight.

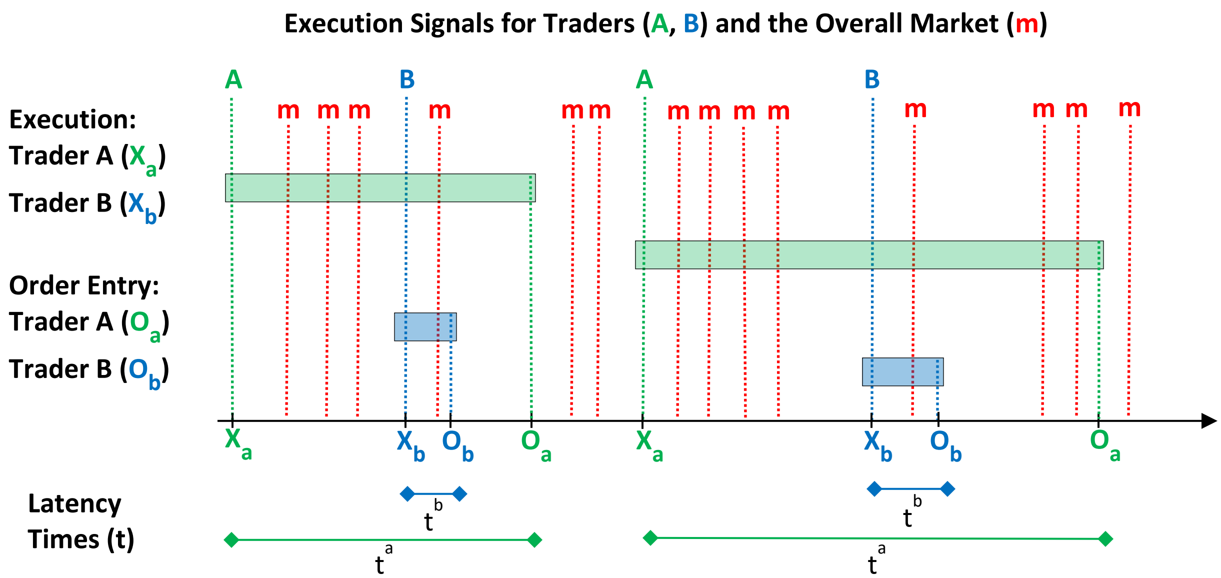

Figure 1: New Order Latency

This figure illustrates new order latency arising after an execution signal. Two traders (A and B) are shown. Execution signals for each trader are marked by vertical dashed lines as belonging to these traders (X) or other traders in the market (m). The new order entry times (t) of each trader are marked on the time lines below the figure and depend on when they re-enter a new order (O). The activity arising in the market during the latency gap for each trader is captured by the shaded boxes.

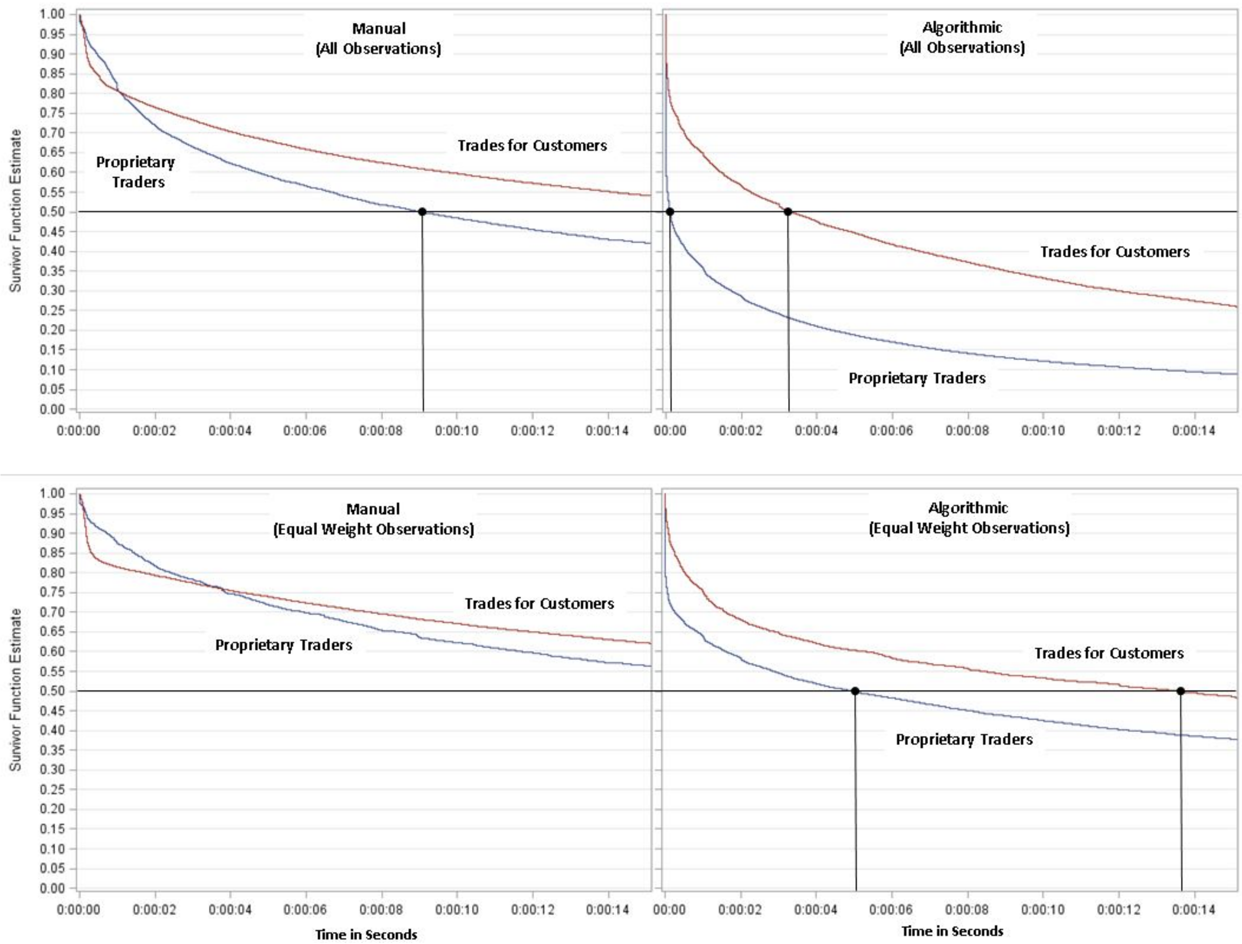

Figure 2: Survival Curves for E-mini Futures: Execution-to-New-Order Latency Gaps

This figure shows survival curves for new order latency defined after an execution removed liquidity from the order book. These data are for the first day of the sample, August 1, 2014. The top two panels show survival curves using all latency gap observations separated by whether the execution is for a proprietary or customer account. The left panel is for manual-entry participants and the right panel is for algorithmic-entry participants. The lower two panels show the same data after using the bootstrap method to generate 500 equally-weighted samples. The average gap latencies for each participant are computed and shown in the bottom panels.

Table 1
Summary Statistics for "New Order Latency" Events

Summary data are presented for the September 2014 expiration of the E-mini futures contract using a sample of six days from August 1-8, 2014. The mean and median of sample covariates are shown. These statistics are measured in the time gap between the last execution and the next new order entry in Panel A and the last cancellation and the next new order entry in Panel B. The time gap is expected to approximate new order latency--the time elapsed until sufficient signals arise to generate a new order. These data include both participants who acted for their own accounts (proprietary traders) and participants who acted on behalf of customers. The fraction of proprietary traders is show for each grouping. The summaries are for all gaps in the sample, and for algorithmic and manual-entry accounts, separately. Covariates shown are cancel, execution, and new order volumes within the gap (measured market wide), inventory at the beginning of a gap, and the absolute value of beginning inventory. Additional statistics measure the percent of gaps that start with a zero inventory and how that inventory changes with the execution (or would without the cancellation), use a market order to execute at the beginning of execution gaps, and set the new order price equal to the price at execution or the limit price on the cancellation message. The number of gaps and participants counts are also shown.

	E-Mini Futures Contract					
	Full Sample		Algo-Entry		Manual-Entry	
Signal Variable	Mean	Median	Mean	Median	Mean	Median
Panel A: New Order Latency Measured from Last Execution to Next New Order Entry						
Gap Latency (seconds)	40.50	0.77	16.87	0.21	139.20	31.30
Begin Gap Inventory	-0.70	0.00	-0.74	0.00	-0.57	0.00
Abs(Inventory)	46.10	5.00	52.40	7.00	19.68	1.00
Execution Order Size	3.34	1.00	3.29	1.00	3.52	1.00
Begin Gap with Zero Inventory	12.3%		10.6%		19.5%	
Execution adds to Long Inventory	32.1%		31.4%		35.1%	
Execution adds to Short Inventory	31.7%		30.6%		36.4%	
Execution reduces Long Inventory	11.7%		13.5%		4.3%	
Execution reduces Short Inventory	12.1%		13.9%		4.7%	
Market Order at Execution Exit	35.2%		32.0%		49.0%	
New Order at Execution Price	46.7%		55.3%		10.5%	
Proprietary Traders	9.2%		37.5%		5.2%	
In Gap:						
Cancel Volume	9,728	728	4,196	493	12,298	3,283
Execution Volume	5,111	349	2,168	238	13,135	3,329
New Order Volume	14,445	1,032	6,183	701	18,394	4,844
Participant Count	19,465		2,752		17,138	
Gap Count	1,868,121		1,508,183		359,938	
Panel B: New Order Latency Measured from Last Cancellation to Next New Order Entry						
Gap Latency (seconds)	16.17	0.55	5.34	0.45	123.67	23.55
Begin Gap Inventory	0.22	0.00	0.25	0.00	-0.09	0.00
Abs(Inventory)	26.56	5.00	28.38	3.00	8.49	0.00
Cancellation Order Size	6.13	2.00	5.99	2.00	7.46	1.00
Begin Gap with Zero Inventory	31.6%		28.8%		58.4%	
Cancel would add to Long Inventory	16.6%		17.4%		9.0%	
Cancel would add to Short Inventory	17.6%		18.5%		9.0%	
Cancel would reduce Long Inventory	17.0%		17.6%		11.4%	
Cancel would reduce Short Inventory	17.2%		17.7%		12.2%	
New Order at Cancel Price	54.3%		58.6%		11.1%	
Proprietary Traders	8.9%		43.4%		4.4%	
In Gap:						
Cancel Volume	3,666	270	1,201	235	26,643	6,877
Execution Volume	1,209	72	365	61	9,083	1,818
New Order Volume	5,457	436	1,799	386	39,561	10,108
Participant Count	14,189		1,883		12,536	
GapCount	2,287,090		2,077,747		209,343	

Table 2
Latency between Order Execution and New Order Entry

Proportional hazard regression estimates are shown for the latency between order execution and new order entry in the E-Mini futures contract. Models are estimated for two cases of population inference: Full sample of orders for order inference and a bootstrap simulation for inferences at the participant level. The bootstrap simulation includes 500 random samples in which a gap for each participant is drawn once in each sample. The simulation gives equal weight to all participants and removes effects caused when a few participants have many gaps. Interaction terms are included to measure how each covariate is affected if the new order is a market order. These effects are shown in the column labeled "Market-order" effects for ease of exposition. Covariates are static and time-varying. Static covariates are dummies for the following: manual participants who are proprietary, algorithmic participants who are proprietary, algorithmic participants handling customer orders, new order equals the last execution price, last execution is a market order, last execution adds to existing long inventory, last execution adds to existing short inventory, last execution subtracts from existing long inventory, and last execution subtracts from existing short inventory. The omitted variables that compare to these dummies are a manual participant who is acting for a customer and the case where inventory reaches zero at the last execution. The quantity of contracts traded at the last execution is also included as a static variable. Time-varying covariates are totals for execution quantity, new order quantity, and cancellation quantity measured during the gap. The model is estimated using a partial likelihood function that takes account of time dependent covariates. The table shows the estimated hazard rate for each covariate and the p-value of the covariate's estimated coefficient. The generalized pseudo r-squared, percentage of censored data, and number of observations/participants are shown at the bottom of the table for each model. For bootstrap results, the 95% confidence interval (95% C.I.) of the average hazard ratio and the generalized r-squared is shown.

	Full Sample Hazard ratio/p-value				Bootstrap Sample Hazard ratio/95% C.I.			
	Model I		Model II		Model III		Model IV	
Covariate	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects
Manual Proprietary Account	1.089 <0.001	1.142 <0.001	0.990 0.179	1.050 <0.001	0.861 (0.81, 0.91)	1.233 (1.14, 1.33)	0.889 (0.83, 0.96)	1.182 (1.08, 1.29)
Algorithmic Proprietary Account	3.139 <0.001	2.101 <0.001	2.095 <0.001	1.574 <0.001	1.550 (1.46, 1.65)	1.833 (1.66, 2.03)	1.385 (1.28, 1.49)	1.527 (1.34, 1.70)
Algorithmic Customer Account	1.416 <0.001	1.539 <0.001	1.157 <0.001	1.352 <0.001	1.171 (1.12, 1.22)	1.247 (1.17, 1.33)	1.147 (1.09, 1.21)	1.244 (1.15, 1.35)
New order at last execution price	1.579 <0.001	3.187 <0.001	1.520 <0.001	2.483 <0.001	1.209 (1.15, 1.27)	2.390 (2.22, 2.56)	1.034 (0.98, 1.09)	2.065 (1.91, 2.23)
Last execution is a market order	0.884 <0.001	0.832 <0.001	0.920 <0.001	0.778 <0.001	0.995 (0.96, 1.02)	0.879 (0.84, 0.91)	0.968 (0.94, 1.00)	0.809 (0.77, 0.85)
Last execution adds to long position	1.281 <0.001	0.920 <0.001	1.263 <0.001	0.875 <0.001	1.691 (1.63, 1.75)	0.798 (0,76, 0.84)	1.513 (1.45, 1.58)	0.691 (0,65, 0.73)
Last execution adds to short position	1.301 <0.001	0.933 <0.001	1.286 <0.001	0.883 <0.001	1.670 (1.61, 1.73)	0.846 (0.81, 0.89)	1.541 (148, 1.61)	0.744 (0.70, 0.79)
Last execution reduces long position	1.604 <0.001	1.120 <0.001	1.490 <0.001	0.998 <0.001	1.922 (1.68, 2.18)	1.392 (1.18, 1.62)	1.513 (1.30, 1.74)	1.034 (0.84, 1.24)
Last execution reduces short position	1.617 <0.001	1.139 <0.001	1.502 <0.001	1.006 <0.001	1.998 (1.74, 2.27)	1.412 (1.19, 1.66)	1.588 (1.37, 1.83)	0.989 (0.80, 1.20)
Quantity traded in last execution	0.997 <0.001	0.999 <0.001	0.998 <0.001	0.999 0.001	0.997 (0.99. 1.00)	1.000 (0.99. 1.00)	0.998 (0.99. 1.00)	1.000 (0.99. 1.02)
Time-varying covariates								
Quantity executed during the gap			1.018 <0.001	1.002 <0.001			0.966 (0.96, 0.97)	1.013 (1.01, 1.02)
New order quantity entered during the gap			1.018 <0.001	1.023 0.001			0.932 (0.92, 0.94)	0.932 (0.93, 0.94)
Quantity cancelled during the gap			0.929 <0.001	0.944 <0.001			0.997 (0.99, 1.01)	1.008 (1.00, 1.01)
Generalized R-Sqrd		32.0%		46.9%		10.8%		37.1%
Percent Censored		4.3%		5.0%		18.9%		18.4%
Observations/Accounts		1,868,121		1,613,993		19,465		18,594

Table 3
Latency between Order Cancellation and New Order Entry

Proportional hazard regression estimates are shown for the latency of new order entry after the participant cancels an order for the E-Mini futures contract. Models are estimated for two cases of population inference: Full sample of orders for order inference and a bootstrap simulation for inferences at the participant level. The bootstrap simulation includes 500 random samples in which a gap for each participant is drawn once in each sample. The simulation gives equal weight to all participants and removes effects caused when a few participants have many gaps. Interaction terms are included to measure how each covariate is affected if the new order is a market order. These effects are shown in the column labeled "Market-order" effects for ease of exposition. Covariates are static and time-varying. Static covariates are dummies for the following: manual participants who are proprietary, algorithmic participants who are proprietary, algorithmic participants handling customer orders, new order price set at the cancellation price, cancellation adds to existing long inventory, cancellation adds to existing short inventory, cancellation subtracts from existing long inventory, and cancellation subtracts from existing short inventory. The omitted variables that compare to these dummies are a manual participant who is acting for a customer and the case where inventory is zero before the cancellation. The quantity cancelled is also included as a static variable. Time-varying covariates are market totals for execution quantity, new order quantity, and cancellation quantity measured during the latency gap. The table shows the estimated hazard rate for each covariate and the p-value of the covariate's estimated coefficient. The generalized pseudo r-squared, percentage of censored data, and number of observations/participants are shown at the bottom of the table for each model. For bootstrap results, the 95% confidence interval (95% C.I.) of the average hazard ratio and the generalized r-squared is shown.

	Full Sample Hazard ratio/p-value				Bootstrap Sample Hazard ratio/95% C.I.			
	Model I		Model II		Model III		Model IV	
Covariate	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects
Manual Proprietary Account	1.079 <0.001	0.871 <0.001	0.978 0.179	0.770 <0.001	1.042 (0.97, 1.12)	0.933 (0.84, 1.04)	1.051 (0.96, 1.14)	0.805 (0.69, 0.92)
Algorithmic Proprietary Account	3.967 <0.001	2.508 <0.001	2.488 <0.001	1.776 <0.001	1.915 (1.81, 2.02)	1.482 (1.31, 1.68)	1.600 (1.45, 1.75)	1.365 (1.14, 1.59)
Algorithmic Customer Account	2.282 <0.001	1.365 <0.001	1.631 <0.001	1.186 <0.001	1.155 (1.09, 1.21)	1.477 (1.36, 1.59)	1.151 (1.07, 1.23)	1.413 (1.22, 1.60)
New order at last cancel price	1.338 <0.001	1.783 <0.001	1.300 <0.001	1.714 <0.001	1.105 (1.04, 1.17)	1.557 (1.41, 1.71)	1.013 (0.95, 1.08)	1.306 (1.16, 1.45)
Cancelled Order adds to long position	1.156 <0.001	1.154 0.885	1.149 <0.001	1.060 <0.001	0.988 (0.94, 1.04)	0.855 (0.80, 0.91)	1.002 (0.93, 1.09)	0.901 (0.81, 0.99)
Cancelled Order adds to short position	1.167 <0.001	1.164 0.843	1.161 <0.001	1.061 <0.001	1.082 (1.03, 1.14)	0.865 (0,80, 0.92)	1.080 (1.01, 1.15)	0.835 (0.76, 0.93)
Cancelled order reduces long position	1.192 <0.001	1.239 <0.001	1.173 <0.001	1.101 <0.001	1.736 (1.65, 1.84)	1.660 (1.53, 1.79)	1.375 (1.27, 1.47)	1.510 (1.32, 1.71)
Cancelled order reduces short position	1.204 <0.001	1.258 <0.001	1.186 <0.001	1.110 <0.001	1.688 (1.61, 1.78)	1.633 (1.52, 1.75)	1.342 (1.26, 1.43)	1.200 (1.08, 1.32)
Quantity cancelled	0.998 <0.001	0.999 <0.001	0.998 <0.001	0.999 0.001	1.000 (0.99, 1.00)	1.000 (1.00, 1.00)	1.000 (1.00, 1.00)	1.000 (0.99, 1.01)
Time-varying marketwide covariates								
Quantity executed during the latency gap			0.827 <0.001	0.802 <0.001			0.781 (0.78, 0.79)	0.779 (0.77, 0.79)
New order quantity during the latency gap			1.102 <0.001	1.127 0.001			1.138 (1.13, 1.14)	1.139 (1.13, 1.15)
Quantity cancelled during the latency gap			0.875 <0.001	0.864 <0.001			0.867 (0.86, 0.87)	0.867 (0.86, 0.87)
Generalized R-Sqrd		20.0%		33.9%		5.5%		58.2%
Percent Censored		1.7%		2.0%		16.1%		17.5%
Observations/Accounts		2,287,090		1,979,870		14,189		12,969

Table 4
Latency between Order Execution and New Order Entry Starting from Zero Inventory

Proportional hazard regression estimates are shown for the latency between order execution and new order entry in the E-Mini futures contract. Models are estimated for two cases of population inference: Zero inventory sample for order inference conditional on a zero inventory at the start of a gap and a bootstrap simulation for inferences at the participant level. The bootstrap simulation includes 500 random samples in which a gap for each participant is drawn once in each sample. The simulation gives equal weight to all participants and removes effects caused when a few participants have many gaps. Interaction terms are included to measure how each covariate is affected if the new order is a market order. These effects are shown in the column labeled "Market-order" effects for ease of exposition. Covariates are static and time-varying. Static covariates are dummies for the following: manual participants who are proprietary, algorithmic participants who are proprietary, algorithmic participants handling customer orders, new order equals the last execution price, and last execution is a market order. The omitted variables that compare to these dummies are a manual participant who is acting for a customer. The quantity of contracts traded at the last execution is also included as a static variable. Time-varying covariates are totals for execution quantity, new order quantity, and cancellation quantity measured during the gap. The model is estimated using a partial likelihood function that takes account of time dependent covariates. The table shows the estimated hazard rate for each covariate and the p-value of the covariate's estimated coefficient. The generalized pseudo r-squared, percentage of censored data, and number of observations/participants are shown at the bottom of the table for each model. For bootstrap results, the 95% confidence interval (95% C.I.) of the average hazard ratio and the generalized r-squared is shown.

	Zero Inventory Sample Hazard ratio/p-value				Bootstrap Sample Average Hazard ratio/95% C.I.			
	Model I		Model II		Model III		Model IV	
Covariate	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects
Manual Proprietary Account	1.393 <0.001	0.988 <0.001	1.267 <0.001	1.036 <0.001	1.261 (1.16, 1.37)	0.962 (0.84, 1.07)	1.121 (1.02, 1.22)	0.990 (0.86, 1.12)
Algorithmic Proprietary Account	5.020 <0.001	3.000 <0.001	3.214 <0.001	2.166 <0.001	2.582 (2.42, 2.74)	1.900 (1.68, 2.14)	2.070 (1.92, 2.23)	1.623 (1.38, 1.86)
Algorithmic Customer Account	1.544 <0.001	0.962 <0.001	1.386 <0.001	0.950 <0.001	1.370 (1.29, 1.45)	1.042 (0.95, 1.14)	1.323 (1.24, 1.41)	1.036 (0.93, 1.15)
New order at last execution price	1.801 <0.001	2.842 <0.001	1.742 <0.001	2.268 <0.001	1.521 (1.42, 1.62)	2.174 (1.93, 2.43)	1.228 (1.14, 1.31)	1.857 (1.61, 2.11)
Last execution is a market order	0.906 <0.001	0.835 <0.001	0.921 <0.001	0.807 <0.001	0.962 (0.93, 0.99)	0.717 (0.69, 0.74)	0.878 (0.84, 0.91)	0.629 (0.60, 0.66)
Quantity traded in last execution	0.997 <0.001	0.994 <0.001	0.999 0.057	0.993 <0.001	0.999 (0.99, 1.00)	0.995 (0.99, 0.99)	1.001 (0.99, 1.00)	0.994 (0.99, 1.00)
Time-varying covariates								
Quantity executed during the gap			0.989 <0.001	0.983 0.045			0.970 (0.96, 0.97)	1.009 (1.00, 1.01)
New order quantity entered during the gap			1.034 <0.001	1.035 0.667			0.946 (0.94, 0.95)	0.936 (0.93, 0.94)
Quantity cancelled during the gap			0.938 <0.001	0.942 0.230			1.007 (0.99, 1.02)	1.004 (1.00, 1.01)
Generalized R-Sqrd		43.7%		53.5%		6.4%		27.5%
Percent Censored		8.2%		8.9%		25.8%		26.1%
Observations/Accounts		230,019		207,043		10,877		10,561

Table 5
Latency between Order Cancellation and New Order Entry Starting from Zero Inventory

Proportional hazard regression estimates are shown for the latency of new order entry after the participant cancels an order for the E-Mini futures contract. Models are estimated for two cases of population inference: Zero inventory sample for order inference conditional on a zero inventory at the start of a gap and a bootstrap simulation for inferences at the participant level. The bootstrap simulation includes 500 random samples in which a gap for each participant is drawn once in each sample. The simulation gives equal weight to all participants and removes effects caused when a few participants have many gaps. Interaction terms are included to measure how each covariate is affected if the new order is a market order. These effects are shown in the column labeled "Market-order" effects for ease of exposition. Covariates are static and time-varying. Static covariates are dummies for the following: manual participants who are proprietary, algorithmic participants who are proprietary, algorithmic participants handling customer orders, and new order equals the last cancel price. The omitted variables that compare to these dummies are a manual participant who is acting for a customer. The quantity of contracts at the last cancel message is also included as a static variable. Time-varying covariates are totals for execution quantity, new order quantity, and cancellation quantity measured during the gap. The model is estimated using a partial likelihood function that takes account of time dependent covariates. The table shows the estimated hazard rate for each covariate and the p-value of the covariate's estimated coefficient. The generalized pseudo r-squared, percentage of censored data, and number of observations/participants are shown at the bottom of the table for each model. For bootstrap results, the 95% confidence interval (95% C.I.) of the average hazard ratio and the generalized r-squared is shown.

	Zero Inventory Sample Hazard ratio/p-value				Bootstrap Sample Hazard ratio/95% C.I.			
	Model I		Model II		Model III		Model IV	
Covariate	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects	Limit Order Effects	Market Order Effects
Manual Proprietary Account	1.190 <0.001	0.888 <0.001	1.125 <0.001	0.746 <0.001	1.130 (1.05, 1.21)	1.026 (0.93, 1.12)	1.084 (0.99, 1.17)	0.838 (0.73, 0.95)
Algorithmic Proprietary Account	5.036 <0.001	2.540 <0.001	3.185 <0.001	1.808 <0.001	2.061 (1.94, 2.19)	1.415 (1.26, 1.61)	1.605 (1.47, 1.74)	1.339 (1.10, 1.59)
Algorithmic Customer Account	1.931 <0.001	1.019 <0.001	1.528 <0.001	1.024 <0.001	1.106 (1.05, 1.16)	1.272 (1.17, 1.38)	1.107 (1.03, 1.20)	1.364 (1.18, 1.54)
New order at last cancel price	0.847 <0.001	1.856 <0.001	0.853 <0.001	1.674 <0.001	1.175 (1.11, 1.24)	1.408 (1.27, 1.56)	1.045 (0.97, 1.11)	1.224 (1.08, 1.37)
Quantity cancelled	0.999 <0.001	0.999 <0.001	0.998 <0.001	0.999 0.001	1.001 (1.00, 1.00)	1.000 (0.99, 1.00)	1.001 (1.00, 1.00)	1.000 (0.99. 1.00)
Time-varying marketwide covariates								
Quantity executed during the latency gap			0.803 <0.001	0.784 <0.001			0.783 (0.78, 0.79)	0.781 (0.77, 0.79)
New order quantity during the latency gap			1.124 <0.001	1.137 <0.001			1.139 (1.13, 1.15)	1.137 (1.13, 1.14)
Quantity cancelled during the latency gap			1.012 <0.001	0.863 0.206			0.864 (0.86, 0.87)	0.869 (0.86, 0.87)
Generalized R-Sqrd		28.2%		45.4%		2.2%		58.9%
Percent Censored		3.3%		3.7%		16.5%		17.9%
Observations/Accounts		722,260		632,581		12,765		11,767