Measuring and Interpreting the Performance of Broker Algorithms

Ian Domowitz
Managing Director
ITG Inc.
380 Madison Ave.
New York, NY 10017
212 588 4000
idomowitz@itginc.com

Henry Yegerman
Director
ITG Inc.
380 Madison Ave.
New York, NY 10017
212 588 4000
hyegerma@itginc.com

August, 2005

We thank Scott Baum of ITG for his insights with respect to trade structure.

The advantages of algorithmic trading have popularized algorithms to the extent that most major brokers, and some technology providers, offer the service in some form. Recent studies suggest that anonymity, reduced market impact, trading efficiency, and lowering the overall cost of trading are drivers of the trend from the customer perspective.[1]

This list of attributes suggests that a link between cost measurement, both pre-trade and post-trade, and algorithmic trading activity should be strong. In fact, firms report that algorithm usage frequently is triggered by some form of cost analysis. Further, transaction cost technology that supports algorithmic trading is found to be the leading factor influencing growth in algorithmic trading.[2]

The focus of this chapter is on information and measurement leading to informed choice and evaluation of algorithmic trading engines. Choice and evaluation are closely linked in our minds. The number of vendors is growing quickly, but not half as fast as the alphabet soup labeling the number of strategies available. Evaluation, hence measurement, must include the choice of strategy and a preferred set of vendors. We begin with some remarks relating to the strategies themselves.

Trade Structure

Algorithm construction is an exercise in structuring a sequence of trades, and the choice of an algorithm follows the same basic principles. At the most abstract level, trade structure is a continuum, ranging from unstructured, opportunistic liquidity search to highly structured, precisely scheduled sequences of trading activity, generally linked to a certain benchmark, such as VWAP. Unstructured liquidity search is associated with real-time information and raw or adjusted decision price benchmarks. Structured approaches embody tight tracking strategies, based on historical data, and are relative to participation strategy benchmarks. Although there are many ways to describe structure, the following factors serve to illustrate the nature of required pre-trade information.

- *Trade Horizon*. Shorter horizons require less structure. For example, a half-hour VWAP trade and a similarly timed pegging and discretion strategy will not yield

[1] See, for example, "Institutional Equity Trading in America 2005: A Buy-Side Perspective," Adam Sussman, Tabb Group, June 2005.
[2] 38 percent of firms responded that algorithm usage is triggered by cost measurement, while 35 percent identified cost analysis supporting usage to be the major driver. See "Marching Up the Learning Curve: The First Buy-Side Algorithmic Trading Survey," Randy L. Grossman, Financial Insights, May 2005.

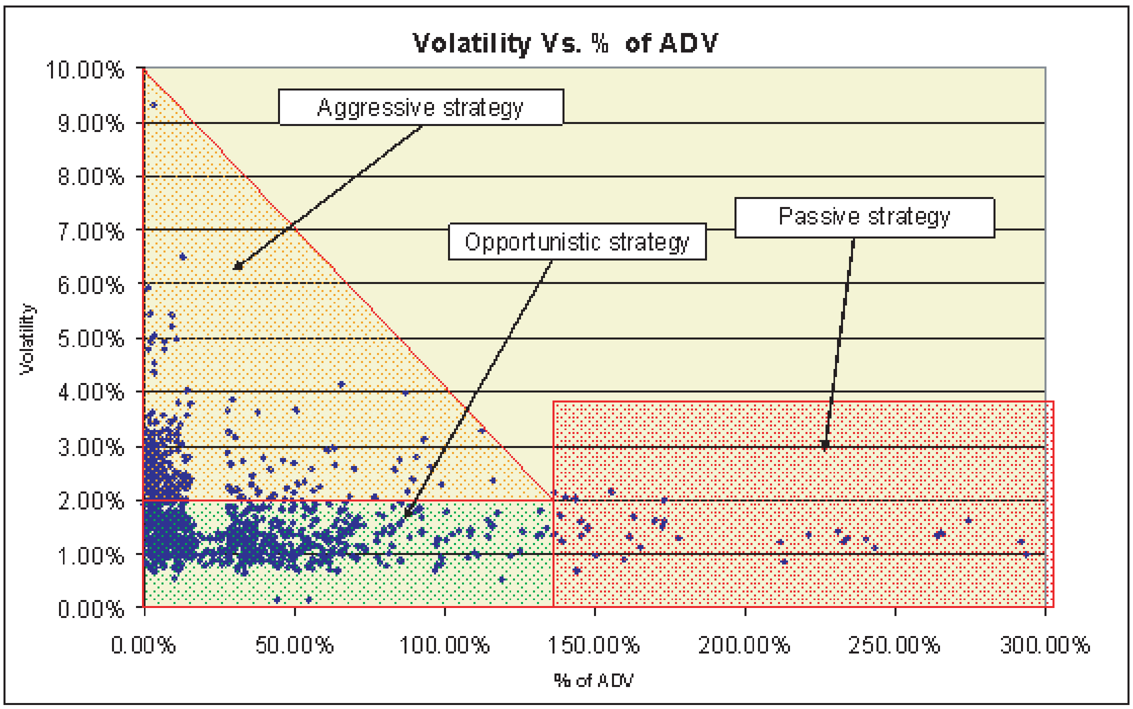
wildly different outcomes. Inputs into the horizon choice include volatility and the order's percentage of ADV, for example, in addition to any specific portfolio instructions.

- *Need to Finish*. The higher the need to finish an order, the more structure is needed, in order to avoid falling behind schedule. The type of pre-trade information here relates more to portfolio manager instructions than to specific analytics.
- *Predictability*. Predictability considerations cover alpha estimates, volume profiles, and transaction costs. The degree of predictability governs the degree to which a horizon and a schedule should be adhered. This consideration requires the use of properties of the distribution of estimates, in addition to averages, such as standard deviation measures.
- *Price Sensitivity*. As price sensitivity increases, structure becomes less useful, due to the need to advertise willingness to trade. Short-term volatility history, as well as real-time deviations, are inputs along this dimension.
- *Risk Tolerance*. This item refers to execution risk versus the benchmark. Greater tolerance generates less need for a structured horizon and schedule. Judgments with respect to risk tolerance are informed by pre-trade information in the form of "trading efficient frontiers," mapping out optimal tradeoffs between risk, cost, and alpha, for varying trade horizons.
- *Need for List Integrity*. The greater the need for time uniform execution across a list, the more structure is required. Portfolio risk increases as one moves, say, from time-weighted average price strategies to liquidity search, the performance of which is measured in terms of implementation shortfall. Pre-trade inputs are similar to those expected in standard portfolio analysis. Stock correlations are needed, as well as acceptable tracking error to an index. Constraints with respect to dollar and/or sector imbalance are factored in through optimization engines integrated with the pre-trade analytics. This also suggests choices (and tradeoffs to be considered) between minimization of total risk, minimization of the tracking error of portfolio residuals against index benchmarks, and minimization of expected cost.

Informing the Trade Structure: Aggressiveness and Opportunity

There are a variety of factors influencing the choice of trade structure; hence there are many different requirements for pre-trade information. As an illustration, we concentrate here, and later in the chapter, only on volatility and the percentage of ADV represented by the order. The former is a market-wide proxy for market conditions, while the latter is a function of the chosen portfolio. Figure 1 depicts the two characteristics for a recent large transition.

Figure 1



The original strategy contemplated breaking the list into as many equal slices as days in the horizon, as set by the transition manager. The goal was to hit VWAP for each name on each day. This strategy produced transaction costs, which were roughly double those expected, based on a pre-trade cost estimate. When costs were measured against the VWAP as a benchmark, the expense of the transition increased another 32 percent relative to the expected cost calculation.

Figure 1 suggests, however, that different trade structures might usefully have been applied to elements of the overall list, as opposed to "one size fits all." Although the horizon governing the entire transition is fixed by the manager, not all pieces need be completed according to that schedule. For example, lower ADV orders, with names exhibiting high price volatility, justify an aggressive stance, arguably with minimal trade structure, exploiting expected lower market impact for small orders, while avoiding the opportunity costs associated with greater volatility and price sensitivity.

The pre-trade view also suggests orders for which a passive structure might be useful, namely for large orders exhibiting small price volatility, hence sensitivity. Market impact costs are expected to be high, while opportunity cost is low. Lack of price sensitivity suggests the ability to signal willingness to trade at a low cost.

While the passive stance might suggest a high degree of structure, in order to ensure completion. Uncertainty with respect to transaction cost estimates for extremely large order sizes argues for less structure, depending now on factors, such as the accuracy of volume distributions for those names, which are not summarized in the preliminary analysis summarized in Figure 1. This point is especially important in considering the lower volatility, lower ADV orders labeled as “opportunistic strategy” in the figure. True opportunism, in the interest of lower costs, might well be justified in the sense of less structure for the higher ADV orders in this category. While a more structured approach, combined with some aggressiveness may work in the case of truly low ADV/low volatility segments of the transition. Although it is clear that additional pre-trade information and research is needed for this particular case, transaction costs may be cut considerably by considering even such a rough cut at the data, relative to a simple uniform strategy over a fixed horizon.

Measurement and Interpretation

In a recent study[3], we analyze the cost of algorithmic trading based on a sample of 2.5 million orders, consisting of almost 10 billion shares traded from January through December of 2004. The data come from over 40 institutions, and cover the performance of six broker-provided algorithmic trading systems. In the aggregate, we find that algorithmic trading has lower transaction costs than alternative means represented in the sample, based on a measure of implementation shortfall. Controlling for trade difficulty, differences in market, side of trade, and volatility regime does not change the result. Algorithmic trading performance relative to a volume participation measure, VWAP, also is quite good, averaging only about two basis points off the benchmark, although certainty of outcome declines sharply with order size. These differences in uncertainty are highlighted in the comparison of performance across vendors of the service, in which equality of average performance also breaks down, once order sizes exceed one percent of ADV.

The study sheds light on various aspects of algorithm performance, and provides some qualitative lessons, as well as quantitative conclusions. Our interest here, however, is in the potential interpretation and use of such results from an integrated trading perspective. As an example, we follow the previous discussion of trade structure choice, expanding the two-dimensional view offered in Figure 1 to three dimensions in Exhibit 1, which is based on the same data used in our earlier report.

Exhibit 1

Algorithmic Trading Costs by Broker, ADV, and Volatility

Cost V. MBA		**Order Pct of ADV**			
Volatility	**Broker**	**Less than 1%**	**1 - 5%**	**5 - 10%**	**Total**
Low	Broker 1	(5)	(9)	(14)	(6)
	Broker 2	(8)	(18)	(26)	(10)
	Broker 3	(2)	(5)	(1)	(2)
	Broker 4	(2)	(22)	(16)	(3)
	Broker 5	(3)	(6)	(23)	(4)
	Broker 6	(9)	(31)	(26)	(16)
Low Total		**(5)**	**(11)**	**(14)**	**(6)**
Medium	Broker 1	(7)	(13)	(9)	(8)
	Broker 2	(7)	(15)	(21)	(9)
	Broker 3	(4)	(46)	(72)	(11)
	Broker 4	(5)	(22)	4	(6)
	Broker 5	(8)	(28)	(52)	(13)
	Broker 6	(10)	(25)	31	(13)
Medium Total		**(7)**	**(15)**	**(13)**	**(9)**
High	Broker 1	(11)	(11)	(31)	(11)
	Broker 2	(9)	(5)	(45)	(9)
	Broker 3	(14)	(38)	(206)	(20)
	Broker 4	(16)	(24)	4	(17)
	Broker 5	(15)	(38)	(38)	(24)
	Broker 6	(19)	(37)	(8)	(23)
High Total		**(12)**	**(17)**	**(34)**	**(14)**
Total		**(7)**	**(14)**	**(19)**	**(9)**

Transaction cost results in basis points, broken down by order percentage of ADV and volatility, are further disaggregated by the broker providing the algorithmic

---

[3] "The Cost of Algorithmic Trading: A First Look at Comparative Performance," by Ian Domowitz and Henry Yegerman, *http://www.itginc.com/research/whitepapers/domowitz/algorithmictrading_2.24.2005.pdf*

trading strategy. The benchmark used in the table is the midpoint of the bid and ask prices at the time the order is received by the trading desk, which is a measure of the implementation shortfall variety[4].

Systematic regularities are difficult to discern, beyond the obvious conclusions that costs rise with volatility for given order sizes, and that costs increase with size for comparable volatility conditions. Irregularities suggest the real message: the relative performance ranking of vendors varies, depending on order type and trading environment, and this information ought to be taken into account in choosing vendors for a given trade structure. This view is consistent with the study cited above, in which a clear relationship between algorithmic trading cost and certainty of outcome relative to a benchmark across brokers is found to be lacking. The leading example illustrated in the table relates to performance in different volatility regimes.

If one were to look only at low volatility outcomes, brokers 3, 4, and 5 are roughly equivalent with respect to performance in the aggregate. This result is maintained for the lowest ADV trades, which dominate the sample here. On the other hand, for the next ADV category, ranging from one to five percent of volume, only brokers 3 and 5 appear to be doing a good job on a relative basis. As trades move into a range defined by relative order ADV of five to 10 percent, broker 3 now clearly outperforms its peers.

In medium volatility environments, the cast of "good" vendors changes, even in the aggregate. Broker 4 remains a top performer, but brokers 3 and 5 appear to be replaced by numbers 1 and 2. The aggregate ranking again is only slightly changed in the lowest ADV category. As trades in the range of one to five percent of ADV are considered, brokers 1 and 2 appear to survive, while number 4 slips. The latter result may be a consequence of some vagaries in this particular sample, since in the range of five to 10 percent of ADV; broker 4 is back, with results even better than number 1, while broker 2 no longer appears to be performing as well.

As we move to the highest volatility environments, brokers 1 and 2 still dominate the aggregate rankings. Interestingly, this performance comparison survives through order sizes up through five percent of ADV, after which broker 4 again appears to dominate in terms of performance.

Exhibit 2 below utilizes the same set of data to measure the risk associated with each broker's algorithms across different levels of volatility and demand for liquidity. Risk

[4] Low volatility is defined as less than 125 basis points of price movement per day. Medium volatility is defined as between 125 – 200 basis points, and high volatility is defined as greater than 200 basis points.

is defined here as the standard deviation of cost outcomes versus the benchmark. A greater standard deviation indicates greater volatility in the cost outcomes of a given broker algorithm relative to the ADV and stock price volatility categories.

Exhibit 2

Standard Deviation of Algorithmic Trading Costs by Broker, ADV, and Volatility

Standard Deviation of Cost V. MBA		Order Pct of ADV			
Volatility	**Broker**	**Less than 1%**	**1 - 5%**	**5 - 10%**	**Total**
Low	Broker 1	30	51	63	31
	Broker 2	25	40	32	26
	Broker 3	20	39	9	20
	Broker 4	21	31	64	21
	Broker 5	23	36	37	23
	Broker 6	31	44	20	32
Low Total		**27**	**48**	**58**	**28**
Medium	Broker 1	42	67	75	44
	Broker 2	36	69	50	38
	Broker 3	26	31	52	26
	Broker 4	29	48	57	29
	Broker 5	31	49	93	33
	Broker 6	38	63	68	40
Medium Total		**38**	**65**	**74**	**39**
High	Broker 1	60	97	113	64
	Broker 2	62	93	137	64
	Broker 3	35	73	133	36
	Broker 4	40	76	86	42
	Broker 5	49	79	89	52
	Broker 6	61	83	164	63
High Total		**56**	**93**	**115**	**60**
Total		**40**	**75**	**89**	**42**

The dispersion of cost outcomes follows common sense intuition. More volatile orders have more risk than less volatile orders and orders demanding more liquidity embody more risk than orders demanding less liquidity. Analogous to the transaction cost results, there is a significant difference between brokers within the different categories, ranging from 11 bps for the easiest orders (Low Volatility and Less than 1% ADV) to 78 bps for the most difficult orders (High Volatility and 5-10% ADV). However, unlike the performance averages, there appears to be a certain consistency of results relative to brokers with respect to the distribution of cost outcomes. Broker 3 ranked either first or second in eight of the nine ADV / Volatility categories, and Broker 4 ranked either first or second in seven of the nine categories. Conversely, Broker 1 ranked either fifth or sixth in six of the nine categories.

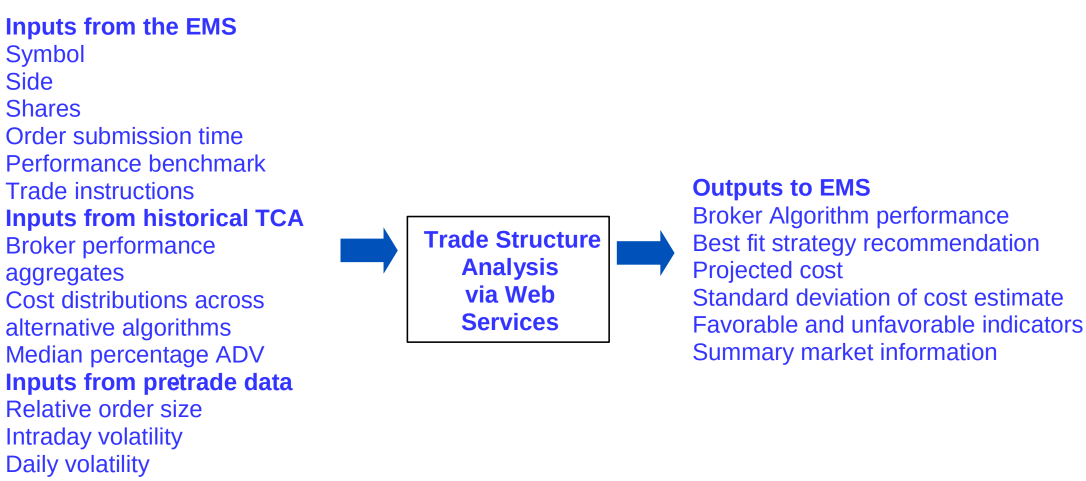
In all of these cases, the extra information is important in differentiating providers. We have focused on two of the factors, volatility and order percentage of ADV, which determine how different brokers model the optimal trade structure in their algorithmic servers. The data suggests there is a difference between the algorithms of various brokers with respect to both trading cost results and the risk associated with achieving those results. This points to the possibility that some broker algorithms may be better suited for specific order types and trading environments than others, and that users of algorithms are not yet identifying which broker algorithms are preferable in different situations.

Tying the Pieces Together

In order to develop a feedback loop that allows algorithm users to make informed choices and evaluate tradeoffs between different broker algorithms, a body of historical performance data is required. Such data may include order information, market information, additional pre-trade analytics, and order characteristics, such as the trade strategy, the broker and the actual algorithm that was employed. An example of the role of data inputs and of historical TCA measurement in the analysis of trade structure is illustrated in Exhibit 3.

Exhibit 3

Transaction Cost Inputs to Strategy Recommendations



The information flow begins with input from the execution management system (EMS). This includes the trading characteristics of individual names in the list, as well as submission times, performance benchmarks, and any specific trade instructions, such as urgency levels. Historical broker performance is matched with transaction cost distributions across alternative algorithms. User-defined parameters and historical cost information are complemented by inputs from pre-trade analytics, including, as in our previous examples, relative order size and volatility information.

The result is information that may be used to judge tradeoffs. A "best fit" strategy recommendation is a place on the trade structure continuum. At this level, the recommendation might be as simple as, for example, "unstructured liquidity search, based on a combination of pegging and discretion." More specific recommendations depend on the granularity of strategy information available by broker. The projected cost and risk of the general strategy are provided. Broker performance is matched against strategy recommendations, producing indicators of favorable and unfavorable broker choice for each strategy option.

Real Time Broker Measurement and Analysis of Algorithmic Trading Costs

Although historical TCA is valuable for analyzing performance trends, the immediacy of the trading environment often demands that information be readily available directly on the trade blotter. A fully structured approach integrates historical trading cost information with real time trading data to provide effective measurement, analysis and decision support tools for broker and algorithm choice. This in turn, provides its own short-term feedback loop of information.

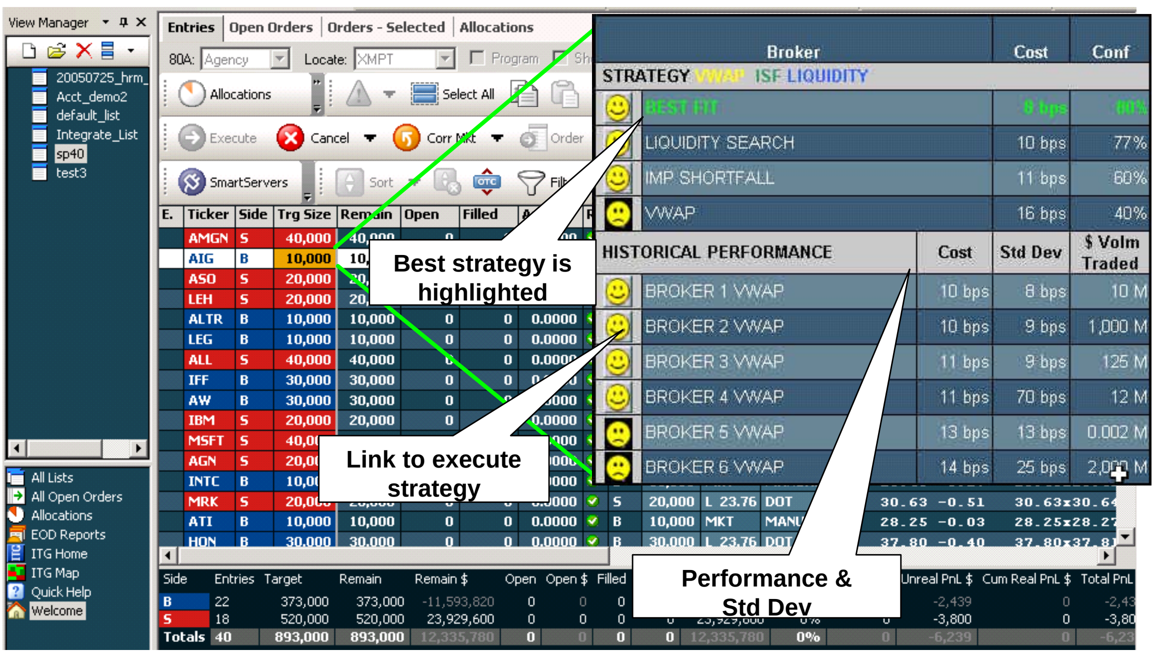
Integration of trading cost analysis tools into the trading blotter makes it possible to use real time inputs to dynamically refine algorithmic strategy selection. Real time inputs can be compared to historical norms for the stocks on the blotter, producing "alerts" that signal deviations from historical pre-trade cost trends. It is then possible to run multiple hypothetical strategies that measure the estimated cost and risk at different confidence levels to achieve the desired result and determine the algorithmic strategy that provides the "best fit" given current market conditions.

It should be noted that trade cost models do not actually specify the preferred strategy, although they may provide relatively clear guidance in that respect. It is a trader's decision to determine which of the set of cost/risk/confidence level tradeoffs is most appropriate given their appetite for price impact and opportunity cost risk.

Once the “best fit” strategy has been identified, it is possible to analyze the historical trading costs of how different broker algorithms have performed for different stocks at different order sizes. Figure 2 illustrates how this type of information might be incorporated into a trade blotter.

Figure 2

Trade Blotter Window with Historical Broker Algorithm Costs



This type of tool provides traders with information as to which brokers and algorithms have performed best for them with similar orders, displaying the historical cost and standard deviation of different algorithms for that stock and strategy. The intra-day trade cost results of the selected algorithm strategy also can be displayed as the executions occur, showing whether the strategy is successfully meeting its benchmark or whether the strategy should be modified to adapt to current market conditions.

Investment and the Human Element

Incorporating cost and risk controls into the algorithmic trading process is an investment. These controls and informational guidance must be tailored to fit a trading organization's style and goals, in terms of an overall portfolio of trading possibilities.

Measurement and analytics integration are only the first steps. The investment consists, in part, in educating the business side involved in the investment, not just the trading, process, and of structuring systems, which efficiently process data and feed information back to the decision-making support tools. It is easy to lose sight of the human element, when speaking of algorithmic trading. In our minds, this would be a critical error.

A structured approach to measurement and analysis in the algorithmic trading space is a piece of the overall portfolio management puzzle. It is an asset that is not easily replicated, even in a world of potentially commoditized strategies. Once put in place, the process can provide competitive advantage to managers across all phases of planning and execution, providing a unique asset that can produce superior investment performance over time.