THEORY OF QUEUING AND THE DETERMINATION OF THE BID-ASK SPREAD:
AN EQUILIBRIUM MICROSTRUCTURE ANALYSIS

Thomas A. Rhee
Professor of Finance
California State University, Long Beach
Long Beach, California 90840, USA
Email: Thomas.rhee@csulb.edu

Abstract

This paper studies the effect of bid-ask spread on the optimal trading strategies. We find that much of the bid-ask spread can be explained by how long it takes before orders get filled, and the anticipated fill time depends on the average order volumes unfilled across all venues. We explain this phenomenon via the Little’s Law in queuing theory. In this paper, we experiment how traders can forecast the order fill time by assuming that order arrivals come in the Poisson stochastic process with the properties of “orderliness” and “memorylessness.” We then develop a set of trading strategies in a partial equilibrium context. Whether to place a limit order or at market is also shown as a function of the predicted order fill time. In particular, we demonstrate that the optimal bid and ask spread is only the size of the tick in the market quote, if the predicted fill time is “short,” in which case the market order strategy appears to be optimal.

Keywords: Limit order trading, Poisson distribution, Little’s Law, queuing theory, bid-ask spread
With the advancement of computer technology, the securities industry has experienced a surging popularity of high frequency algorithmic trading. In response to the industry’s demand, recently, a new line of financial engineering research started to emerge in the area of microstructure of securities trading. See for example some recent research by Guilbaud and Pham (2011),\textsuperscript{1} Avellaneda and Stoikov (2008),\textsuperscript{2} Stoll (2003)\textsuperscript{3} and O’Hara (1997).\textsuperscript{4} Much earlier scant and yet notable researches on the capital market phenomenon of the optimal bid-ask spread are also found in Ho and Stoll (1980, 1981),\textsuperscript{5} and Ho and Macris (1984),\textsuperscript{6} Copeland and Galai (1983),\textsuperscript{7} Glosten (1994),\textsuperscript{8} and Handa and Schwartz (1996).\textsuperscript{9}

Typically, these researches are focused on the optimal limit and market order strategies in the context of minimizing inventory risks. These researches, however, miss some important aspects of real world trading: limit orders placed, unless order prices are “marketable,” add liquidity, while the market orders remove liquidity. Consequently, many exchanges reward market makers in the form of fee rebates for creating liquidity and “penalize” them by charging fees for removing liquidity.\textsuperscript{10} Clearly, it would be beneficial to place limit orders but such orders may not be filled for some time or ever, if the market moves away from the trader.\textsuperscript{11} This not only creates inventory risk but also put market makers out of their buying power limit in the algorithmic trading. Given this, the research by Ho and Stoll, Guilbaud and Pham, and Avellaneda and Stoikov ignores the aspect of liquidity, and the liquidity is not free. Many market makers are incentivized with rebates to provide liquidity, as it ensures the market continuity and eventually, the success of the exchange. Therefore, the liquidity is well priced in the bid-ask spread. Readers are cautioned, however, that the term here, the liquidity, has very little to do with how “marketable” a security is necessarily unlike many researchers identify the bid-ask spread as a measure of the security’s marketability. What is relevant here is the market continuity rather than the security’s marketability.

\textsuperscript{1} Fabien Guilbaud and Huyen Pham, “Optimal High Frequency Trading With Limit And Market Orders”, working paper series, University of Paris, 2011.
\textsuperscript{2} Marco Avellaneda and Sasha Stoikov, “High Frequency Trading In A Limit Order Book”, Quantitative Finance, 8(3), 217-224.
\textsuperscript{8} Lawrence R. Glosten, “Is the Electronic Open Limit Order Book Inevitable?” the Journal of Finance, 49 (4), September 1994, 1127-1161.
\textsuperscript{10} It should also be noted that some exchanges are also known as “inverted,” in the sense that taking liquidity is rewarded while liquidity providers are penalized as in BSX and EDGX, etc.
\textsuperscript{11} Generally, traders have the option to have their limit orders rerouted to the inverted market at no cost, when the decision to do so is profitable.
Trading profits are realized only if orders are filled. Handa and Schwartz see that the market order offers the guaranteed profit/loss opportunity, assuming of course there are enough shares available in the market, but costs money for the fills under certainty. In the meantime, the limit orders, if filled, can bring a larger trading profit in addition to the rebates received for the liquidity created. However, these potential gains are uncertain. Therefore, the relevant trading profit/loss is the expected profit, when there is a chance that the order may not be filled. Furthermore, there are hidden limit order costs as well, when placing a firm, as opposed to subject, limit order\(^{12}\) to buy resembles writing a call option on stocks at zero cost. Therefore, the limit orders have their own costs, which can be quite substantial, in exchange for a higher expected trading profit. Again, our intuition is that these costs and benefits associated with the market and/or limit orders must be well embedded in the bid-ask spread.

The purpose of this paper is to show how the bid-ask spread is determined in the marketplace in equilibrium and show, in the end, how the expected fill time for various orders influence the spread. In what follows, we first argue that the equilibrium spread between the market inside bid and the ask, which refers to the greatest bid and the lowest ask, is at least as much as the size of the tick,\(^{13}\) even when all trades are driven only by liquidity. Obviously, there are significant bid-ask spreads, even when the Direct Market Access is available even to small investors. If that were the case, the theory of bid and ask goes much beyond this liquidity consideration, although it is an important issue that needs to be reckoned with.

We believe that in the multiparty trading environment, the decision to place the market vs. limit orders influences the bid-ask spread in the economy. So, we examine the behavior of the rate of the arrival of various orders to buy or sell, and the rate of their departures. In this case, we find the queuing theory quite useful to learn that the average number of orders in a queuing system, \(L\), equals the average rate at which orders arrive, \(\lambda\), multiplied by the average time that an order spends in the system without being filled, \(W\), i.e. \(L = \lambda W\). This is the famous Little’s Law.\(^{14}\) By assuming for simplicity that all orders, both for the buys and sells, follow a stochastic Poisson process, we calculate the mean holding time for orders and suggest an algorithm to predict the optimal bid-ask spread.

We now first look at how the bid and ask spread is determined when all trades are liquidity driven.

**The Bid/Ask Spread in Liquidity Driven Trades**

Assume that the allowed tick size is given by \(s\), and the bid-ask spread is given by \(ns\), where \(n\) is an integer representing the number of ticks between the bid and ask. In this case, we will first note that

\(^{12}\) Some limit orders can be subject to change or even withdrawn. They are called the subject orders. However, this is rare and definitely, is not common among market makers.

\(^{13}\) This is one (1) cent for common stocks in U.S. Some other securities such as bonds, futures, and options, may have the tick size different from 1 cent, e.g. 1/8 of a dollar, 1/32 of a dollar, etc.

the equilibrium optimal bid-ask spread is where \( n = 1 \). A simple and straightforward proof is given here by assuming for simplicity, without the loss of generality, that for every buy order there exists a sell order in the market and that the market will print the trade whenever the order price matches. We will assume that all orders get filled when the price crosses without delay. Later, we discuss this in the context of queuing theory and show how the expected fill time is associated with the equilibrium bid-ask spread, if it takes a while before orders get filled.

**Theorem** – Assuming that all limit orders are liquidity driven and can be filled as placed without any delay and at no other particular cost, the equilibrium bid/ask spread in the competitive capital market is at least as great as the minimum allowed tick size, i.e. \( n = 1 \) with \( s = 1 \) cent.

**Proof.** If \( n = 2 \), place a buy order of \( x \) number of shares at ask \(-\) 1, while placing a sell order of \( y \) number of shares at bid \( +1 \), where \( x = y \). The profit is \((\text{bid} + 1) - (\text{ask} - 1) = \text{bid} - \text{ask} + 2 = -(\text{ask} - \text{bid}) + 2 = -2 + 2 = 0\). Notice that the buy limit order may not be placed at ask, since it would have been a “marketable” limit, in which case the order is in essence a market order. The buy limit order could also have been placed at ask \( -2 \), in which case, the profit could have been \(-(\text{ask} - \text{bid}) + 4 = -2 + 4 = 2 > 0\). Competition in the marketplace drives the bid down and the ask up. In equilibrium, a positive profit becomes hard to get. Therefore, placing a buy limit at ask \(-1\), and a sell limit at bid \(+1\) would be at best the only viable option. It should be noted that the process is a zero-sum game when multiple parties are involved.

Similarly, if \( n = 3 \), place a buy order of \( x \) number of shares at ask \(-2\), while placing a sell order of \( y \) number of shares at bid \(+1\), where \( x = y \). In the efficient capital market, this could be the best strategy, as before. The profit is \((\text{bid} + 1) - (\text{ask} - 2) = \text{bid} - \text{ask} + 3 = -(\text{ask} - \text{bid}) + 3 = -3 + 3 = 0\).

If \( n = 4 \), place a buy order of \( x \) number of shares at ask \(-2\), while placing a sell order of \( y \) number of shares at bid \(+2\), where \( x = y \). The profit is \((\text{bid} + 2) - (\text{ask} - 2) = \text{bid} - \text{ask} + 4 = -(\text{ask} - \text{bid}) + 4 = -4 + 4 = 0\). And if \( n = 5 \), place a buy order of \( x \) number of shares at ask \(-3\), while placing a sell order of \( y \) number of shares at bid \(+2\), where \( x = y \). The profit is \((\text{bid} + 2) - (\text{ask} - 3) = \text{bid} - \text{ask} + 5 = -(\text{ask} - \text{bid}) + 5 = -5 + 5 = 0\), etc.

Thus, if \( n \neq 0 \), it must be true that \( n = 1 \). Therefore, if an individual trader places a sell order at bid \(+\ (n/2)\) and a buy order at ask \(-\ (n/2)\), then, the trader’s profit is, in equilibrium, \((\text{bid} + (n/2)) - (\text{ask} - (n/2)) \leq 0\), or the bid-ask spread, \((\text{ask} - \text{bid}) \geq n = 1\). This establishes the lowest boundary for the spread. In reality, however, providing liquidity is awarded and taking or removing it is penalized. If \((x - y) = 0\), meaning that for each buy or sell order, there are corresponding sell or buy orders, then, it must be true that either buy or sell order must precede the other, and thus, if the order that preceded added liquidity to the market, the order that followed must have taken liquidity. If \( \pi \) is the market maker’s rebate incomes received from exchanges and \( SR \) is paid per share for liquidity creation from a nonmarketable limit order, and it costs \( SC \) a share for taking liquidity, the economy that places limit orders on one side and places equivalent market orders on the opposite side must have the equilibrium bid-ask spread, \((\text{ask} - \text{bid}) \geq n + (R - C) = (R - C) + 1\).

**Application to Trading Strategies**

The arguments that we used to prove that the equilibrium \((\text{ask} - \text{bid}) \geq n = 1\) can be used in reality as a trading strategy, especially when traders have difficulties to have orders filled.
Suppose that over a given trading horizon, a market maker wishes to buy \( q \) shares at a price limit and cannot have his orders filled, either because the bid and ask do not budge or because the market has now moved away from the price limit, that is, the market is on the rise. How would the market maker create liquidity and provide continuity in the market for the stock that he or she makes the market for?

Suppose that one buys \( x = q + \Delta b \) shares at \( \text{ask} - \text{roundup}(n/2) \) and sell \( y = \Delta s \) shares for \( \text{bid} + \text{rounddown}(n/2) \), where both \( \Delta b \) and \( \Delta s \) are nonnegative and \( n \) is an odd integer. As stated, the value of \( n = 1 \) in equilibrium. If \( n \) is even, buy \( x = q + \Delta b \) shares at \( \text{ask} - (n/2) \) and sell \( y = \Delta s \) shares for \( \text{bid} + (n/2) \).

Then, if \( \Delta b = \Delta s \equiv \Delta q \) such that the net inventory is \( x - y = q \) which is what the market maker wants to have filled, and if \( C - R = \emptyset \), then, the market maker’s rebate function is given by eqn (1), regardless of which order has preceded, i.e. ether a \( \Delta b \) buy order or a \( \Delta s \) sell order.

\[
\pi = R(q + \Delta q) - C\Delta q = Rq - \emptyset \Delta q
\]  
(1)

Obviously, the profit maximizing value of \( \Delta q \) would be zero. On the other hand, the value of \( \Delta q \), which gives a zero profit is

\[
\Delta q = \frac{Rq}{\emptyset}
\]

For example, NASDAQ charges an exchange fee of $0.0030 per share for taking liquidity and rebates $0.0020 per share for adding liquidity. Rounding down the formula for \( \Delta q \), the maximum tolerable values of \( \Delta q \), which will result in zero profit for each possible buy quantities are as follows.

<table>
<thead>
<tr>
<th>Venue</th>
<th>Taking liquidity</th>
<th>Adding liquidity</th>
<th>Phi</th>
<th>100</th>
<th>200</th>
<th>300</th>
<th>400</th>
<th>500</th>
<th>600</th>
<th>700</th>
<th>800</th>
<th>900</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASDAQ</td>
<td>0.0030</td>
<td>0.0020</td>
<td>0.0010</td>
<td>200</td>
<td>400</td>
<td>600</td>
<td>800</td>
<td>1000</td>
<td>1200</td>
<td>1400</td>
<td>1600</td>
<td>1800</td>
<td>2000</td>
</tr>
</tbody>
</table>

Thus, for example, if one wants to buy 300 shares on the NASDAQ venue, a combination of 900 shares of buy orders and 600 shares of sell orders will bring zero profit. In reality, \( \Delta q \) should be set something less than 600. In fact, \( \Delta q = 0 \) will give the maximum profit. Thus, what we do is set the value \( \Delta q \) at something like \( \Delta q = 100 \), a number less than 600. That will result in a positive profit of $0.50. That means buying 400 shares and selling 100 shares. Temporarily, this could lure others to sell 300 shares by creating liquidity in the market.

Similar arguments can be made for the unfilled sell orders. For example, one may wish to sell \( q \) shares and may not have the orders filled. Unfortunately, orders are executed in orders in which they arrive in the market and in fact, the orders that the market maker places would be the last order that he must await until before all other orders are filled before him. Therefore, we need to look at the stochastic process in which orders arrive and depart. In other words, all orders are not filled instantaneously without delay. We believe that the stochastic process of order arrivals and fills should be an important part in the determination of the equilibrium bid-ask spread.
Queuing Theory in Limit Order Trading

We now apply the fundamental law in queuing theory. Now enter the Little’s Law.

The trading venue is like a channel server, and we argue that the average number of orders in a queuing system, $L$, equals the average rate at which orders arrive, $\lambda$, multiplied by the average (waiting) time, $W$, that an order spends in the system without being filled, i.e. $L = \lambda W$. This relationship widely known as the Little’s Law can be quite useful in that the average number of orders in a queuing system can be construed as the number of unfilled orders that are still awaiting to be filled. Noting that the orders are filled or “depart” through the supply and demand, the average number of unfilled orders results from an equilibrium process. The Little’s Law says that the average number of unfilled orders depends on the rate at which orders arrive and the average time, $W$, that an order is in the trading exchange venue, which also depends on how orders are filled according to a first-in, first-out (FIFO) rule.\textsuperscript{15} As usual, the Little’s Law solves for the last remaining quantity, when two quantities are known.

In this paper, we assume that orders arrive in a fixed parameter Poisson process, and we compute the Poisson parameter $\lambda$ from the Maximum Likelihood Estimation technique. We then illustrate how the average waiting time is computed through an equilibrium process. We then show that the average waiting order fill time, $W$ and the Poisson parameter $\lambda$ will determine the average number of orders in a queuing system, $L$. The author believes that this research is the first of its kind to offer an explanation about how the bid and ask spread is determined or will move.

To be consistent with the characteristics of the Poisson process, we assume that (1) the probability that each order will occur in a small time interval, or a trading horizon,\textsuperscript{16} is proportional to the length of the interval; (2) the probability that more than one orders will occur in a small interval is substantially smaller than the probability that a single order will occur, i.e. “orderliness,” unless some compound processes are accompanied, that is, unless the order price and/or size differs with special order qualifiers; (3) the number of orders in non-overlapping time interval is independent, i.e. “memorylessness,” as each order also follows a Markov process; and (4) the expected number of orders between time $t$ and time $t + s$ is independent of $t$.

Just as in the case of the traditional queuing problem, all orders are assumed to be “conserved” in the economy’s trading system as a whole. If orders are withdrawn or cancelled, we will treat them as if they have been “departed.” If arriving buy orders are met by arriving sell orders, we will also treat them as if they have been “departed.” Therefore, the average number of unfilled orders results from the other traffic, that is, the sell traffic to the buyers and the buy traffic to the sellers. If any limit orders are not filled, there has been a de-synchronization of these two queues, and when this de-synchronization occurs, we say that the market has moved away from the price limit leaving one side of the queues in the channel server.

\textsuperscript{15} In practice, there is a security trade execution order in which the first bid or offer price is executed before other bid and offer prices. Priority executions are undertaken regardless of the volume of the order, meaning that a larger order that does not get to be executed first based solely on its size. Market rules often require the first trade received to be executed first. If two bids or offers are received at the same time, the bid or offer at the larger volume is given priority and is executed first. In this case, the smaller bid or offer is rejected, and the broker placing the bid is told that the rejection occurred because there was a stock ahead.

\textsuperscript{16} Orders are often placed on a \textit{Good till Time} (GTT) basis.
Equilibrium Model of Expected Order Fill Times

Therefore, if \( N(t) \) denotes the cumulative number of arrivals by time \( t \) in the trading horizon, \( H \), and \( D(t) \) is the cumulative number of departures (or fills), including the number of orders being withdrawn, where \( N(0) = 0, N(H) = N, D(0) = 0, D(H) = 0 \), and \( N(t) \geq D(t) \), then the number of orders in the system at time \( t \):

\[
L(t) = N(t) - D(t)
\]

Therefore, the average number of unfilled outstanding orders over \( H \) is

\[
L = \frac{1}{H} \times \int_{t=0}^{H} (N(t) - D(t)) \, dt.
\]

It will be convenient to quantify the number of arrivals and departures in terms of the number of shares placed to buy and sell. This means that when the market does not clear over a short time horizon, \( H \), the average number of unfilled orders, \( L \), represents the disequilibrium, that is, either excess supply of or excess demand for shares.

The number of order arrivals before time \( t \), \( N(t) \), is an integer valued non-decreasing random function of time from 0 to \( t \). Here, a Poisson process is \( \lambda(t) \), which is the expected number of events or arrivals. Assuming for simplicity that the arrival process is homogeneous, \( \lambda(t) = \lambda \) in which case \( N(t) \) has a Poisson distribution with expected value \( \lambda t \).

The following must hold.

If \( N(t) \) is a discrete Poisson distribution, the (waiting) time of the \( j \)th arrival with the sequence of arrival \( \{s_1, s_2, ..., s_N\} \) is a continuous probability distribution.\(^{17}\) The Poisson distribution has the property often known as the memorylessness that the conditional probability that we need to wait more than so many more minutes, let’s say, another 10 minutes, before the first arrival, given that the first arrival has not yet happened after some earlier minutes, let’s say, 30 minutes, is no different from the initial probability that we need to wait more than 10 minutes for the first arrival. In other words,

\[
\text{Prob}[s_1 > 40, | s_1 > 30] = \text{Prob}[s_1 > 10]
\]

Because the inter-arrival times are exponentially distributed, the time between the 4th and 9th arrival (for instance) is distributed as the sum of exponential random variables, i.e. 5th order gamma distribution. Also, these conditions imply that the probability distribution of the number of events in the interval \([t_1, t_2]\), which is also written as \( N(t_2) − N(t_1) \) is Poisson distributed with parameter \( \lambda(t_2 − t_1) \) for homogeneous processes. Thus, for \( t = t_2 − t_1 \),

\(^{17}\) We have a situation of a 1-dimensional Poisson process, which involves both the discrete Poisson distributed random variables that count the number of arrivals in each time interval, and the continuous Erlang-distributed waiting times. It is known that the mean and the variance of a Poisson random variable are \( \lambda \) each. Skewness is \( \lambda^{-1/2} \), while kurtosis is \( \lambda^{-1} \).
\[ \text{Prob}[N(t) = k] \equiv p(k, \lambda t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}. \]

Over a longer period of time, \( t \), this represents the probability that the event occurs exactly \( k \) times. Noting that the mean of the Poisson distribution is \( \lambda t \), for example, if the probability of the arrival of a buy order is \( \lambda = 2\% \) per second, then, the probability of seeing no other buy orders arriving in any given second is

\[ p(k, \lambda t) = p(0, 0.02 \times 1) = \frac{e^{-0.02 \times 1} (0.02 \times 1)^0}{0!} = 0.9802. \]

The probability of seeing exactly two buy orders over the next 10 second period is

\[ p(k, \lambda t) = p(2, 0.02 \times 10) = \frac{e^{-0.02 \times 10} (0.02 \times 10)^2}{2!} = 0.0164. \]

Thus, for example, if we have to wait more than 100 minutes before we see another buy order that takes place 10 times, i.e. \( s_j \equiv s_{10} > t = 100 \), it must be true that the number of arrivals before \( t \) must be less than 10 times, i.e. \( N(t) \equiv N(100) < x = 10 \). Clearly, the number of arrivals before \( t \) is less than \( x \), if and only if the waiting time until the \( j^{th} \) arrival is more than \( t \). Thus,

\[ \text{Prob}[N(t) < x] = \text{Prob}[s_j > t] \]

Put it differently, consider the most simplistic case where the rate function \( \lambda^a(t) = \lambda^a \) and \( x = 1 \). Then, knowledge of the Poisson distribution results in

\[ \text{Prob}[s_1 > t] = \text{Prob}[N(t) = 0] = e^{-\lambda^a t}, \]

which is an exponential distribution. Thus, the average rate of arrival has expected value of this exponential distribution, i.e. \( 1/\lambda^a \). In other words, if the average rate of arrival is 6 per minute, the average waiting time until the first arrival is 1/6 minutes.

Similar arguments can be made also for the random process, \( D(t) \). Thus, if \( N(t) \) and \( D(t) \) are two independent random variables, both following a Poisson distribution with parameters \( \lambda^a \) and \( \lambda^d \), respectively, then the “orderliness” property of Poisson means that \( L(t) = N(t) - D(t) \) also follows a Poisson distribution with parameter \( \lambda^a - \lambda^d \). We will estimate these values of \( \lambda^a \) and \( \lambda^d \), the aggregate flow of buy orders and sell orders later on in this paper.

In reality, no orders get filled immediately, and thus, there always is a chance that \( L(t) > 0 \), i.e. the excess demand in the market, with the probability \( p \), and also, \( L(t) < 0 \), i.e. the excess supply in the market, with the probability \( 1 - p \). This means that the price may be rising immediately with \( L(t) > 0 \) or may be falling with \( L(t) < 0 \). The conventional wisdom is that with \( L(t) > 0 \), the limit order to buy may not be filled; and so, the buy order should be at market. Similarly, all sell orders should be at market if \( L(t) < 0 \). Unfortunately, however, choosing a particular order type is not as simple, as this argument assumes that the trade has already taken place at a price. We now consider a few important market-making dynamics in an important environment where all limit orders are at “good till time...
“certain (GTT)” and all market orders are placed in the form of the *marketable* limit orders as “all or nothing (AOL)” and “immediate or cancel (IOC).”

To illustrate, assume that the market maker’s reservation buy price is higher than the current ask and places a buy order at market, which gets subsequently filled. If the total quantity traded in the market as a whole is $Q^*$ as shown in Figure 1, it is clear that the price must be falling, since at the currently trade price at the market ask, $L(t) = N(t) - D(t) < 0$. The next buy order should be at limit, while the next sell order must be at market. If the initial market buy order is not filled, however, we would use the market buy order and sell limit order. But this may not be the case if the buy order was placed at a price limit below the point $c$, and was executed. In this case, $L(t) > 0$, and the price will rise. So, the price can send a conflicting signal, unless the buy limit was at a price above the point $c$. Note also that if the bid limit was below the point $b$, the buy order would not have been filled unless the order price is above the reservation sell price of those market makers other than the marginal sellers. In reality, these offer prices are the second best, the second to second best, the second to the second to the second best offers, which are shown in the dealer’s limit order books, and perhaps, those offers in the dark pools.

Figure 1: The bid-ask spread is the vertical distance between the point $a$ and $b$. The actual trade will take place somewhere between $p_{ask}$ and $p_{bid}$.

Now suppose that the initial order was on the sell side. If the sell order was at market, i.e. at the current bid, and gets subsequently filled, the market must be anticipating that the price will rise. $L(t) = N(t) - D(t) > 0$. We are assuming here that the market bid is higher than the seller’s reservation price. The

---

18 Often called *dark liquidity*, the term refers to bid/ask sizes available only to institutional traders, whose information is not available, “dark”, on the trading *graphical user interface* (GUI) available to regular individual traders.
same result follows unless the initial sell order was a limit order placed above the point \( c \). Obviously, if the order price was above \( c \), \( L(t) < 0 \) and the price can send a conflicting signal.

To summarize, we now introduce a notational convention that if the buy order at time \( t \) is filled at \( t + h \), the event is denoted as \( B^i(t, t + h) = 1 \), or \( B^i(t, t + h) = 0 \), where \( i \) is either a limit order (\( \ell \)) or a market order (\( m \)). Similarly, for the sell order as well, i.e. \( S^i(t, t + h) = 1 \) and \( S^i(t, t + h) = 0 \). The market vs. limit order algorithm is summarized as follows.

\[
\begin{align*}
\text{Buy limit or Sell at market at } t & \quad \begin{cases} 
\text{if } B^m(t - h, t) = 1, & \text{or} \\
\text{if } B^\ell(t - h, t; > c) = 1, & \text{or} \\
\text{if } B^\ell(t - h, t; < c) = 0, & \text{or} \\
\text{if } S^\ell(t - h, t; > c) = 1
\end{cases} \\
\text{Buy at market or Sell limit at } t & \quad \begin{cases} 
\text{if } S^m(t - h, t) = 1, & \text{or} \\
\text{if } S^\ell(t - h, t; < c) = 1, & \text{or} \\
\text{if } S^\ell(t - h, t; > c) = 0, & \text{or} \\
\text{if } B^\ell(t - h, t; < c) = 1
\end{cases}
\end{align*}
\]

In other words, the strategy of placing a limit buy or a market sell order is predicated upon the fact that the price may fall shortly, i.e. \( L(t) < 0 \). Similarly, the strategy of placing a market buy order or a limit sell order is predicated upon the fact that the price may rise shortly, i.e. \( L(t) > 0 \). Note that we are implicitly assuming that all market orders are executed quickly. The notation, \( > c \) or \( < c \), signifies that orders are placed at above or below \( c \), where \( c \) is the equilibrium market trade price, which is unknown now. The point is that the price can be a misleading signal, and hence, may not be used to decide on whether the order should be placed either at market or at limit. Given this, we argue that the single most important statistic to whether the order should be placed either at market or at limit is the average number of outstanding shares still waiting to be filled, \( L(t) \). The queuing theory makes this possible.

The Queuing Theory and the Little’s Law

The Little’s law is quite useful to traders. Using the Little’s original work, we now model the average time in the system for each order. Define the sequence of arrival time for \( N \) orders as \( \{s_1, s_2, \ldots, s_N\} \); and similarly, the sequence of fill times for \( N \) orders as \( \{c_1, c_2, \ldots, c_N\} \). Then, the time spent in the system for the \( j \)th arriving order is \( W_j = c_j - s_j \), where the fill and the arrival time for the \( j \)th order is \( c_j \) and \( s_j \), respectively. The average time in the system over all the orders is:

\[
W = \frac{1}{N} \times \left( \sum_{j=1}^{N} c_j - \sum_{j=1}^{N} s_j \right).
\]

If \( c_j \) is the fill times for the \( j \)th filled order, then the sequence of the fill times for the \( N \) orders is \( \{c^1, c^2, \ldots, c^N\} \), which may be a permutation or reordering of the sequence\( \{c_1, c_2, \ldots, c_N\} \) from the mathematical standpoint. If each order is filled in sequence, \( c_j = c^j \). But it is not the case in general, especially if we consider multiple trading venues with many ECNs, and if the order size also matters in addition to the time of arrival.
Define the \(j^{th}\) wait time as \(W^j = c^j - s_j\) equal to the difference between the departure time for the \(j^{th}\) fill and the arrival time for the \(j^{th}\) order. Now consider the average of these wait times:

\[
\frac{1}{N} \sum_{j=1}^{N} W^j = \frac{1}{N} \left( \sum_{j=1}^{N} c^j - \sum_{j=1}^{N} s_j \right)
\]

Since \(\sum_{j=1}^{N} c^j = \sum_{j=1}^{N} c_j\),

\[
W = \frac{1}{N} \sum_{j=1}^{N} W^j = \frac{1}{N} \left( \sum_{j=1}^{N} c^j - \sum_{j=1}^{N} s_j \right)
\]

Now,

\[
\sum_{j=1}^{N} W^j = \frac{1}{H} \int_{j=1}^{H} \left( N(t) - D(t) \right) dt
\]

Then, Little’s Law is derived as follows.

\[
L = \frac{1}{H} \times \int_{j=1}^{H} \left( N(t) - D(t) \right) dt = \frac{1}{H} \sum_{j=1}^{N} W^j = \frac{N}{H} \times W = \lambda W
\]

This relationship holds over a finite time window with nonstationary arrivals and no notion of any steady state for the system for any \(N\), including, for example, \(N = 1\). But as Professor Little states, the relationship is exact, only after the fact, however. And thus, the average length of queue or the average number of orders in the trading system over a given trading horizon \(H\) can be used as a useful statistic that explains how the bid-ask spread is determined. The law says, in fact, that the higher the value of \(\lambda\), and the longer the average waiting time for orders to be filled, the larger the average number of orders or the length of queue on the trading venue.

Recall that we earlier stated that if \(N(t)\) and \(D(t)\) are two independent random variables, both following a Poisson distribution with parameters \(\lambda^a\) and \(\lambda^d\), respectively, then the “orderliness” property of Poisson means that \(L(t)\), which is defined as \(N(t) - D(t)\), also follows a Poisson distribution with parameter \(\lambda^a - \lambda^d\). Luckily, we can estimate these values quite simply by finding the value of \(\lambda\), which maximizes the function (often known as the Maximum Likelihood Estimation function) given by

\[
\mathcal{L}(\lambda) = \prod_{i=1}^{N} f(k_i; \lambda) = \prod_{i=1}^{N} \frac{\lambda^{k_i} e^{-\lambda}}{k_i!} = \frac{\lambda^{\sum_{i=1}^{N} k_i} e^{-\lambda N}}{\prod_{i=1}^{N} k_i!}
\]
where the sums and products are from $i = 1$ to $N$. Taking the logarithm of $L$ and then taking the derivative with respect to $\lambda$ and equating it to zero yields the MLE estimate of $\lambda$. The solution turns out to be the average value of $k_i$'s. From the property of the characteristic function, it is also seen that the expected value of $\lambda_{MLE}$ is itself equal to $\lambda$; and hence, the average value of $k_i$'s is an unbiased estimator of $\lambda$.

**Numerical Example**

We now use actual data to estimate the mean holding waiting time before an order is filled on the average. In this paper, we have decided to look at one particular stock's real-time market data for January 31, 2014 from 9:30 in the morning when the market opens until 1:30 pm for four hours. The data we chose was Apple Computer (AAPL), as the symbol has reasonably active trades. To estimate the values of $\lambda^a$ and $\lambda^d$, the aggregate flow of buy and sell orders, we have assumed a short trading horizon of five minutes. We also assumed that the number of events for buy orders and sell orders is the bid and ask events over the five subsequent trading minutes. The result is reported in Table 1.

**Table 1: Computation of the Expected Waiting Time for Order Fills**

<table>
<thead>
<tr>
<th>Trade Time</th>
<th>Total Buy Flows in shares Column (1)</th>
<th>Total Sell Flows in shares Column (2)</th>
<th>Unmatched Shares per sec Column (3)</th>
<th>Avg Unfilled Shares per sec Column (4)</th>
<th>No of Bid Events Column (5)</th>
<th>Avg Bid Events per sec Column (6)</th>
<th>No of Ask Events Column (7)</th>
<th>Avg Ask Events per sec Column (8)</th>
<th>Avg Unfilled Events per sec Column (9)</th>
<th>Avg Waiting Time to Fill in sec Column (10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>935</td>
<td>2369</td>
<td>3887</td>
<td>-1518</td>
<td>-5.06</td>
<td>736</td>
<td>2.45</td>
<td>582</td>
<td>1.94</td>
<td>0.51</td>
<td>9.86</td>
</tr>
<tr>
<td>940</td>
<td>921</td>
<td>1639</td>
<td>-718</td>
<td>-2.39</td>
<td>431</td>
<td>1.44</td>
<td>386</td>
<td>1.29</td>
<td>0.15</td>
<td>15.96</td>
</tr>
<tr>
<td>945</td>
<td>928</td>
<td>876</td>
<td>52</td>
<td>0.17</td>
<td>403</td>
<td>1.34</td>
<td>373</td>
<td>1.24</td>
<td>0.10</td>
<td>1.73</td>
</tr>
<tr>
<td>950</td>
<td>914</td>
<td>611</td>
<td>303</td>
<td>1.01</td>
<td>405</td>
<td>1.35</td>
<td>347</td>
<td>1.16</td>
<td>0.19</td>
<td>5.22</td>
</tr>
<tr>
<td>955</td>
<td>1188</td>
<td>655</td>
<td>533</td>
<td>1.78</td>
<td>426</td>
<td>1.42</td>
<td>294</td>
<td>0.98</td>
<td>0.44</td>
<td>4.04</td>
</tr>
<tr>
<td>1000</td>
<td>1114</td>
<td>945</td>
<td>169</td>
<td>0.56</td>
<td>431</td>
<td>1.44</td>
<td>324</td>
<td>1.08</td>
<td>0.36</td>
<td>1.58</td>
</tr>
<tr>
<td>1005</td>
<td>756</td>
<td>1360</td>
<td>-604</td>
<td>-2.01</td>
<td>349</td>
<td>1.16</td>
<td>396</td>
<td>1.32</td>
<td>-0.16</td>
<td>12.85</td>
</tr>
<tr>
<td>1010</td>
<td>678</td>
<td>719</td>
<td>-41</td>
<td>-0.14</td>
<td>298</td>
<td>0.99</td>
<td>325</td>
<td>1.08</td>
<td>-0.09</td>
<td>1.52</td>
</tr>
</tbody>
</table>

19 Let us assume that pure-chance arrivals and pure-chance terminations have the following properties that (1) the number of buy order arrivals in a given time has a Poisson distribution or Poisson traffic that $P(\alpha) = \frac{\lambda^a\alpha^k}{k!} e^{-\lambda}$ where $\alpha$ is the number of buy order arrivals in time between $t_1$ and $t_2$; and $\lambda$ is the mean number of buy order arrivals in time $t$; and (2) the number of buy order departures, i.e. the number of arrivals of sell orders or rate of withdrawals of the existing buy orders in a given time, also has a Poisson distribution, $P(d) = \frac{\lambda^d}{d!} e^{-\lambda}$ where $d$ is the number of buy order departures in time $t_1$ and $t_2$; and $\lambda$ is the mean number of buy order departures in time $t$. Then, the intervals, $t_1$ and $t_2$, between the order arrivals and departures are intervals between independent, identically distributed random events. It is known that these intervals have a negative exponential distribution, $P[T \geq t_2 - t_1] = e^{-(t_2 - t_1)/\lambda}$ where $h$ is the mean holding time (MHT). In this section, we estimate the MHT by using the Little’s Law.

20 We acknowledge that Tara Advisors LLC has provided some of their internal data source for this research.

21 We have decided to limit this study only to 4 hours, as increasing the sample size would not add anything to the research.
<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1015</td>
<td>688</td>
<td>1194</td>
<td>-506</td>
<td>-1.69</td>
<td>261</td>
<td>0.87</td>
<td>295</td>
<td>0.98</td>
<td>-0.11</td>
<td>14.88</td>
</tr>
<tr>
<td>1020</td>
<td>733</td>
<td>510</td>
<td>223</td>
<td>0.74</td>
<td>281</td>
<td>0.94</td>
<td>267</td>
<td>0.89</td>
<td>0.05</td>
<td>15.93</td>
</tr>
<tr>
<td>1025</td>
<td>512</td>
<td>724</td>
<td>-212</td>
<td>-0.71</td>
<td>242</td>
<td>0.81</td>
<td>284</td>
<td>0.95</td>
<td>-0.14</td>
<td>5.05</td>
</tr>
<tr>
<td>1030</td>
<td>807</td>
<td>556</td>
<td>251</td>
<td>0.84</td>
<td>328</td>
<td>1.09</td>
<td>277</td>
<td>0.92</td>
<td>0.17</td>
<td>4.92</td>
</tr>
<tr>
<td>1035</td>
<td>6606</td>
<td>653</td>
<td>5953</td>
<td>19.84</td>
<td>412</td>
<td>1.37</td>
<td>279</td>
<td>0.93</td>
<td>0.44</td>
<td>44.76</td>
</tr>
<tr>
<td>1045</td>
<td>1151</td>
<td>950</td>
<td>201</td>
<td>0.67</td>
<td>381</td>
<td>1.27</td>
<td>340</td>
<td>1.13</td>
<td>0.14</td>
<td>4.90</td>
</tr>
<tr>
<td>1050</td>
<td>510</td>
<td>768</td>
<td>-258</td>
<td>-0.86</td>
<td>198</td>
<td>0.66</td>
<td>270</td>
<td>0.90</td>
<td>-0.24</td>
<td>3.58</td>
</tr>
<tr>
<td>1055</td>
<td>689</td>
<td>1152</td>
<td>-463</td>
<td>-1.54</td>
<td>238</td>
<td>0.79</td>
<td>328</td>
<td>1.09</td>
<td>-0.30</td>
<td>5.14</td>
</tr>
<tr>
<td>1100</td>
<td>511</td>
<td>509</td>
<td>2</td>
<td>0.01</td>
<td>211</td>
<td>0.70</td>
<td>192</td>
<td>0.64</td>
<td>0.06</td>
<td>0.11</td>
</tr>
<tr>
<td>1105</td>
<td>338</td>
<td>444</td>
<td>-106</td>
<td>-0.35</td>
<td>171</td>
<td>0.57</td>
<td>201</td>
<td>0.67</td>
<td>-0.10</td>
<td>3.53</td>
</tr>
<tr>
<td>1110</td>
<td>463</td>
<td>601</td>
<td>-138</td>
<td>-0.46</td>
<td>191</td>
<td>0.64</td>
<td>241</td>
<td>0.80</td>
<td>-0.17</td>
<td>2.76</td>
</tr>
<tr>
<td>1115</td>
<td>380</td>
<td>355</td>
<td>25</td>
<td>0.08</td>
<td>194</td>
<td>0.65</td>
<td>153</td>
<td>0.51</td>
<td>0.14</td>
<td>0.61</td>
</tr>
<tr>
<td>1120</td>
<td>461</td>
<td>545</td>
<td>-84</td>
<td>-0.28</td>
<td>207</td>
<td>0.69</td>
<td>211</td>
<td>0.70</td>
<td>-0.01</td>
<td>21.00</td>
</tr>
<tr>
<td>1125</td>
<td>270</td>
<td>360</td>
<td>-90</td>
<td>-0.30</td>
<td>149</td>
<td>0.50</td>
<td>183</td>
<td>0.61</td>
<td>-0.11</td>
<td>2.65</td>
</tr>
<tr>
<td>1130</td>
<td>455</td>
<td>505</td>
<td>-50</td>
<td>-0.17</td>
<td>216</td>
<td>0.72</td>
<td>211</td>
<td>0.70</td>
<td>0.02</td>
<td>10.00</td>
</tr>
<tr>
<td>1135</td>
<td>597</td>
<td>520</td>
<td>77</td>
<td>0.26</td>
<td>262</td>
<td>0.87</td>
<td>253</td>
<td>0.84</td>
<td>0.03</td>
<td>8.56</td>
</tr>
<tr>
<td>1140</td>
<td>372</td>
<td>518</td>
<td>-146</td>
<td>-0.49</td>
<td>169</td>
<td>0.56</td>
<td>191</td>
<td>0.64</td>
<td>-0.07</td>
<td>6.64</td>
</tr>
<tr>
<td>1145</td>
<td>253</td>
<td>480</td>
<td>-227</td>
<td>-0.76</td>
<td>144</td>
<td>0.48</td>
<td>156</td>
<td>0.52</td>
<td>-0.04</td>
<td>18.92</td>
</tr>
<tr>
<td>1150</td>
<td>533</td>
<td>578</td>
<td>-45</td>
<td>-0.15</td>
<td>218</td>
<td>0.73</td>
<td>236</td>
<td>0.79</td>
<td>-0.06</td>
<td>2.50</td>
</tr>
<tr>
<td>1155</td>
<td>535</td>
<td>591</td>
<td>-56</td>
<td>-0.19</td>
<td>227</td>
<td>0.76</td>
<td>199</td>
<td>0.66</td>
<td>0.09</td>
<td>2.00</td>
</tr>
<tr>
<td>1200</td>
<td>299</td>
<td>564</td>
<td>-265</td>
<td>-0.88</td>
<td>177</td>
<td>0.59</td>
<td>203</td>
<td>0.68</td>
<td>-0.09</td>
<td>10.19</td>
</tr>
<tr>
<td>1205</td>
<td>759</td>
<td>1002</td>
<td>-243</td>
<td>-0.81</td>
<td>283</td>
<td>0.94</td>
<td>327</td>
<td>1.09</td>
<td>-0.15</td>
<td>5.52</td>
</tr>
<tr>
<td>1210</td>
<td>999</td>
<td>968</td>
<td>31</td>
<td>0.10</td>
<td>329</td>
<td>1.10</td>
<td>275</td>
<td>0.92</td>
<td>0.18</td>
<td>0.57</td>
</tr>
<tr>
<td>1215</td>
<td>1136</td>
<td>2102</td>
<td>-966</td>
<td>-3.22</td>
<td>379</td>
<td>1.26</td>
<td>334</td>
<td>1.11</td>
<td>0.15</td>
<td>21.47</td>
</tr>
<tr>
<td>1220</td>
<td>1238</td>
<td>1537</td>
<td>-299</td>
<td>-1.00</td>
<td>401</td>
<td>1.34</td>
<td>376</td>
<td>1.25</td>
<td>0.08</td>
<td>11.96</td>
</tr>
<tr>
<td>1225</td>
<td>836</td>
<td>592</td>
<td>244</td>
<td>0.81</td>
<td>252</td>
<td>0.84</td>
<td>264</td>
<td>0.88</td>
<td>-0.04</td>
<td>20.33</td>
</tr>
<tr>
<td>1230</td>
<td>428</td>
<td>521</td>
<td>-93</td>
<td>-0.31</td>
<td>202</td>
<td>0.67</td>
<td>242</td>
<td>0.81</td>
<td>-0.13</td>
<td>2.33</td>
</tr>
<tr>
<td>1235</td>
<td>607</td>
<td>394</td>
<td>213</td>
<td>0.71</td>
<td>206</td>
<td>0.69</td>
<td>186</td>
<td>0.62</td>
<td>0.07</td>
<td>10.65</td>
</tr>
<tr>
<td>1240</td>
<td>580</td>
<td>533</td>
<td>47</td>
<td>0.16</td>
<td>206</td>
<td>0.69</td>
<td>226</td>
<td>0.75</td>
<td>-0.07</td>
<td>2.35</td>
</tr>
<tr>
<td>1245</td>
<td>671</td>
<td>658</td>
<td>13</td>
<td>0.04</td>
<td>266</td>
<td>0.89</td>
<td>256</td>
<td>0.85</td>
<td>0.03</td>
<td>1.30</td>
</tr>
<tr>
<td>1250</td>
<td>556</td>
<td>433</td>
<td>123</td>
<td>0.41</td>
<td>219</td>
<td>0.73</td>
<td>177</td>
<td>0.59</td>
<td>0.14</td>
<td>2.93</td>
</tr>
<tr>
<td>1255</td>
<td>322</td>
<td>339</td>
<td>-17</td>
<td>-0.06</td>
<td>155</td>
<td>0.52</td>
<td>137</td>
<td>0.46</td>
<td>0.06</td>
<td>0.94</td>
</tr>
<tr>
<td>1300</td>
<td>428</td>
<td>278</td>
<td>150</td>
<td>0.50</td>
<td>165</td>
<td>0.55</td>
<td>152</td>
<td>0.51</td>
<td>0.04</td>
<td>11.54</td>
</tr>
<tr>
<td>1305</td>
<td>575</td>
<td>609</td>
<td>-34</td>
<td>-0.11</td>
<td>271</td>
<td>0.90</td>
<td>246</td>
<td>0.82</td>
<td>0.08</td>
<td>1.36</td>
</tr>
<tr>
<td>1310</td>
<td>519</td>
<td>426</td>
<td>93</td>
<td>0.31</td>
<td>219</td>
<td>0.73</td>
<td>182</td>
<td>0.61</td>
<td>0.12</td>
<td>2.51</td>
</tr>
<tr>
<td>1315</td>
<td>356</td>
<td>438</td>
<td>-82</td>
<td>-0.27</td>
<td>159</td>
<td>0.53</td>
<td>158</td>
<td>0.53</td>
<td>0.00</td>
<td>82.00</td>
</tr>
<tr>
<td>1320</td>
<td>469</td>
<td>482</td>
<td>-13</td>
<td>-0.04</td>
<td>215</td>
<td>0.72</td>
<td>170</td>
<td>0.57</td>
<td>0.15</td>
<td>0.29</td>
</tr>
<tr>
<td>1325</td>
<td>463</td>
<td>390</td>
<td>73</td>
<td>0.24</td>
<td>208</td>
<td>0.69</td>
<td>177</td>
<td>0.59</td>
<td>0.10</td>
<td>2.35</td>
</tr>
<tr>
<td>1330</td>
<td>418</td>
<td>371</td>
<td>47</td>
<td>0.16</td>
<td>196</td>
<td>0.65</td>
<td>172</td>
<td>0.57</td>
<td>0.08</td>
<td>1.96</td>
</tr>
</tbody>
</table>

Column 1: Total number of bid shares over the 5-minute trading horizon
Column 2: Total number of ask shares over the 5-minute trading horizon
Column 3: Difference between Column (1) and Column (2)
Column 4: Column (3) ÷ 300 seconds
Column 5: Total number of bid events over the 5-minute trading horizon
Column 6: Column 5 ÷ 300 seconds
Column 7: Total number of ask events over the 5-minute trading horizon
Column 8: Column 7 ÷ 300 seconds
Column 9: Column (6) – Column (8)
Column 10: Column (4) ÷ Column (9), i.e. Little’s Law
The final computed waiting time before any orders are filled is shown in Column (10). It has the minimum value of 0.11 second and the maximum value of 82 seconds with an average of 9.02 seconds. Actually, this should interest many traders, as they are now able to estimate mean waiting time before orders get filled in real time.

Most importantly, however, knowledge of \( L(t) = \lambda(t)W(t) \) can predict either or both the ask and the bid, provided that in equilibrium, where the initial equilibrium price is \( c \),

\[
\frac{\Delta N(t)}{\Delta P} \bigg|_{c,L=0} = -\alpha; \quad \text{and} \quad \frac{\Delta D(t)}{\Delta P} \bigg|_{c,L=0} = \beta
\]

Thus, if the trade price equals the market inside ask,

\[
p_{\text{ask}} = c - \left( \frac{1}{\alpha + \beta} \right)L^*, \text{where } L^* < 0; \quad L^* = \lambda^*W^*
\]

Similarly,

\[
p_{\text{bid}} = c + \left( \frac{1}{\alpha + \beta} \right)L', \text{where } L' > 0; \quad L' = \lambda'W'
\]

Knowledge of the forthcoming market bid and ask should greatly benefit the traders. The limit order vs. market order strategy is as stated earlier.

**Summary and Conclusion**

In order to put these into perspective, we now recall the boundary condition that we established earlier for the bid-ask spread, that is,

\[
\text{(ask - bid)} \geq n + (R - C)
\]

With time delay for order fills, then, we now conjecture that the integer \( n > 1 \). In other words, with delayed fills, the economy’s average number of unfilled outstanding orders must impact the value of \( n \), i.e. \( n = f(L) \). By subtracting \( p_{\text{bid}} \) from \( p_{\text{ask}} \), it is clear that

\[
\text{(ask - bid)} \geq \gamma(L' - L^*) + (C - R)
\]

As a result, the arbitrage opportunity expands.

This paper has studied the optimal limit order trading strategies in the context of the partial equilibrium. If all orders are instantaneously filled without any friction, we first showed that the optimal bid and ask spread is only the size of the tick in the market quote. However, as soon as we introduce a possibility that the order can only be filled with delay, we needed to describe the stochastic process with Poisson with the properties of “orderliness” and “memorylessness,” which appear to illustrate perfectly how the orders arrive and depart in a trading system. We then integrated this Poisson process into the Little’s Law to compute the optimal waiting time before any orders get filled. We hope that the use of the
average unfilled orders in the market was effective. We then further showed how this result can be used to explain the bid and ask spread and to formulate trading strategies.

Reference


