ITG

Let's Play Hide-and-Seek: The Location and Size of Undisclosed Limit Order Volume

STEVE BONGIOVANNI, MILAN BORKOVEC, AND ROBERT D. SINCLAIR

July 9, 2006

Abstract

With the advent of "smart" algorithmic trading systems driven in part by more transparent data offerings from market venues, information leakage from order placement is the nightmare of any market participant. Market participants are camouflaging their intent more than ever by strategic placement of hidden volume throughout the order book via iceberg and/or discretionary limit orders. This paper identifies simple, stylized facts which will allow other market participants to evaluate the likelihood of finding hidden volume. Based on our model, we can predict the hidden volume and also assess the probability that a market order will be executed within the spread and better than the mid-quote. The cost per immediate execution can be better assessed. Moreover, we can identify the delicate balance between submitting a hidden versus visible limit order that an informed trader may use to stimulate the market towards an equilibrium goal without revealing too much information.

STEVE BONGIOVANNI is a senior research analyst at ITG Solutions Network, Inc., 44 Farnsworth Street, Boston MA 02210, Tel: (617) 692-6745; Email: Steve_Bongiovanni@itginc.com

MILAN BORKOVEC is Co-Head of the Financial Engineering Group at ITG Solutions Network, Inc., 44 Farnsworth Street, Boston MA 02210, Tel: (617) 692-6733; E-mail: Milan_Borkovec@itginc.com

ROBERT SINCLAIR is a senior research analyst at ITG Solutions Network, Inc., 44 Farnsworth Street, Boston MA 02210, Tel: (617) 692-6765; E-mail: Robert_Sinclair@itginc.com

The authors wish to thank Ian Domowitz, Henry Yegerman, Yossi Brendes, Vitaly Serbin, Scott McIntire, Zhaoyang Zhao, Gabriel Butler and Hitesh Mittal for their support and comments.

1. OVERVIEW

With the current trend towards more transparency in order driven electronic markets (the latest examples being the new level 2 and real-time data products offered by NASDAQ and NYSE) it is clear that the market is demanding greater transparency and currency of market information. The limit order book offers the ability for market participants to observe the level of market liquidity by seeing the price and quantity of unexecuted limit orders. By utilizing this data, market participants can implement a range of "game theoretical" strategies and choose limit orders with specified price, quantity, and timing, thus allowing them to minimize execution costs and uncertainty, hide market information, and possibly move the market towards the desired price.

Given the concerns associated with information leakage due to order placements, most market venues are allowing market participants to enter limit orders without revealing the full share volume size and/or the associated price level ("iceberg", "undisclosed", or "discretionary" limit orders). This brings with it the complex interrelationship between exposure risk (adverse selection), market liquidity, and the need for transparency.

From a market design point of view, hidden limit orders represent a real trade-off between liquidity and transparency. Trading systems need to attract liquidity and trading activity. The availability of hidden limit orders encourages limit order traders to supply liquidity when they might be hesitant to fully disclose their trading interests thus increasing the liquidity on the system. However, hidden limit orders by their nature do not add information to the market and thus do not help in the market's transparency. Specifically, hidden orders inside the spread will not attract activity to a venue, since most order routing systems can only operate on visible information. So, hidden limit orders clearly diminish the supposed benefits of transparent order driven markets: price efficiency, low costs of market monitoring and less information asymmetries (Madhavan [2000]).

The concept of hiding transaction fingerprints has been around for several years but has recently seen increased popularity due to the advent of algorithmic trading systems such as ITG's "Dark Server^SM" or CSFB's "Guerilla" which utilize continuous mid-point crosses from Dark Books. For illiquid stocks, which have larger intra-day volatility, the concept of hiding allows the market participant to be able to transact with minimum market impact.

Hidden limit orders have become an important limit order type. Hasbrouck and Saar [2002] show that hidden orders account for more than 12% of all orders executed on Island, and Tuttle [2005] reports that hidden liquidity represents 20% of the inside depth in the Nasdaq 100 stocks. D'Hondt, De Winne, and Francois-Heude [2004] find that hidden depth on Euronext Paris accounts for 45% of the total depth available at the best five quotes and 55% of the total depth at the best limits.

These findings suggest that there are underlying factors that cause a market participant to use a hidden versus a visible limit order, considering the controversial rationality behind using hidden limit orders. Consistent with previous literature, there are two main beliefs for the existence of hidden limit orders. The first argument is that hidden limit orders are used by large liquidity traders to reduce their exposure risk by hiding their intent to trade. In other words, liquidity traders use hidden limit orders as a self-protective strategy against other more informed traders. The other argument is that hidden limit orders are mostly submitted by informed traders to conceal their insider information. By placing (aggressive) hidden limit orders, market participants with insider information can trade fast and almost unobserved. Therefore, informed traders may prefer using undisclosed versus displayed limit orders for certain market conditions.

Taking into account undisclosed limit orders can dramatically change the picture of the limit order book at any given time of the day. From Figure 1, it can be easily concluded that if we wanted to instantaneously execute a buy market order for 1000 shares of company Argonaut Group Inc., the cost associated with that trade (benchmarked on the existing mid-quote) would be \$0.05 per share. However, if we were able to reconstruct the book in a way that included the hidden shares using information from the prevailing market conditions, we would then see that the "true" cost for the 1000 shares is actually only \$0.045 per share (see Section 4 for more details).

A trader seeing the "true" limit order book instead of Figure 1 might be willing to consider the opportunity cost relative to the market dynamics associated with removing only a portion of the desired volume from within the spread which leads to improvement in per share transaction cost.[1] This trader would also be able to evaluate the probability that an order is filled within or below the existing visible best ask price.

Our main goals in this paper are to estimate the probability of the existence of hidden volume, to calculate this volume between the best bid and ask, and to predict the actual location (price) of this hidden volume. With this information, we are able to construct and display a complete limit order book, which includes the expected hidden volume at the appropriate price levels.

Reconstructing the limit order book around the best levels allows us to measure the "true" execution prices if we place market orders or marketable limit orders. The possibility of getting better than expected execution prices has two main implications:

- quantifying best execution (ignoring execution improvement can be misleading when comparing execution quality across markets with significant non-displayed additional liquidity)
- undisclosed volume is an integral part of the pricing process. We will see that price improvement is the largest when spreads are narrow and the volatility is large. We are able to understand the traders' behavior relative to the additional information in the limit order book. We believe that the findings are of great interest not only in terms of modeling pure order driven markets and characterizing traders' behavior, but also in giving us an advantage as it relates to implementing an automatic search (liquidity and asymmetric information) algorithm. This research sheds some light on some basic questions of pre-trade transparency and the challenges faced when developing better trading algorithms

(Yang and Borkovec [2005], Domowitz and Yegerman [2005a]) to improve trading performance (Yang and Jiu [2006], Domowitz and Yegerman [2005b]).[2]

This paper is organized as follows. The data used is described in Section 2 and in Section 3, we highlight the static empirical evidence associated with hidden volume and its placement. In Section 4, we reconstruct the limit order book and highlight possible applications. The paper concludes in Section 5 where it also identifies future research.

2. The Data

Our research data includes three months of Comstock level 2 (L2) real time information from ARCA. We use two months data from June to July 2005 to estimate our empirical models and use August 2005 for out-of-sample testing. With the same information available to any market participant using the limit order book, we receive messages that reflect additions, cancellations, and modifications of all orders on the book. To match trades with limit orders, we examine L2 messages within a 2-second time bandwidth to find the order cancellation or modification message which corresponds in price, size, and exchange to a particular trade. If a trade is matched to an order message, the side classification of this trade is obtained from the message; this trade would be classified as being visible. If there is more than one match, we assume that the correct match is the one which is closest in absolute time difference between the time stamp and the message time. If a trade couldn't be matched with a limit order message, then the trade is classified as a *hidden* trade (coming from a hidden/undisclosed limit order). To determine the side of the hidden trade we use a generalization of the Lee and Ready [1991] algorithm.[3]

Table 1 describes the tickers used in our model, based on market capitalization. More than 78% of the tickers chosen belong to small cap stocks, 13% belong to median cap stocks, and the remaining 9% belong to large cap stocks[4]. For each large cap stock, the average number of trades per day is 7,900 with an average trade size of 920 shares. As for each small cap stock, the average number of trades per day is approximately 280 trades with an average trade size of 520 shares. Of the trading activities for the small cap stocks, 28% of all traded volume is classified as hidden while the number is only 21% for the large cap stocks. Approximately 96% of all orders added to the book are eventually cancelled. Of the cancelled orders, approximately 10% can be classified as fleeting orders[5]. The data also shows that on average, hidden orders have a larger size in comparison to orders that are fully displayed. This result is consistent with Harris' [1998] findings that traders often restrict displayed orders, especially for orders with larger expected remainders.

Given that our model is trade-based in nature, we think that classifying stocks by market capitalization is inadequate since stocks within the same market cap group can differ significantly in trade volumes. Therefore, instead of the commonly used market capitalization, we group stocks based on their 21-day median trade share volume; this allows us to classify stocks with similar trade volume within the same group. To get a representative sample of tickers across the universe, we rank all available tickers (approximately 7,000) according to their 21-day median trade volume at the beginning of our sample period and then divide this universe into eleven liquidity groups with *Liquidity Group 0* representing the least liquid stocks and *Liquidity Group 10* representing the most liquid stocks. For each of the eleven liquidity groups, we then randomly select a sample of tickers to use in our pooled data model. Loosely speaking, small cap stocks belong to Liquidity Group 0 to Liquidity Group 7, mid cap stocks belong to Liquidity Group 8 to Liquidity Group 9, and large cap stocks belong to Liquidity Group 10.

In order to justify our grouping, we examine order placement in each liquidity group to see if there is a clear difference in how limit orders are placed across different liquidity groups. We classify limit order placement into three categories: (1) *AT,* which represents limit orders being placed at the best level, (2) *BETTER,* which represents limit orders being placed between best bid and ask, and (3) *AWAY,* which represents orders being placed at prices worse than the best levels. From Figure 2 we can easily see that the placement pattern is not similar across any liquidity group. For the lowest liquidity group, more than 28% of all new limit orders were placed *AT* the best bid and ask level while for the most liquid group, 48% of all new limit orders were placed *AT* the best bid and ask level. The difference in percentage reflects the differences in share trade volume, urgency to get order completed, and the competition within the liquidity group.

In this *pooled data analysis,*[6] we look for specific factors that we expect to affect the probability of hidden order placement. One intuitive hypothesis is that hidden orders are more frequently used for stocks with a high exposure risk (Harris [1998]). We speculate that in a market with low volatility, hidden orders reduce the chances of being front-run and thus we expect volatility to play an active role in our analysis. In markets that enforce time precedence, front-running can be very expensive. Since front-running is expected to be more expensive for stocks with relatively low prices, we also expect the use of hidden orders to be higher for those stocks. As for uninformed traders, the option value of limit orders is affected by factors like volatility [*Mid-Quote Volatility*], trading activity [*Depth Size, Time Since Last Trade, and Spread*] and time to total (partial) execution. We also speculate that the order exposure risk is related to the expected time for an order to be (totally) executed and that the frequency of orders that are partially displayed is related to the trading frequency of a stock.

The next concern is associated with the time of the day. Are there privileged periods over the trading day to enter hidden orders on the market? Do we see market participants placing limit orders at specific periods of the day?[7] We divide the trading day into thirteen 30-minute time bins and examine the order placement pattern. Clearly from Figure 3, the time of day seems to explain where an order might be placed. At the opening of the market, with no real information, a market participant might be equally likely (33%) to place a limit order *BETTER*, *AT*, or *AWAY*. As the day progresses, we see that by 3:30-4:00pm, the probability of a market participant placing a limit order within the best bid and ask drops to 18%. When we look at the average number of limit orders per ticker placed throughout the day, we note that most limit orders are placed at the first 2.5 hours of the trading day. This pattern is consistent across all liquidity groups.

When we examine the pattern associated with the number of the orders placed based on the time period of the day, it seems to mimic the spread curve. This suggests that the time bin might not be a factor associated with limit order placement and that what we are observing is really limit order placement relative to the spread. The spread also captures the market impatience and is possibly the first hint that there might be asymmetric information among the market players[8].

Glosten and Milgrom [1985] is among the many papers that identify that information asymmetries among investors influences the bid-ask spread. Large spread would seem to suggest that there is little or no market information or activity. If we examine the commonly known spread profile, the spread is on average largest at the opening of the market (the information searching period) and, as the day progresses and information is captured through the market transactions, the spread declines and reaches its lowest level by the end of the trading day.

A good hidden volume predictor must potentially take all of the previously discussed variables into consideration. Some of these variables describe the stock price dynamics, while others describe the "fundamental" or historical characteristics of the stock. The next section will discuss in more detail our model and its associated input variables that have been found meaningful.

3. Model

3.1 Size of Undisclosed Limit Order Volume

In this section, we describe the model and look at the empirical results associated with estimating the size of the hidden volume and its location (placement) between the spread. To achieve this goal, we identify (a) all trades that have been executed through undisclosed limit orders and (b) their associated market conditions. Given that we want to model the discrete choice of placing a hidden versus a visible limit order, we use a probability regression model that maps trade volume with market conditions. We consider different trading horizons (trading instantaneously, or within a 1-, 2-, 3-, 4-, or 5-minute period). A regression model that only uses the hidden trade volume that is actually executed would produce estimators which are biased downward. To correct that aspect, we specify necessary censoring conditions. To evaluate the model's strength and to identify a few stylized facts, (1) we examine the prior belief and match it with the empirical results to see if these results are consistent across all liquidity groups, and (2) we estimate McFadden's LRI to approximate a pseudo R^2 for assessing the goodness of fit.[9]

Based on our research, we identify a few stylized facts that relate to modeling hidden volume. The effective spread and volatility measures capture the level of front-running and any abnormal market movement which could be associated with asymmetric information, herding, market corrections, or short-term movements. Less than normal effective spread indicates that many market participants are front-running and hence, to camouflage some of the liquidity demand, the hidden volume would be greater. As for volatility, high volatility reflects the market uncertainty and the possibility of hidden volume being executed away from the mid-quote. With high volatility levels, we expect that a market participant will place more hidden volume, since the probability for being executed increases and no information or strategy is revealed to the market. We expect the coefficients associated with effective spread to be negative and the coefficients associated with volatility to be positive. This is consistent with our findings in Table 2.

We know that market participants are monitoring the changes in the shape of the limit order book. We expect that they are tracking order additions, cancellations, depth, previous 15-seconds mid-quote returns, and the misalignment of the mid-quote associated relative to the composite market. These variables act as the frontline variables to capturing market dynamics and participants' gaming/strategy. Our results suggest that more additions than cancellations of limit orders is a signal that there are players in the market that hope that such actions stimulate the market, perhaps to attract the market towards their undisclosed volume.

When the mid-quote is misaligned, we observe that if the ECN's mid-quote is less than the composite mid-quote price, the expected buy hidden limit order volume will be less than that of an ECN which has a mid-quote that is equal or even greater that the composite mid-quote price. In other words, we expect that hidden buy (sell) limit order volume follows the ECN with the highest (lowest) mid-quote price.

Apart from these variables, we examine the previous hidden volume executed. At first glance, one might dismiss this variable as being invisible, and hence not a reliable explanatory variable. However, this would be mixing the concept of "hidden" with that of "invisible" orders. After an execution against hidden volume takes place, there is a telltale trade tick which is printed. Our research indicates that if we find hidden volume, there is a good chance that there will be more hidden volume in the near future. Some of the results of this research are presented in Table 2.

3.2 Location of Undisclosed Limit Order Volume

In the previous section, we modeled the size of hidden volume that is assumed to be located between the best bid and ask. In this section, we go further by estimating where this volume might be located between the best bid and ask. To achieve this goal, we divide the spread into uniformly spaced regions and use the explanatory variables to estimate the probability that an order is placed in that region.

Our premise is that market participants observe the market conditions and from that, decide where to place hidden volume. Hence, changes in the state of the limit order book cause participants to reevaluate their placement strategies. We assume that the "fundamental" factors, absolute historical spread (in cents), and volatility contribute to identifying where hidden orders are being placed on the book. Our assumption is that placement is based on perceived market conditions such as absolute spread and volatility. This model maps the location of a hidden order with existing market conditions. Table 3 gives a brief relationship between a few market variables and the placement of hidden volume.

If we examine stocks in Liquidity Group 8 and divide the bid-ask spread into six equally sized groups, then Figure 4 illustrates how the placement of hidden volume changes with the spread where Region 1 includes the best ask price and Region 6 includes the Region prior to the best bid price. As spread increases beyond its normal levels, hidden volume is more likely to be redistributed within the spread. This pattern holds true across all liquidity groups.

3.3 Testing the Model

To examine the strength of the model in predicting the hidden volume and its placement on the order book, we aggregate all (partial) executions that have the same sign (buy or sell) and occur around the same time on ARCA. For each of these trade clusters, we identify and save the existing market conditions and calculate the cluster volume, the share-weighted average execution price, and the average execution price that is derived from the observed unadjusted limit order book at the start of the cluster execution. The difference between the limit order book's perceived execution price v_i and the actual execution price p_i is what we refer to as the *Virtual Price Error* $v_i - p_i$. *Virtual Price Errors* are typically positive and give a first impression about the usability of the displayed limit order book alone. Next, we reconstruct the limit order book based on the prevailing market conditions using the models that have been discussed in Section 3.1 and 3.2 and we include the estimated undisclosed limit order volume at the appropriate price levels. Based on our adjusted limit order book, we then determine the estimated execution price v_i^{new} and the *New Virtual Price Error* $v_i^{new} - p_i$.

The above procedure enables us to: (a) have a one-to-one comparison between actual and estimated prices, and (b) evaluate the superiority of the adjusted limit order book over the unadjusted limit order book. To assess accuracy, we study varying scenarios for each liquidity group, considering different Time Of Day and different levels of Volatility, Spread, or Limit Order Volume Activity. More precisely, for each liquidity group and scenario we compute the average Virtual Price Error and average New Virtual Price Errors for all subgroups of each scenario. Figure 5 shows the graphical comparison between the average *Virtual Price Error* and average *New Virtual Price Error* for all stocks that have been classified in Liquidity Group 2. The trading day is broken-up into seventy-eight 5-minute bins and we present the average *Virtual Price Error* and average *New Virtual Price Error* in each of these subgroups. In order to capture the *Virtual Price Error* distribution, we also graph the 5% and 95% levels. We see that on average, there is an error of approximately a 5-cent difference between what a trader would believe to be the instantaneous trade execution price if he only looks at the displayed limit order book, and the actual execution price. That bias is eliminated when we use our model to adjust for the hidden volume. Similar results hold for all other scenarios that we have examined which, for brevity, we omit here.

4. Example of Short-term Price Impact using the Reconstructed Book

In what follows, we illustrate and discuss the practical importance of our model described in the previous section. We create a static example associated with placing a market order and we examine both the cost and the price impact associated with executing the order instantaneously and within a 2- and 5-minute bin period.

The price impact PI_i of a market order i is defined as the difference between the last execution price p_i^{final} and the mid-quote m_i immediately prior to the market order i. Similarly, the cost C_i of a market order i is defined as the difference between the share-weighted average execution price $\overline{p}_i$ and the mid-quote m_i immediately prior to the market order i. More precisely,

$$PI_i = \delta_i \cdot \left(p_i^{final} - m_i \right) \text{ and } C_i = \delta_i \cdot \left(\overline{p}_i - m_i \right),$$

where $\delta_i = 1$ for a buy market order and $\delta_i = -1$ otherwise.

From Figure 1, executing a buy market order of 1000 shares of the company **Argonaut Group Inc.**[9] would have a price impact of \$0.10 and the cost per share would be \$0.05. For illustration purposes, we assume that at 10:40 am, Argonaut Group Inc is very actively traded with an effective spread of one deviation less than average and volatility being one half of a deviation higher than normal. Moreover, we assume that in the previous 5-minute bin, of the 1000 shares traded we classify 30% as being hidden shares. From the specified market conditions, we estimate that there will be around 60 hidden shares instantaneously available between the best bid and ask and 147 and 280 hidden shares of sell limit orders within the next 2 and 5 minutes, respectively. Table 4 gives the complete breakdown.

Next, we look at where we can locate the estimated hidden sell limit order volume, that is, at which price level we could expect the undisclosed sell limit orders. Using the probability model discussed in Section 3.2, we estimate the probability associated with each price level between the best bid and ask. Given the market conditions for Argonaut Group Inc. at 10:40am, Table 4 shows that there is approximately 4.6% of hidden sell limit order volume located at the price level $35.00 and that 15.45% of total hidden sell limit order volume is available at or below $35.02. As for the best ask price ($35.05), we find that there is approximately 74% of the total hidden sell limit order volume located at that level. If we reconstruct the limit order book to include the hidden volume (we multiply the probabilities by the total hidden sell limit order volume), the price per share traded for the instantaneous trade model (with the assumption of locating 59 shares) is $35.065 and the cost is $0.045, which is lower than the estimated cost for the unadjusted book. The price impact remains as $0.10. Figure 6 gives a complete graphical representation of the trading horizon, the hidden sell limit order volume, and the average execution price. Some of the results are presented in Table 4.

Figure 6 shows the average execution price based on the unadjusted (static) limit order book in comparison to the adjusted book that takes into account trading instantaneously and with 1-, 2-, 3-, 4-, and 5-minute horizon. As the trading horizon increases, and the market players execute against the estimated volume within the next 5 minutes, the average expected execution price falls to $35.052 and the expect price impact decreases to $0.05.

5. Conclusion

While market venues strive to achieve greater transparency by offering market data products with more granular and current information, market participants, in their demand for minimal information leakage, continue to hide their trading intent by placing hidden limit orders. This conflict between market transparency and traders' secrecy complicates the task of algorithmic trading systems of seeking out both liquidity and best execution, and it complicates the tasks of the market participant by obscuring the true liquidity of the market.

The rising popularity of placing undisclosed limit orders instead of displayed limit orders, has greatly limited the usefulness of the limit order book when it comes to the transparency of market participants' actions. We have shown that using a "simple" limit order book is insufficient for estimating true liquidity and transaction costs. Furthermore, we note that utilizing an analysis of the "simple" limit order book and ignoring the undisclosed limit order volume actually alters the execution optimization and the transaction and opportunity cost reduction, with a bias towards lower available liquidity and higher transaction costs.

It can easily be inferred from our results that ignoring the probability of undisclosed volume within the spread greatly limits (at best) the ability of algorithmic trading systems and smart order routing systems to find the best execution price for market orders. Algorithmic trading systems must either uniformly search across different market venues (at a great opportunity cost) or devise smarter ways to seek out available liquidity. As for 'smart' order routing systems that do not take into account the probability of undisclosed volume within the spread - well, they aren't.

This paper highlights a few stylized facts that provide market participants an easy way to monitor the limit order book for liquidity. In summary we see, in general, that the Effective Spread is negatively correlated to the hidden liquidity and that the Mid-Quote Volatility, Additions Between Best Bid And Ask, and the Additions Less Cancellations are all positively correlated to hidden volume. Although these facts do not permit a reconstruction of the "real" book they are nonetheless useful in getting at least a sense of the overall hidden liquidity, if not also its size and location.

The model discussed in this paper goes further than the stylized facts. It could be used to enhance the existing picture of the limit order book by synthesizing the "real" limit order book based on the probability of hidden volume, including its location (in the book), size, and venue. The major implications of utilizing this extra information are twofold and apply both to automated trading systems and market participants' strategies. We offer some simple examples here. For algorithmic trading systems and smart routing systems it would be obvious that given equal price and liquidity across multiple venues, the best choice of which venue to route to would be the one with the highest probability of undisclosed volume within the spread; we would consider these systems to be "smart". For market participants, this enhanced book of information can be used to guide a liquidity trader to the best sources of liquidity and it could potentially be used to drive enhanced limit order trading models. By utilizing this enhanced book in choosing to place a limit order, and in selecting its price, size, time, venue, and whether or not it should be displayed, the market participant has a more realistic view of how others will respond to his action.

So, in the game of market liquidity hide-and-seek, we are sure that algorithmic trading systems and market participants that ignore hidden order liquidity are fundamentally handicapped and are stumbling around in the dark, when compared to systems and players that utilize the spotlight that hidden volume models shine on the limit order book.

6. References

BROCKMAN, P., AND D. Y. CHUNG, 1999, "Bid-ask spread component in an order-driven environment", *The Journal of Financial Research*, 22, 2, 227-246.

D'HONDT, CATHERINE, DE WINNE, RUDY AND ALAIN FRANCOIS-HEUDE, 2004, "Hidden Orders on Euronext: Nothing is Quite as it Seems" (May 2004). Available at SSRN: http://ssrn.com/abstract=379362.

DOMOWITZ, IAN, AND HENRY YEGERMAN, 2005a, "The Cost of Algorithmic Trading: A First Look at Comparative Performance", In Brian Bruce, ed., *Algorithmic Trading: Precision, Control, Execution*, New York: Institutional Investor

DOMOWITZ, IAN, AND HENRY YEGERMAN, 2005b, "Measuring and Interpreting the Performance of Broker Algorithms", *ITG Inc. Research Report*

GEORGE, T., G. KAUL, AND M. NIMALENDRAN, 1991, "Estimating the components of the bid-ask spread: A new approach", *Review of Financial Studies*, 4, 623-656.

GLOSTEN, L., AND L. HARRIS, 1988, "Estimating the components of the bid-ask spread", *Journal of Financial Economics*, 21, 123-142.

GLOSTEN, L., 1987, "Components of the bid-ask spread and the statistical. properties of transaction prices", *Journal of Finance*, 42, 1293-1307.

GLOSTEN, L., AND PAUL MILGROM, 1985, "Bid, Ask and Transaction Price in a specialist market with heterogeneously informed investor", *Journal of Financial Economics*, 14 (10), 71-100.

GREENE, WILLIAM, 2000, *Econometrics Analysis*, 4th Edition. Prentice-Hall, Inc (New Jersey)

HASBROUCK, JOEL, AND GIDEON SAAR, 2002, "Limit Order and Volatility in a Hybrid Market: The Island ECN", working paper

HASBROUCK, JOEL, AND GIDEON SAAR, 2004,, "Technology and Liquidity Provision: The Blurring of Traditional Definitions", working paper, New York University.

HARRIS, LAWRENCE, 1998, "Optimal Dynamic Order Submission Strategies in Some Stylized Trading Problems," Financial Markets, Institutions & Instruments.

LEE, CHARLES M.C., AND READY MARK J., 1991, "Inferring Trade Direction from Intraday Data", *The Journal of Finance,* 46(2), 733-746.

MADHAVAN, ANANTH, 2000, "Market microstructure: a survey", *Journal of Financial Markets*, 3, 205-258.

PARDO, ANGELO, AND ROBERT PASCUAL, 2004, "On the hidden side of liquidity", *Working Paper Series 2004*, FIN-04-004, Leonard N. Stern School of Business, New York University.

PASCUAL, ROBERTO AND DAVID VEREDAS, 2004, "What Pieces of Limit Order Book Information Matter in Explaining the Behavior of Limit-Order Traders and Market-Order Traders?" working paper. Available at SSRN: http://ssrn.com/abstract=489508

Tuttle, Laura. "Hidden Orders, Trading Costs and Information". working paper, University of Kansas, 2005

YANG, JIAN, AND MILAN BORKOVEC, 2005, "Algorithm Trading: Opportunities and Challenges", *Financial Engineering News,* 46, 14-15.

YANG, JIAN, AND BRET JIU, 2006, "Algorithm Selection: A Quantitative Approach", In Brian Bruce, ed., *Algorithmic Trading II*, New York: Institutional Investor

END NOTES

(1) We found that most of the explanatory power of the book is concentrated within the dynamics associated with the visible best quotes. This finding is consistent with Pascual and Veredas [2004].

(2) Madhavan [2000, page 234] defines pre-trade transparency as "the wide dissemination of current bid and ask quotations, depths, and possibly also information such as the existence of large order imbalance."

(3) To establish a level of reliability for our side classification algorithm, we test our algorithm against proprietary ITG execution data where we know the side. Our matching algorithm was approximately 90-95% accurate.

(4) We define a stock as being small cap if its market capitalization is less than $1.5 billion. If the market capitalization was greater than $1.5 billion but less than $10 billion, we consider those stocks as median cap stocks. All other stocks are classified as large cap stocks.

(5) Fleeting orders are orders which are added and cancelled from the book within less or equal 2 seconds. The time stamp on each order is defined in seconds (Hasbrouck and Saar [2004]).

(6) The pooled data consists of more than 325 tickers viewed across two months of data.

(7) The model could be extended to capture the anomalies associated with days of the week and month of the year. We could extend the model to take care of timing associated with rebalancing portfolios.

(8) Glosten [1987], Glosten and Harris [1988], George, Kaul, and Nimalendran [1991], Brockman and Chung [1999] consider decomposition of bid-ask spread.

(9) See Greene [2000] for examples of this proxy for R^2.

(10) Argaunot Group Inc. has been selected for illustrative purposes at random.

TABLE 1

CHARACTERISTIC (DAILY AVERAGE)	LARGE CAP STOCKS (CAP > $10 BILLION)	MID CAP STOCKS ($10 BILLION <CAP<$1.5 BILLION)	SMALL CAP STOCKS (CAP<$1.5 BILLION)
TRADES:			
Number of Trades	207,000	87,800	72,000
Size of Trade (Visible)	370	300	200
Size of Trade (Hidden)	550	450	320
Percentage of Trade Hidden	21%	23%	28%
FLEETING/CANCELLATION:			
Number of Fleeting Orders	328,500	126,800	160,100
Fleeting Orders/Total Cancelled Orders	11%	9%	13%
Cancelled Orders/Total Added Orders	96%	96%	97%
Number of Stocks	26	43	260

Table 1 represents data from Comstock's ARCA data feed on 329 tickers from June-August 2005. The tickers are divided into three market capitalization categories where the small capitalization category represents tickers with market capitalizations less than $1.5 billion. If the market capitalizations of stocks were greater than $1.5 billion but less than $10 billion, the tickers are classified as mid capitalization stocks. All other tickers are classified as large cap stocks. Fleeting orders are orders which are added and cancelled from the limit order book within a two seconds period. Hidden trades are trades which can't be matched to a limit order execution.

TABLE 2

			Mid-Quote		Standardized			
Liquidity	**Goodness of Fit R^2**	**Lagged Hidden Volume**	**-1**	**0**	**Addition less Cancellation**	**Addition between and at the best Bid and Ask**	**Mid-Quote Volatility**	**Effective Spread**
0-2	0.05	**0.1807** (0.0151)	**-2.9563** (0.2336)	**-1.2840** (0.1650)	**0.2043** (0.0523)	**0.2155** (0.0588)	**0.0205**^{ns} (0.0763)	**-0.0064**^{ns} (0.0783)
3	0.05	**0.1233** (0.0112)	**-0.9911** (0.1329)	**-2.9905** (0.1888)	**0.4151** (0.0386)	**0.2096** (0.0425)	**0.1557** (0.0541)	**-0.1923** (0.0553)
4	0.07	**0.1072** (0.0077)	**-0.9270** (0.0585)	**-0.3364** (0.0427)	**0.1302** (0.0113)	**0.1057** (0.0116)	**0.1059** (0.0172)	**-0.0477** (0.0150)
5	0.08	**0.1746** (0.0063)	**-0.5122** (0.0316)	**-0.2447** (0.0063)	**0.0857** (0.0055)	**0.0464** (0.0055)	**0.0420** (0.0076)	**-0.0613** (0.0060)
6	0.08	**0.2037** (0.0043)	**-0.2909** (0.0171)	**-0.0985** (0.0120)	**0.0696** (0.0029)	**0.0429** (0.0030)	**0.0424** (0.0019)	**-0.0323** (0.0030)
7	0.08	**0.1387** (0.0047)	**-0.2429** (0.0144)	**-0.1146** (0.0098)	**0.0350** (0.0022)	**0.0326** (0.0023)	**0.0328** (0.0013)	**-0.0236** (0.0021)
8	0.08	**0.1633** (0.0039)	**-0.1931** (0.0134)	**-0.0872** (0.0091)	**0.0235** (0.0015)	**0.0294** (0.0016)	**0.0179** (0.0022)	**-0.0172** (0.0013)
9-10	0.06	**0.3164** (0.0033)	**-0.0868** (0.0097)	**-0.0145** (0.0057)	**0.0119** (0.0008)	**0.0128** (0.0010)	**0.0091** (0.0012)	**-0.0070** (0.0006)

Table 2 reports a subset of the variables used in the model for predicting hidden sell limit order volume. The numbers in parentheses are the standard errors for the parameters. The coefficient values with the superscript *ns* indicate that these numbers are not significant at the 95% confidence level. For the variable MID-QUOTE, we assign a 1 if ARCA's mid-quote is greater than that of the composite mid-quote, we assign a 0 if the mid-quote is equal, and -1 otherwise. All variables are standardized by their corresponding historical 3 months means and standard deviations, i.e. $X_{(\text{standard})} = \frac{x - \bar{x}}{\sigma(x)}$ where $\bar{x}$ is the mean and $\sigma(x)$ is the standard deviation of x.

TABLE 3

			Standardized		
Liquidity	**Goodness of Fit R^2**	**Return in last 15-seconds**	**Addition less Cancellation**	**Addition between and at the best Bid and Ask**	**Imbalance in Depth**
0-2	0.18	**-0.0425**^{ns} (0.0373)	**-0.1120** (0.0460)	**0.2322** (0.1126)	**-0.0878**^{ns} (0.0673)
3	0.19	**-0.0854** (0.0430)	**-0.0668** (0.0258)	**0.1050**^{ns} (0.0707)	**-0.0923** (0.0404)
4	0.19	**-0.3023** (0.0449)	**-0.0719** (0.0194)	**0.1584** (0.0476)	**-0.1864** (0.0306)
5	0.20	**-0.2860** (0.0590)	**-0.1125** (0.0160)	**0.0274** (0.0516)	**-0.0172**^{ns} (0.0254)
6	0.20	**-0.3757** (0.0502)	**-0.0894** (0.0104)	**0.1501** (0.0323)	**-0.0791** (0.0107)
7	0.20	**-0.2443** (0.0571)	**-0.0931** (0.0108)	**0.1194** (0.0352)	**-0.0964** (0.0187)
8	0.23	**-0.6894** (0.0569)	**-0.1184** (0.00781)	**0.2262** (0.0258)	**-0.0266** (0.0130)
9-10	0.19	**-0.7267** (0.0589)	**-0.0963** (0.00444)	**0.1192** (0.0141)	**-0.0334** (0.0335)

Table 3 reports a subset of the variables used in the model for predicting the location of sell limit order volume. The numbers in parentheses are the standard errors for the parameters. The coefficient values with the superscript *ns* indicate that these numbers are not significant at the 95% confidence level. All variables are standardized by their corresponding historical 3 months means and standard deviations, i.e. $X_{(standard)} = \frac{x - \bar{x}}{\sigma(x)}$ where $\bar{x}$ is the mean and $\sigma(x)$ is the standard deviation of *x*. The variable "Return in last 15-second" is the time weighted percentage mid-quote return within the previous 15 seconds prior to execution.

TABLE 4

PRICE	INSTANTANEOUS	2-MINUTES	5-MINUTES	PROBABILITY
$35.00	3	7	13	0.046319
$35.01	2	4	8	0.027683
$35.02	5	12	23	0.080532
$35.03	6	15	29	0.105157
$35.05	43	109	207	0.740309

Given the existing trading conditions for Argonaut Group Inc at 10:40am, Table 4 presents the amount of hidden sell limit order volume and its location for different execution horizons. The estimated hidden volume is expected to be available for a market order being executed instantaneously, within 2, or 5 minutes. If there is hidden volume, column "PROBABILITY" shows the likelihood that hidden volume is located at that price.

FIGURE 1
Limit Order Book: Argonaut Group Inc. (AGII)

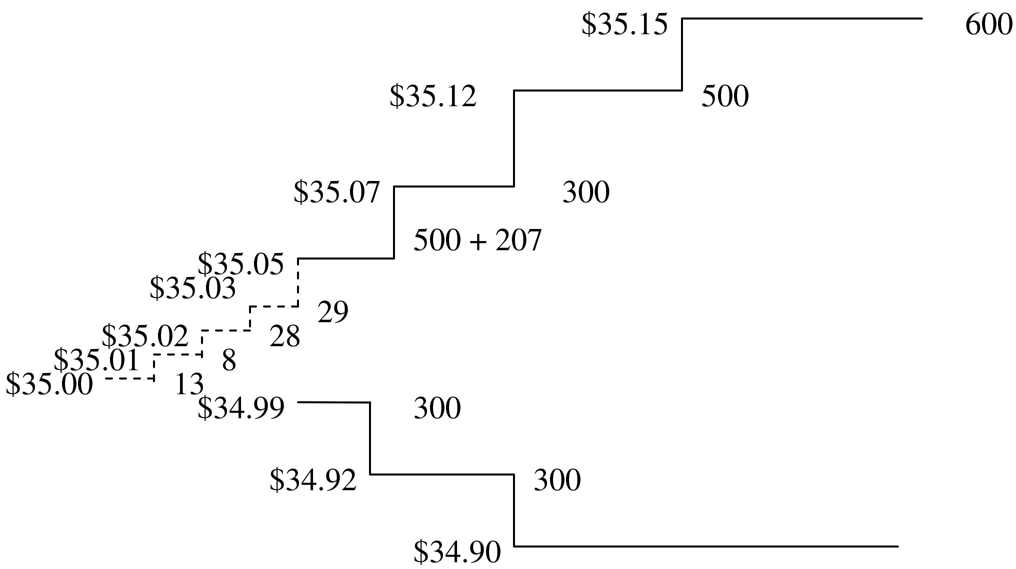

Figure 1 shows a simple example of the limit order book for AGII at 10:40am which incorporates the hidden sell limit order volume associated with executing a market buy order. The solid lines represent the disclosed liquidity while the dotted lines are associated with the price and undisclosed liquidity.

FIGURE 2

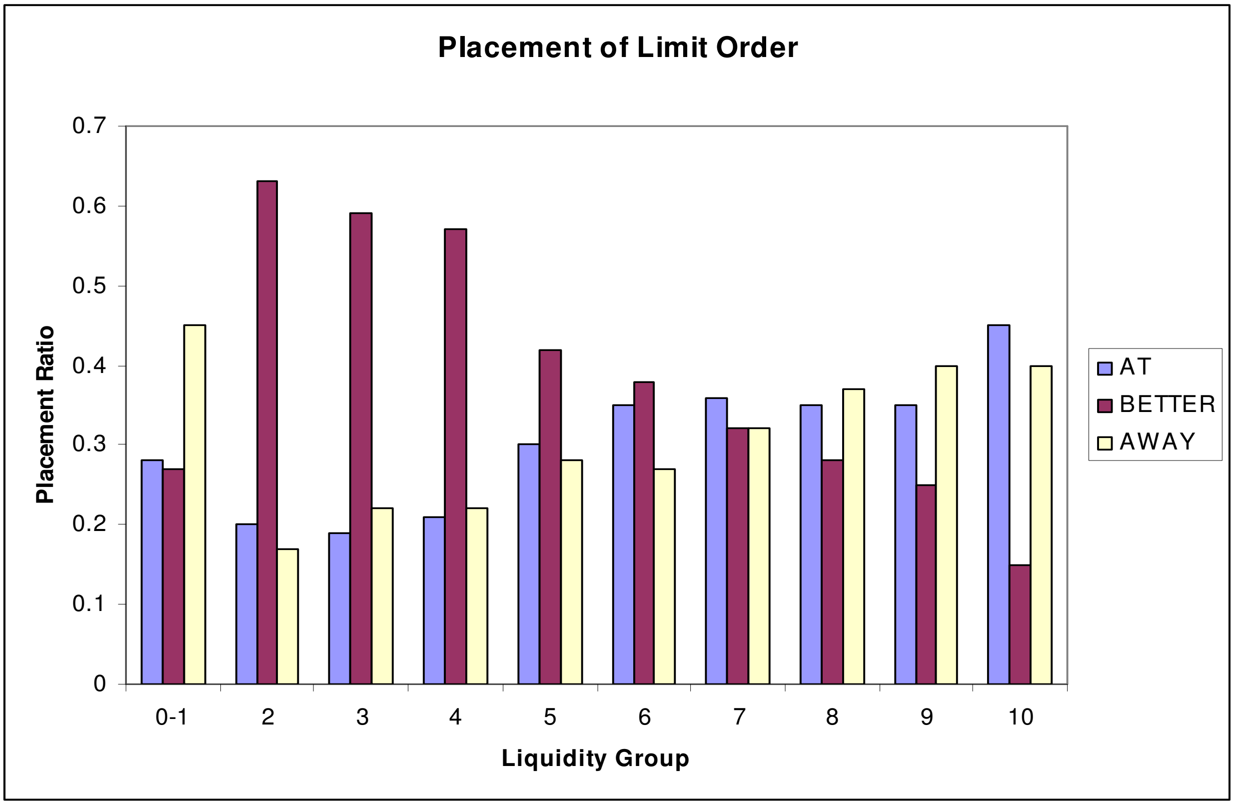

Figure 2 maps placement probabilities against the 21-days median trade share volume where *Liquidity Group 0* represents the least liquid stocks and *Liquidity Group 10* represents the most liquid stocks. We classify limit order placements into three categories: (1) *AT,* which represents limit orders being placed at the best level, (2) *BETTER,* which represents limit orders being placed between best bid and ask, and (3) *AWAY,* which represents orders being placed at prices worse than the best levels.

FIGURE 3

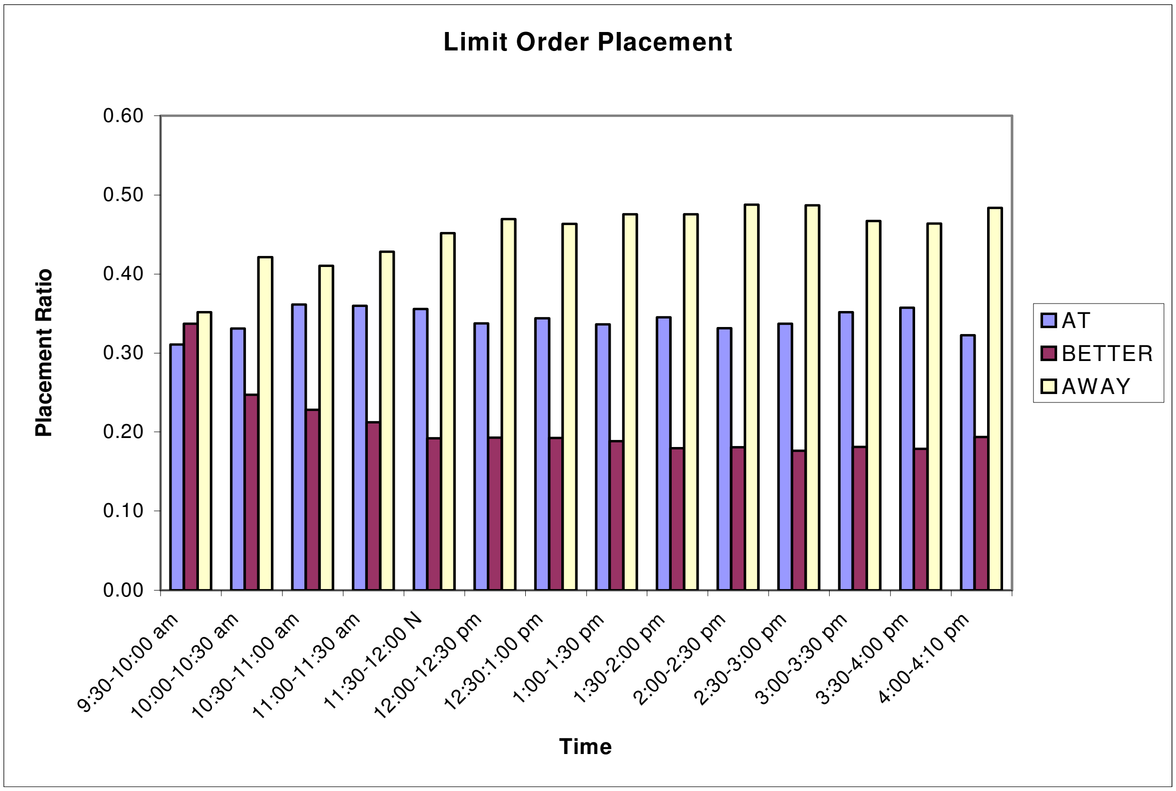

Figure 3 maps placement probabilities against the time of the day. We classify limit order placements into three categories: (1) *AT,* which represents limit orders being placed at the best level, (2) *BETTER,* which represents limit orders being placed between best bid and ask, and (3) *AWAY,* which represents orders being placed at prices worse than the best levels. Although the chart does not distinguish among different liquidity group, this placement pattern is relatively consistent across different groups.

FIGURE 4

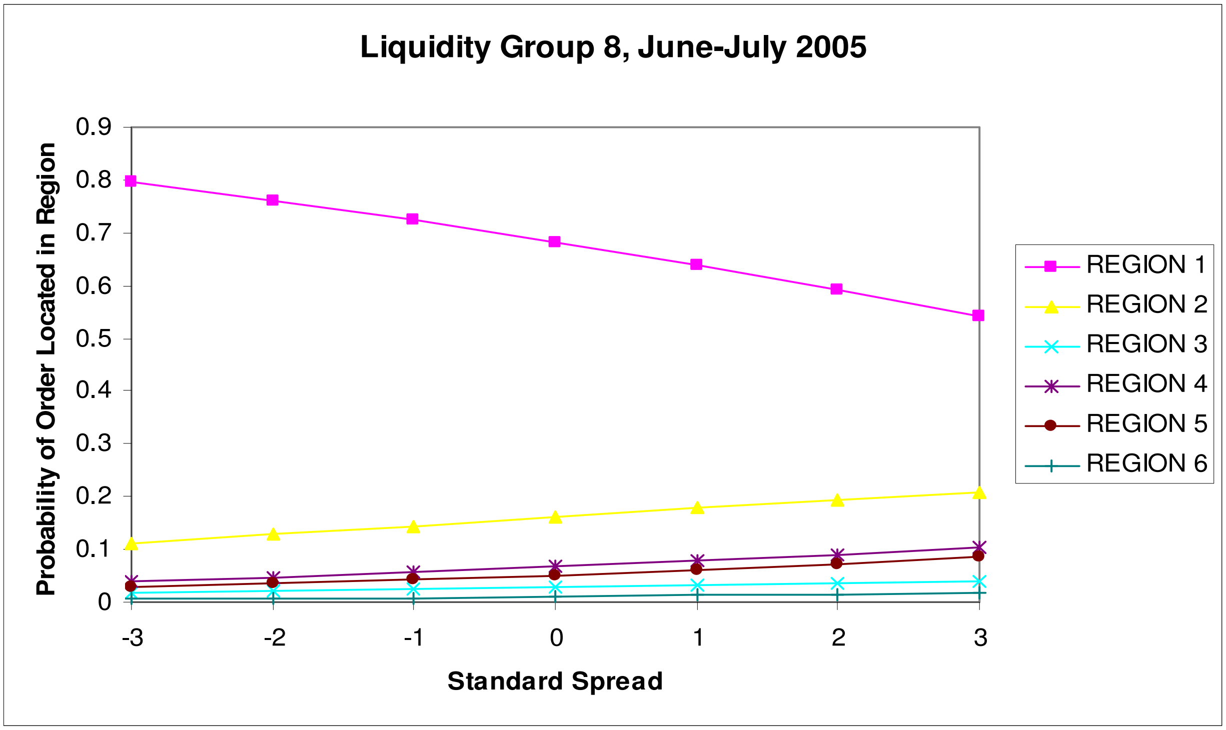

Figure 4 shows the probability that an undisclosed sell limit order is located within a particular region of the bid-ask spread. The bid-ask spread is subdivided into six equally sized groups where region 1 includes the best ask price and region 6 includes the region prior to best bid price.

FIGURE 5

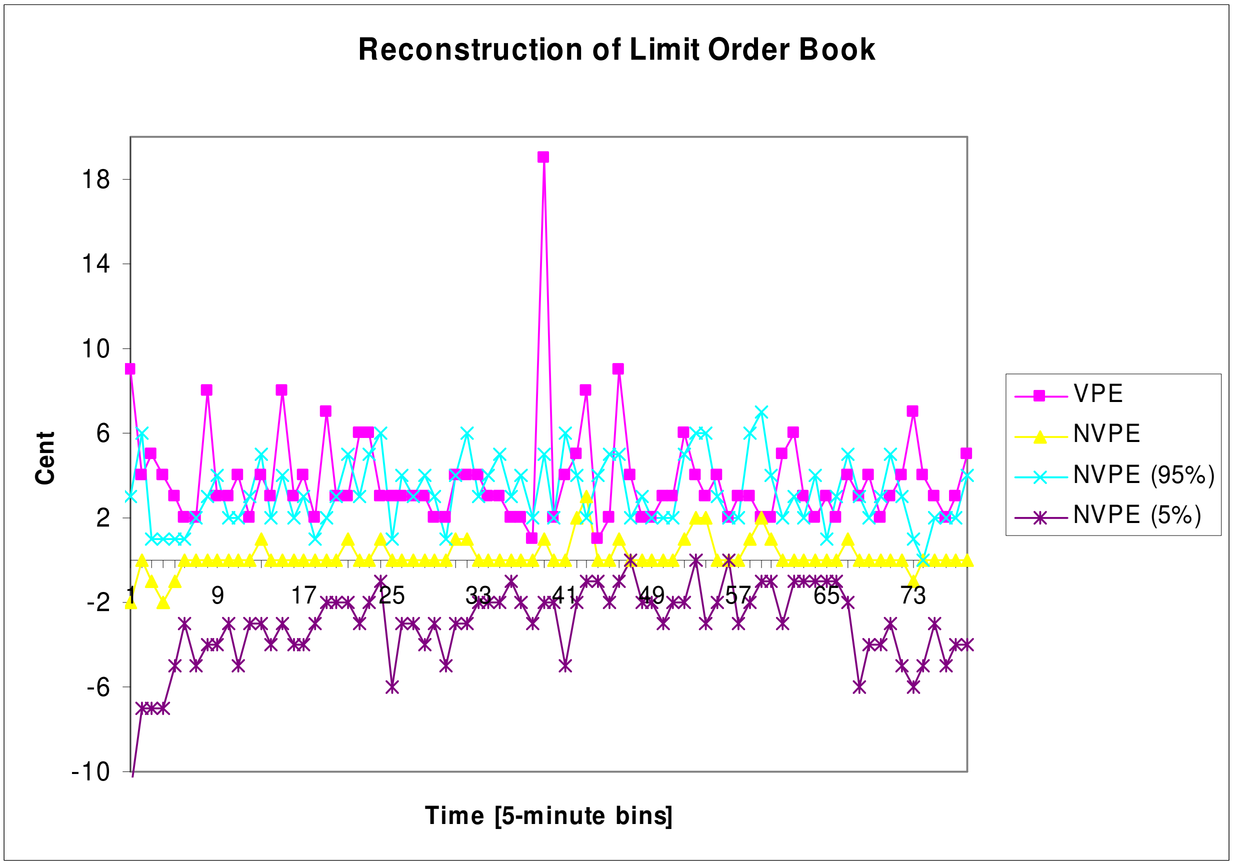

Figure 5 compares the average Virtual Price Error (VPE, the difference between the perceived and actual execution price) and average New Virtual Price Error (NVPE, the difference between the reconstructed limit book's perceived execution price and the actual execution price) for all stocks that have been classified in Liquidity Group 2. The trading day is broken-up into seventy-eight 5-minute bins and we present the average Virtual Price Error and average New Virtual Price Error in each of these subgroups. So as to capture the Virtual Price Error distribution, we also graph the 5% and 95% confidence levels of the NVPE.

FIGURE 6

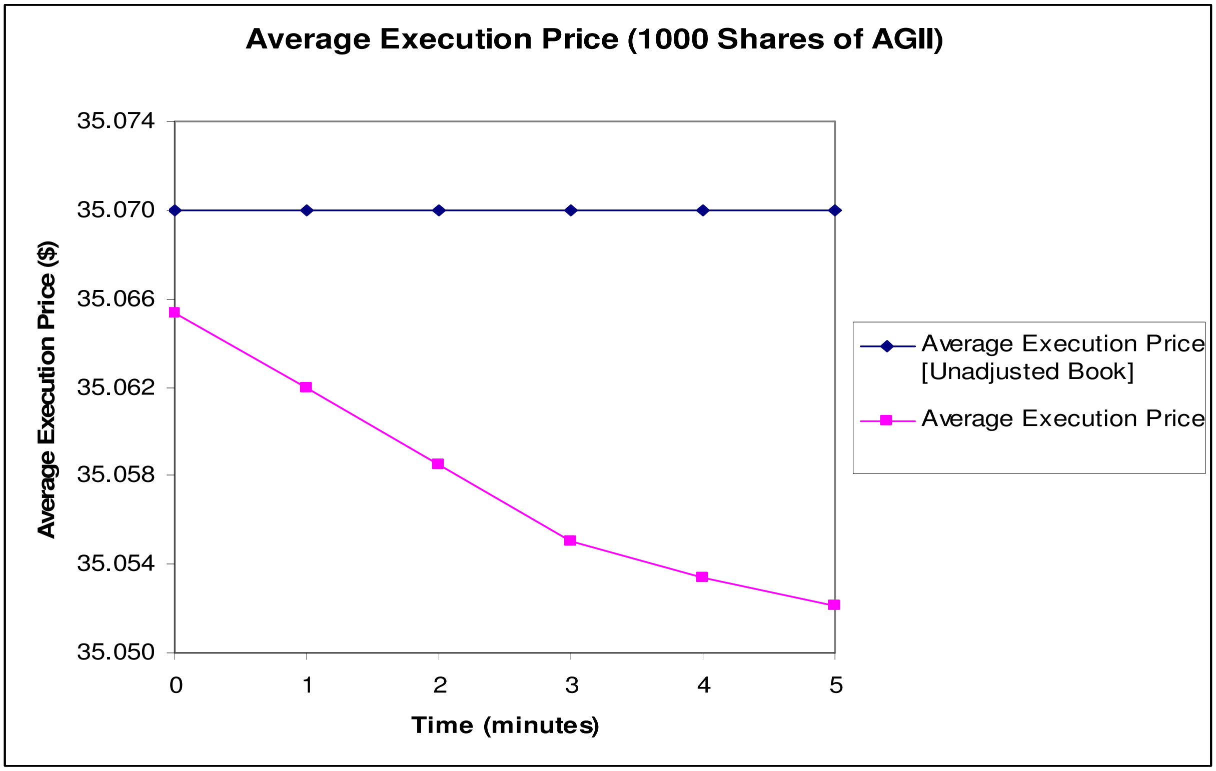

Figure 6 provides the average execution price associated with trading 1000 shares of AGII using different execution horizons at 10:40am. The average execution price of the unadjusted (static) limit order book is compared with estimates that include undisclosed limit order volume when trading instantaneously and with 1-, 2-, 3-, 4-, and 5-minute horizon.