Cancellation Strategies and Adverse Selection Costs

Raymond P. H. Fishe, Esen Onur, and Tugkan Tuzun

Abstract

After the Dodd-Frank legislation of 2010, regulators focused attention on market participants whose trading strategies involve a large number of order cancellations relative to their actual executions. Many observers interpret such behavior as suggestive of "spoofing"; that is, entering orders without intent to trade. However, a cancelled order is the primary outcome for over 90% of all order submissions according to the Securities and Exchange Commission, so strategies other than spoofing are likely at play. Specifically, order cancellations can be strategically used to reduce adverse selection costs, particularly for participants working larger order sizes in a high frequency environment with algorithms tuned to detect parent/child order submissions. The results of this research are expected to offer insight into the degree to which participants use cancellation (and modification) messages to optimally avoid adverse selection costs, how their latency affects their ability to achieve this, and how a recent change in the public data feed by the exchange affects market participants' cancellation strategies.

Methodology

This study examines the use of cancel messages[1] by traders in the e-Mini S&P 500 futures market, one of the largest equity derivative markets in the world. By creating pseudo-order books, we propose to examine whether cancel decisions help to avoid adverse selection costs. These pseudo-books allow us to track whether a particular cancelled order would have executed within the trading session. Thus, we may calculate whether adverse selection costs are avoided or whether participants missed profitable trading opportunities with their cancel decisions. This approach is an advance on previous methods that impute shortfall costs using executions at a closing price or a bid-ask midpoint near a cancellation.

[1] We also plan to implement the methods developed for cancellation messages to modification messages. Many modifications are effectively a combination of a cancel message and a new order submission as it places the order in a different position in the order book queue.

It is well-documented in the literature that certain market participants prefer to invest in being faster by trying to minimize the latency between decisions to place or cancel an order and the order execution. It is hypothesized that market participants investing in a low-latency setup technology, such as co-location to the matching engine, might be more successful at avoiding adverse selection than those who are subject to longer latency when trading. We propose to add to this literature by documenting whether speed helps market participants avoid being adversely selected.

In addition, we propose to examine how an important change to the E-mini market public data feed affected such cancel strategies. Beginning in 2017, the Chicago Mercantile Exchange (CME) Group instituted a method of tracking the position of an order on the book. This new data feed, known as "market by order" (MBO), allows a participant to identify exactly the number and the size of orders in front and behind her own order(s). Thus, the incentive to cancel to avoid execution may be refined to allow an order to remain on the book longer than what would have been feasible under the older data feed. It is hypothesized that this data feed will decrease the adverse selection costs avoided by cancellation because participants have additional time to evaluate market trends and make optimal timing decisions.

Data

The data to be examined in this study are part of the regulatory oversight feeds received by the Commodity Futures Trading Commission (CFTC). The CFTC receives daily files from regulated exchanges on trades, positions, and recently on order message flows. These files contain unique identifiers for market participants, not only allowing us to classify traders according to their latency but also giving us the ability to aggregate all their orders in our sample and compute an accurate adverse selection cost estimate.

For a typical trading day, files received by the CFTC contain millions of records and require software to manage such large data. The proposed project will use a massively parallel processing computer cluster service provided by a big data analytics company to analyze the raw data and recreate the pseudo-books using efficient sort and merge procedures that implement specialized algorithms developed by the

researchers in PostgreSQL for this project. Statistical analysis will be done using the SAS software on CFTC's servers.