QSG® Quantitative Services Group LLC
Trading Research Group
Research Note
November 2009

Beware of the VWAP Trap

Tick-level evidence that 'High Frequency Trading' is negatively impacting volume participation algorithms.

Investment Research | QSG Analyst Team

This study reveals that significantly higher impact costs and trading velocity are incurred for VWAP algorithms when compared to Arrival Price algorithms, especially when applied to liquid, low price stocks. This is contrary to the popular perception that VWAP strategies reduce trading costs. The results suggest that High Frequency Trading (HFT) strategies are materially contributing to these increased costs. The report also illustrates that the increased order parceling related to VWAP strategies has negative ramifications, including increased tick risk, larger statistical footprints and the possibility of costly trading spirals. These insights are particularly relevant given the lack of empirical evidence isolating the impact of HFT strategies on institutions employing non-HFT strategies. The analysis is significantly strengthened by employing a powerful set of tick-based algorithm evaluation measures.

HFT has dominated the securities trading industry headlines for much of 2009. In reviewing the portrayal of this activity by the industry's media sources, we discovered a mixture of concerned curiosity and outright disdain for HFT from outspoken representatives of the buy-side and agency-only execution providers. Recently, support for the practice has emerged from a few buy-side firms and many of HFT's more public practitioners. Industry blogs and self styled 'white papers' have gotten the attention of Congress, and the SEC is considering regulatory changes, already taking steps to curb an HFT-related activity called 'Flash Orders'. The uproar surrounding HFT has been created largely by anecdotal evidence from traders and the soaring profits reported by the HFT operations of a few firms forced to disclose their results to the public. Unfortunately, there is little empirical evidence on the effect that high-frequency trading has on traditional institutional trading desks that are employing non-HFT strategies. In this report, we leverage QSG's proprietary tick-based transaction cost attribution methodology to reveal empirical evidence confirming that significant increases in market impact costs are being experienced by certain types of institutional size trades. This report, the first in a series, will focus on a surprising segment experiencing HFT impact, executed through automated algorithms in liquid US stocks - VWAP targeted orders.

The Evolution of High Frequency Trading Strategies

The definition of high-frequency trading is both broad and evolving. As the name suggests, it generally involves trading strategies that rely on rapid, large scale order executions facilitated by advanced computing and communications technology. The significant data, order routing and communications infrastructure that supports HFT strategies are designed to virtually eliminate execution delays or 'latency'. At its most basic level, the capacity to execute instantaneously creates the ability to arbitrage prices across execution venues, of which greater than 40 exist in the US alone. This profit making opportunity is then augmented by two additional strategies that are pursued exclusively or in combination: electronic market making and statistical arbitrage.

Electronic market makers create two-sided markets with the goal of profiting from the spread between the prices at which they buy and sell. Through changes in regulation, increases in volume, the introduction of decimalization and the automation of the trading floor, market making has significantly evolved over the last decade. The challenges to the business model of market makers include dramatically narrowed bid/ask spreads, fragmented markets and an increasing number of competitors. Today, market makers rely on sophisticated statistical models and trading algorithms that have automated the previously manual decisions as trading flows have increased dramatically and executions are measured in milliseconds. In addition, exchanges and ECN's provide incentives for electronic market makers to provide liquidity by offering a rebate on trades, usually about $0.20 per 100 shares. Firms like Global Electronic Trading Company LLC (GETCO), Tradebot Systems Inc., Hudson River Trading LLC and Wolverine Trading LLC are examples of large electronic market making firms.

The other broad category of HFT profits come from statistical arbitrage (stat arb). Similar to electronic market making activities, stat arb shops using HFT strategies leverage a combination of low-latency trading technologies and an automated decision engine. The statistical patterns that these strategies exploit generally occur intra-day and don't require overnight positions. Statistical arbitrage strategies in a high-frequency setting analyze price, volume, depth of book and trading velocity patterns to identify exploitable price trends. These price trends can be very short and shallow, created by a temporary liquidity imbalance, or they can be much longer in term and quite large, created by a large institutional buy/sell order. In the former, the role of the HFT stat-arb strategy is very similar to the market maker's role. In the latter, the role of the HFT stat-arb strategy is very similar to trading ahead of the institutional order. In such cases, the stat-arb profits are not necessarily derived from a spread premium and rebate; they profit by 'surfing' the price moves available by trading alongside the institutional orderflow. Identifying such orderflow can shift the probability of profit in favor of the HFT and in some cases create a significant liquidity imbalance of its own. These circumstances have the potential to dramatically increase institutional trading costs.

The one factor consistent across all HFT strategies is that they benefit from increased volumes and micro-second execution advantages. The majority of profits are made on razor thin margins. For example, market makers are often executing trades with gross margins of 0.05% or nominally between one and two-tenths of a penny. Of course most of the volume in equity markets is concentrated in the largest capitalization stocks with the narrowest spreads and the lowest price volatility. These characteristics attract the attention of HFT strategies; essentially, trading begets trading in these names. Non-HFT trades (now thought to be less than 50% of overall equity trading) are necessary for HFT strategies to flourish. In order to manage the trading velocity required by HFT strategies, firms must often submit and cancel thousands of orders per second while simultaneously monitoring the market data that drives their automated trading systems. To get the broadest and fastest access, many HFT firms subscribe to enhanced market center data feeds like NASDAQ's ITCH or BATS' FASTPITCH, in place of feeds from the slower consolidated Securities Information Processor (SIP). In addition to the expensive data feeds, these firms have to invest heavily in their internal systems and data processing engines in order to monitor and process data on thousands of securities simultaneously. The time frames under consideration are so short that many of these firms have taken the step to co-locate their technology inside the buildings housing the exchanges.

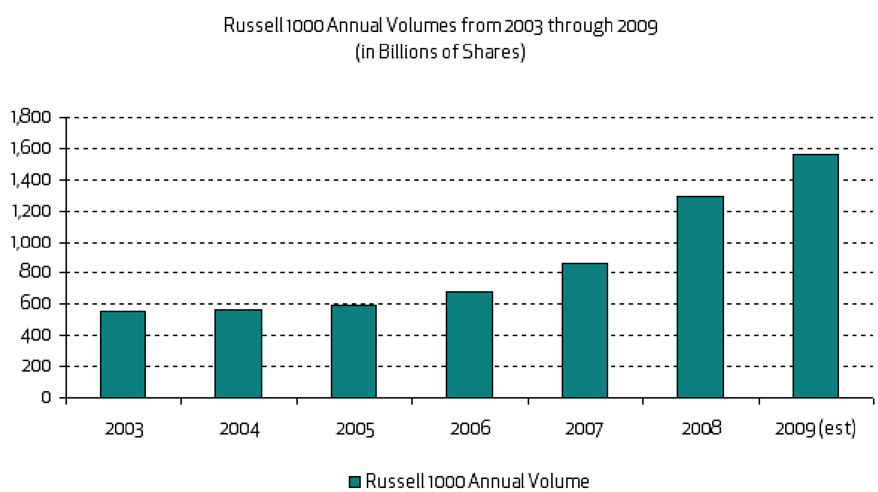

Figure 1: Annual total volumes for stocks in the Russell 1000 from 2003 through 2009.

Much of the media reports surrounding HFT have been focused on the potentially negative impact of such strategies. Proponents of the strategies usually point to the dramatic increases in volume (often equated with liquidity) and declining average bid/ask spreads. However, the collapse in average trade sizes, which dramatically increases the overall number of executions, is often overlooked. This increased trading velocity generates both increased profit opportunities for market makers and greater signaling opportunities for stat-arb strategies.

The Interaction of HFT and Institutional Orderflow

US equity volumes have grown dramatically in recent years, much of it driven by the emergence of HFT strategies. Notable is the significant overall increase in volume and the number of trades, especially following 2005 (see Figure 1 on page 2).

The prevailing estimates of the daily volume impact of HFT strategies now range between 50% and 70% of daily volume in the US. Of course, as HFT strategy estimates cross the 50% barrier, it is clear that HFT participants are increasingly trading amongst themselves. If electronic market makers are primarily providing liquidity, taking the other side of natural orderflow, then clearly stat-arb strategies that are demanding liquidity (as they seek to identify and exploit orderflow patterns) would account for HFT strategies pressing through the 50% barrier.

While electronic market makers are supposedly 'providing liquidity when it would not otherwise be available', what are the effects of their presence when natural liquidity is present? The 'interpositioning' effects of high-frequency market makers can cause inflated volumes in very short periods of time. Instead of a single transaction occurring between two natural sides, a market maker's speed in submitting and cancelling orders may cause two or three trades to occur. The volume inflation and reduction in average trade size that results from these activities especially influences participation-based trading strategies as these algorithms are designed to track market volume. It may also encourage trading algorithms to cut execution sizes into even smaller lots, significantly increasing the number of trades, leading to a new source of exploitable 'information' and increased exposure to 'adverse tick' risk.

It's well known that sophisticated stat-arb models routinely monitor market data and the depth of limit order books to detect asymmetries in trading interests. The goal is to exploit and profit from them before the flows reverse and larger traders have a chance to finish their orders. These HFT strategies increase the costs of completing institutional trades and often introduce 'adverse selection' as orders are completed in names that are moving contrary to the institutional trader's investment goals. Our study seeks to illustrate the role of high-frequency trading in the implicit transaction costs associated with participation-based algorithms.

Algorithmic Performance Evaluation: Data & Methodology

QSG's tick-based T-Cost Pro® transaction cost attribution methodology delivers a significant evaluation advantage. This methodology breaks Implementation Shortfall into two pieces, the footprint left behind by the series of trades executed by a client and the price impact of all the competing trades executed over the trading period. This powerful attribution allows us to isolate key trading costs and reduce the influence of idiosyncratic price movement. This is the key to comparing transaction costs over time.

The data driving the analysis is QSG client executions including data regarding the trading algorithm employed between January 1, 2009 and October 23, 2009. Our sample set included over more than 95,000 orders and represented greater than $30 billion in executed value. To further illustrate the topic, we analyzed the trades between September 10 through October 23, 2009, using the 'adverse selection' analysis methodology introduced by Henri Waelbroek et al in a recent study from Pipeline Financial and AllianceBernstein titled 'Adverse Selection vs. Opportunistic Savings in Dark Aggregators'.

The T-Cost Pro tick-based attribution methodology matches fill-level client execution data and trade and quote data using a proprietary matching algorithm. Once the data is synchronized, the T-Cost Pro system calculates the cumulative 'Liquidity Charge' or footprint that resulted from the client executions. This impact cost is separated from the price impact, or 'Timing Consequence', of the competing trades that are responsible for the remainder of the price drift over the execution period. The technique considers the impact made by each individual execution in the order and accumulates the impact throughout the life of the order.

Avg Tracking Error	The average performance differential between the algorithm's execution prices and a 10% participation-weighted VWAP benchmark (negative indicates 'opportunistic savings' and positive indicates 'adverse selection')
Avg Liquidity Charge	The component of Implementation Shortfall attributed to the direct price impact of the algorithm, as calculated by QSG's T-Cost Pro tick-based cost attribution technique.
Avg Implementation Shortfall	The average total shortfall of the algorithm's execution prices from the price of the stock at the time the broker algorithm acknowledged the order of the trader

Figure 2: Key Terms

The sum of the cumulative Liquidity Charge and the Timing Consequence equals the Implementation Shortfall. This attribution allows us to investigate costs at a level that extends well beyond what we can do with the traditional benchmark measures and is particularly useful in illustrating the performance characteristics of algorithmic trading strategies. By separating costs into these two elements we can examine both the liquidity management characteristics and the price trend reaction characteristics of an algorithm. This technique is also better suited for the complex challenge of contextualizing 'Best Execution' analysis.

In addition to the T-Cost Pro attribution, we calculated a benchmark methodology proposed by Waelbroeck et al, which compares the execution prices obtained by an algorithm to a 'passive' 10% volume participation benchmark. This 'participation weighted' benchmark is a theoretical price calculated as if the algorithm executed 10% of each market trade from the time the order was placed until the order share size was completed. The concept of the PWP (Participation-Weighted Price) benchmark had been introduced in earlier studies by Abel Noser. The Waelbroek analysis compares the PWP benchmark to the price at 'arrival time' and calls the difference the 'Implementation Shortfall' of PWP (IS_{PWP}). This IS_{PWP} is then compared to the Implementation Shortfall of the algorithmic order (IS_{ALGO}). When IS_{PWP} is greater than IS_{ALGO}, the algorithmic order is attributed with 'Opportunistic Savings' and when IS_{PWP} is less than IS_{ALGO}, the order is attributed with 'Adverse Selection'. The difference between these two shortfalls is also referred to as 'Tracking Error'.

This theoretically derived participation benchmark is an advance in evaluating algorithmic performance. However, the VWAP benchmark remains best suited for understanding performance characteristics as a function of varying volume participation. When the strategy's or algorithm's goals vary from the calculation principals of the benchmark, the value of the analysis is dramatically diminished. It's for these reasons that we prefer a multi-dimensional analysis that includes both benchmark and the 'bottom-up' analysis found in the T-Cost Pro attribution. We will combine the insights of these techniques for a more comprehensive analysis.

2008's steep decline in stock prices of large capitalization stocks has introduced an interesting phenomenon. Since most stocks are quoted in penny increments, the minimum tick size for most stocks is also $0.01. This introduces an interesting profit margin bias for HFT strategies. Since profits accrue to HFT on a share basis, low priced, high turnover stocks often have the potential to provide improved profit margins. To account for this we divided our trade dataset by stock price, splitting it into stocks less than/greater than $10 for all trades less than 5% of the day's volume. We suspect that HFT activity will reflect the margin bias and be more intense in the lower priced stocks. As a preliminary evaluation of this assumption, we measured the Market Trade Velocity (MTV) in the stocks traded in each sample set. We found that stocks less than $10 had 11% more executions per second on average than stocks greater than $10.

In addition to the cost attribution and performance metrics we've outlined, we will also conduct an in-depth review of order parceling activities and introduce metrics that provide insight into parceling decisions and their relation to realized transaction costs. The following metrics are calculated for each trade in the sample and then equal-weighted averages are established by algorithm type.

To ease comparisons we categorized the algorithms into two groups, VWAP algorithms and Arrival Price (Implementation Shortfall) algorithms. VWAP algorithms are engineered to execute in-line with market volumes through time, while Arrival Price algorithms are designed to execute near the Arrival Price with less regard for targeted volume patterns and time intervals. Arrival Price algorithm orderflow is often more concentrated, occurring earlier in the execution period. These algorithms are purported to have higher levels of 'market impact' than VWAP algorithms, which have reputations for executing at smaller interval volume participation rates.

Analysis Results

The first set of results is for the subset of data where we calculated both the T-Cost Pro attribution measures and the participation weighted metrics. The metrics are presented in percentage terms (basis points) so it should be noted that transaction costs in smaller price stocks will tend to be greater than higher price stocks; this separation improves the quality of our comparisons (Figure 3).

It is significant that the VWAP costs are larger for all measures during this period. This is contrary to the perception that VWAP implementations are an efficient way to reduce trading costs, especially market impact. While the average liquidity charge difference between Arrival Price and VWAP algorithms for the greater than $10 category isn't nominally large, it is greater than 50% greater on a relative basis.

In the less than $10 category, the performance difference is striking and significant across all three metrics. The idea that the footprint created by VWAP strategies could be almost three times that of the Arrival Price algorithms has large ramifications for the automated strategies and the possible influence of HFT. Of particular importance in the less than $10 subset is the comparison of Liquidity Charge to Implementation Shortfall. Given that the measured Liquidity Charge is greater than the total Implementation Shortfall, it is clear that the algorithm's own impact on price is the driver of these transaction costs. These comparisons to the participation weighted

Percent of ADV	The average size of an algorithmic order in the sample as measured by the order's percent of daily volume.
Trade Duration	The time in minutes to complete the order.
Total Strikes	The amount of individual fills into which the algorithm parceled the order
Strike Participation Rate	The amount of individual fills into which the algorithm parceled the order, taken as a percentage of the total fills within the algorithm's execution horizon
Trade Size Ratio	Calculated as the algorithm's average fill size divided by the average trade size in the market over the algorithm's execution period
Adverse Tick Ratio	Calculated as the total number of adverse ticks occurring in the algorithm's execution horizon divided by the total number of prints during the execution horizon (excluding the client adverse executions and prints) minus the number of adverse ticks (upticks for Buys, downticks for Sells) on which the algorithm executed divided by the total number of executions; negative values mean 'more adverse ticks than the market'

Figure 3: Summary of Trading Cost Measures

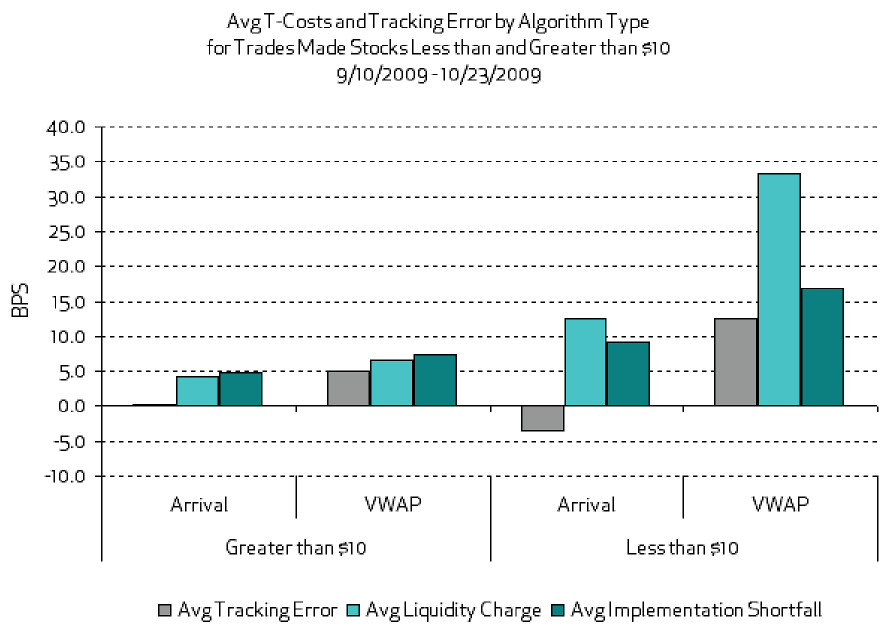

Figure 4: Average T-Costs and Tracking Error by algo for trades made in stocks less than and greater than $10 for the period 9/10/2009 through 10/23/2009.

	Algo Type	Avg Tracking Error	Avg Liquidity Charge	Avg Implementation Shortfall
Greater than $10	Arrival	0.3	4.2	4.8
	VWAP	5.0	6.6	7.4
Less than $10	Arrival	-3.5	12.5	9.1
	VWAP	12.6	33.3	17.0

Table 1: Average T-Costs and Tracking Error by algo for trades made in stocks less than and greater than $10 for the period 9/10/2009 through 10/23/2009.

	Algo Type	Orders	Avg Pct DV	Avg Duration (min)	Avg Strikes	Avg Strike Participation Rate	Avg Trade Size Ratio	Avg Pct Adverse Tick Ratio
Greater than $10	Arrival	8196	1.0%	16	49	4%	100%	-8%
	VWAP	3548	0.8%	179	85	2%	99%	-21%
Less than $10	Arrival	1053	1.3%	41	61	7%	109%	-3%
	VWAP	436	1.5%	202	229	8%	96%	-16%

Table 2: Average 'tick metrics' by algo for trades made in stocks less than and greater than $10 for the period 9/10/2009 through 10/23/2009.

	Algo Type	Orders	Avg Pct DV	Avg Duration (min)	Avg Strikes	Avg Strike Participation Rate	Avg Trade Size Ratio	Avg Pct Adverse Tick Ratio
Greater than $10	Arrival	54960	1.0%	22	71	5%	103%	-6%
	VWAP	29266	1.0%	116	103	3%	98%	-21%
Less than $10	Arrival	7149	1.2%	37	102	8%	102%	-2%
	VWAP	4087	1.4%	119	198	6%	94%	-19%

Table 3: Average 'tick metrics' by algo for trades made in stocks less than and greater than $10 for the period 1/1/2009 - 10/23/2009.

benchmark indicate that the Arrival Price algorithms add value in the less than $10 group and slightly underperform for the larger price stocks. The large positive Tracking Error for the VWAP algorithms shows that these orders on average underperform a 10% participation weighted VWAP by 13 bps in the less than $10 subset. QSG's attribution to Liquidity Charge and Implementation Shortfall suggest that this cost is not due to price drift over the execution period but rather the VWAP algorithm's own impact on price.

To better understand the origin of these surprising results it's important to analyze the trade parceling characteristics related to the two categories.

Of most interest in Table 2 & 3 is the sharp increase in trading velocity (strike participation) for VWAP algorithms in the less than $10 subset compared to those in the greater than $10 subset. The average trade duration only increases by 13%, while the average number of child order executions (strikes) increases by 170%. This is also reflected in the Average Strike Participation, which increases by a factor of 1.75 for Arrival Price algorithms and by a factor of 4 for VWAP algorithms. The much larger number of executions related to VWAP strategies for similar sized orders and trade durations indicates the hyperactive parceling activities of these algorithms in the less than $10 subset of trades. We have found some VWAP algorithms to execute multiple times per second to keep up with volume, exposing the order to the additional 'adverse tick' risk that drives trading costs.

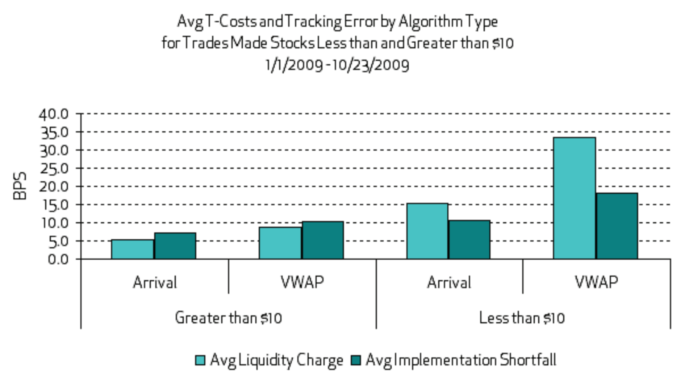

Figure 4: Average T-Costs by algo for trades made in stocks less than and greater than $10 for the period 1/1/2009 through 10/23/2009.

While both algorithm categories executed near the market average trade size for trades in stocks greater than $10, the values diverge in the low priced stocks. Arrival Price algorithms registered an average trade size (109%) advantage, while the VWAP algorithms show a gap (96%) in the measure. These statistics, showing greater strike participation and smaller execution sizes are indicative of higher velocity trading, which both attracts and is caused by high-frequency order flow. Importantly, it is apparent that VWAP algorithms incur a much greater percent of adverse ticks than does the market during their trading interval, especially when compared to Arrival Price algorithms in the dataset having trades in stocks less than $10.

	Algo Type	Avg Liquidity Charge	Avg Implementation Shortfall
Greater than $10	Arrival	5.4	7.2
	VWAP	8.7	10.2
Less than $10	Arrival	15.4	10.6
	VWAP	33.3	18.1

Table 3: Average T-Costs by algo for trades made in stocks less than and greater than $10 for the period 1/1/2009 through 10/23/2009.

It should be noted that although the VWAP algorithm group shows a slightly larger average order size than the Arrival Price algorithm group in the less than $10 subset, we employed two-sample t-tests and found these averages have no significant difference at the 5% level. We also found all cost and tick metrics averages to be significantly different at a 5% level.

When the analysis period is expanded to 2009 year-to-date, it's clear that the trend in high cost VWAP orders persists, especially in the lower price stock category.

While the average cumulative Liquidity Charge is over 60% greater for VWAP algorithms compared to Arrival Price algorithms in the high price category, they increase greater than 110% in the low price segment. As was discovered previously, the average number of strikes per order increased dramatically for the VWAP algorithms in the less than $10 subset while the trade duration remained nearly constant. In context of Liquidity Charge, this is compelling evidence of adverse execution quality for VWAP algorithms, especially when considering the average adverse tick ratios of 20% for VWAP algorithms year-to-date. The fact that the VWAP adverse tick ratios are three to nine times larger than those of the Arrival Price algorithms is of great concern as these values are consistent with systematic liquidity imbalance biases and predatory competition. In the year-to-date analysis period, we again found all cost and tick metrics averages for algorithms in the same subset to be significantly different at a 5% level.

Conclusion

Our results indicate that there is a significant difference in the costs and trading velocity of VWAP algorithms when compared to Arrival Price algorithms, especially when applied to low price stocks. The Tracking Error (PWP) measure confirms most VWAP algorithms are challenged to beat a 10% PWP benchmark. This result is consistent with both the negative impact of certain HFT strategies and the possibility of positive momentum in institutional orderflow. This exposes the limitation of such benchmarks in their ability to separate an order's price impact from price drift associated with other trades over the trading interval. The limitation is overcome by the T-Cost Pro Liquidity Charge measure which is capable of isolating the cumulative impact of each order, revealing that the average impacts for VWAP algorithms are nearly double those of Arrival Price algorithms. This is a significant revelation to proponents of VWAP algorithms as a low impact strategy that flies under the predatory HFT radar.

The details of the study uncover an important artifact from today's trading environment: increased order parceling has three negative ramifications. First, more 'strikes', or executions per order, increase a client's exposure to adverse ticks and this tick risk translates into higher impact costs. Second, more strikes increase the chances of leaving a statistical footprint that can be exploited by the 'tape reading' HFT algorithms. Third, should HFT strategies identify the order and begin to trade in anticipation of the orderflow, this will begin a positive feedback loop that can significantly change an algorithm's behavior and invite even more predatory orderflow.

The incentives to take action to prevent such impact costs are compelling. If the impact costs realized by the VWAP algorithms were reduced to the level of the Arrival Price algorithms, the clients in this study would have reduced trading costs by $35 million. This study also highlights the value of rigorous algorithmic evaluation measures. The proper trading analysis measures empower equity managers to retake control of the execution process with confidence, avoiding the errors of anecdotal decision making that aren't supported by facts of an increasingly challenging trading landscape.

For additional information, please contact:

QSG® Quantitative Services Group LLC
qsgresearch@qsg.com
1560 Wall Street Suite 334
Naperville, IL 60563
PH 630.637.8088
FAX 630.637.8333
www.qsg.com