Venue Analysis: What is it Good for?

"Venue analysis" has become a focus for buy-side traders concerned about execution quality. In this note we look at several popular venue analysis methodologies and show that they do not help traders understand or improve their performance. One factor appears to have a significant influence on execution quality: venue fee structure.

Background

One of the biggest changes in market structure since Reg NMS is the proliferation of exchanges, ECNs, and dark pools. This growth has come with a stream of negative press about questionable practices, from the SEC's censure of Pipeline, a dark pool, for trading through an affiliate against client flow while claiming that it provided "natural block liquidity," to Michael Lewis' *Flash Boys*. This headline risk has led to heightened focus on trading venues and how they can affect investors whose assets are traded there. It has also fed demand for tools to understand and monitor venues.

Another major change since Reg NMS is that the top-tier algorithmic providers all offer services of comparable quality, for example with VWAP shortfall of one or two basis points. Quantitative comparison of high-quality execution algorithms is beyond the reach of most firms, given real-world complexities such as limit prices on parent orders, as well as the massive number of similar trades they would need to execute to uncover small differences in performance. As a result, buy-side firms have turned to venue analysis as a proxy for execution quality.

The buy side takes two approaches to venue analysis. The first approach is to ask brokers to report on the venues where they route trades. These summaries typically include statistics like fill rate, percentage executed at each venue, and post-trade reversion.

The second approach is to collect raw data from brokers, fill by fill. In the FIX messages that brokers use to report on executions, for example, the execution venue is specified in tag 30. In principle, this allows for an independent, granular evaluation of routing behavior and venue quality.

In either case, the buy-side firm wants to evaluate the venues where its brokers are routing trades and to encourage those brokers to cut "bad" ones from the rotation. Instead of providing useful insight into venue quality, however, metrics commonly used to analyze venues largely reflect the behavior of algorithms and how a client uses them rather than the quality of those venues.

In this note we look at fill rate and pre/post-trade returns, and find that these two popular metrics do not provide any useful information about the intrinsic quality of a venue or of a provider's execution algorithms.

Fill Rate

Fill rate, defined as the fraction of shares sent to a venue that are actually filled, is widely assumed to reflect a venue's liquidity, which should be a key factor in determining venue quality. By looking at fill rates, a buy-side firm hopes to discover whether its broker is wasting time, incurring opportunity cost, or leaking information by sending orders to venues that don't add value.

While it has the virtue of simplicity, fill rate doesn't take into account how long a given order was resting at a venue unfilled or how small a fraction of market volume

FIGURE 1

	STRATEGY 1			STRATEGY 2			STRATEGY 3		
	SENT	FILLED	FILL RATE	SENT	FILLED	FILL RATE	SENT	FILLED	FILL RATE
Venue 1	300	300	100%	2000	2000	100%	4700	300	6%
Venue 2	400	400	100%	2000	2000	100%	4700	500	11%
Venue 3	300	300	100%	2000	2000	100%	4900	200	4%
Block venue	5000	5000	100%	0	0	N/A	15000	5000	33%

that this order (or the executed volume) represents.

Consider the following scenario: A client sends a 6,000-share order in a small-cap name to a broker's dark-pool aggregator. This aggregator has access to three "ordinary" dark pools that receive a regular stream of 200-share orders as well as to one "block" pool with a minimum fill size of 5,000 shares.

The algorithm can either send 5,000 shares to the block pool and 300 or 400 to each of the ordinary pools, or it can send 2,000 shares to each ordinary pool and skip the block pool altogether. The former approach could mean missing out on liquidity in the ordinary pools once those first orders are exhausted, but the latter risks missing a valuable block.

A hybrid strategy is better than either: The algorithm should distribute the 6,000 shares among the three ordinary pools but, once every five minutes (or however often it estimates that large blocks arrive in the block pool), pull 5,000 shares from the ordinary pools for a few milliseconds and send an order to the block pool as a test. If no order is waiting there, the algorithm can redistribute that 5,000-share block back to the ordinary pools.

The hybrid approach results in more orders being sent to all four venues, due to the momentary cancellation of existing orders in order to probe the block pool. These cancellations will be reflected in lower fill rates for each venue, even as the overall fill rate for the parent order is improved (Figure 1).

Fill rate does not say anything meaningful about venue quality because algorithms use various order placement strategies that have a much greater effect on fill rate than the venue does. In each of the three cases, we assume the identical arrival of contra orders at the venues, but get very different fill rates because the hybrid strategy pulses orders, effectively inflating the quantity sent. Fill rates at the venues also do not provide any meaningful measure of the overall execution quality of a liquidity sourcing algorithm. A trader who is interested in the overall execution quality gets little useful information from venue fill rates because the behavior of algorithms can vary significantly.

Pre-Fill Momentum and Post-Fill Reversion

Another popular and seemingly sophisticated approach to venue analysis looks at average pre- and post-fill returns. Intuitively, you don't want to see a run-up right before a fill, which would suggest that someone is gaming the market at this venue or perhaps the broker is pushing the stock, and you don't want to see reversion right after a fill, which would suggest gaming or adverse selection. Regardless of the cause, pre-trade momentum and post-trade reversion make the timing of a fill look unfortunate relative to other prices available in the market around the same time. If you can associate these pre- and post-trade patterns with a particular venue, then, other things being equal, you'd want to avoid that venue, right?

Unfortunately, when looking at fill data, things are rarely equal. The biggest determinant of the rate of pre- and post-trade returns is the set of rules that brokers use to make their timing and routing decisions.

Figure 2 shows actual pre- and post-trade returns for three different algorithms customized for Pragma clients with different trading objectives. You can see that pre- and post-trade returns vary significantly, but all of these executions happened **at the same exchange, using the same order type and during the same period of time**.

What could cause such different patterns of pre- and post-trade returns for the same order type at the same venue? Client B is trading a passive, VWAP-style strategy, and gets the vast majority of orders filled using passive limit orders (which are not shown on this chart). Note that the average return over a random four-minute window from the client's order is just a tiny fraction of the six basis points shown in the graph. The steep return shown leading up to and following an aggressive order is the result of a selection bias; the algorithm trades aggressively at precisely such times when passive orders have not been filled and it needs to catch up—that is, when the market has been trending away. Client A's algorithm, in contrast, places

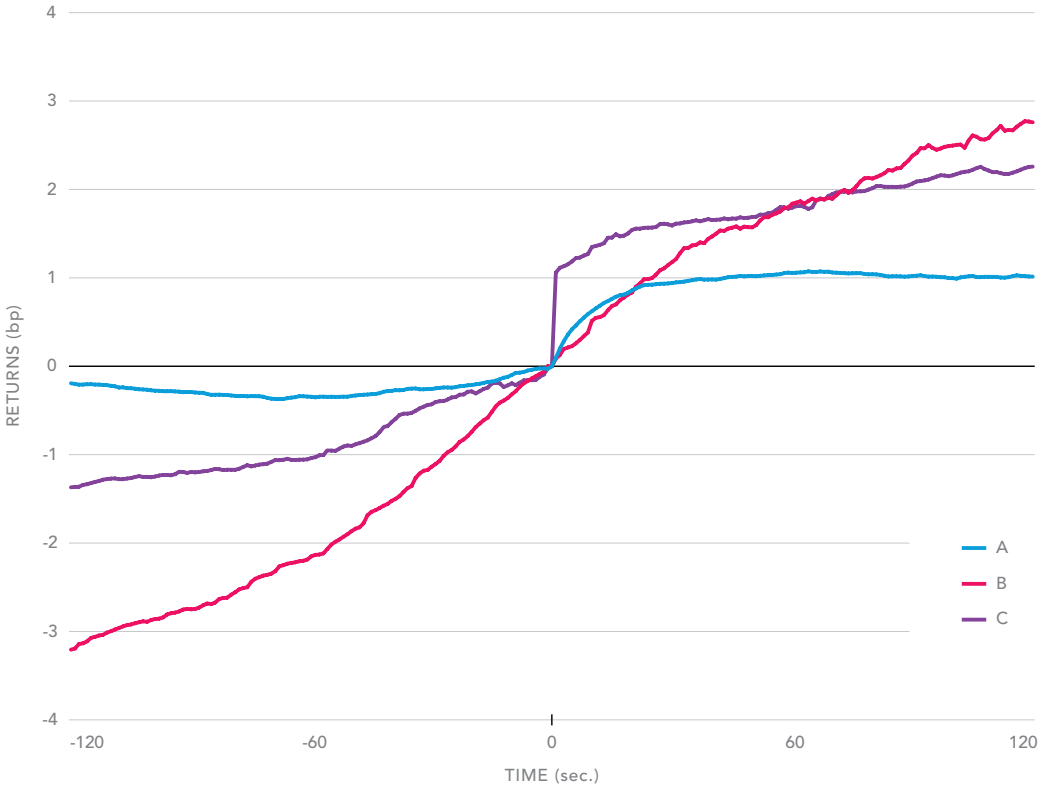

FIGURE 2
Pre- and Post-Trade Returns for three Different Customized Algos at One Venue – Aggressive Executions (Prices are the best offer).

aggressive orders for reasons exogenous to local price movement. The graph shows a roughly one-basis-point market impact of a small aggressive order, and no other price trend before or after the fill. Finally, Client C is using an aggressive opportunistic algorithm that uses aggressive sweeps, which often take out the price.

Through these examples, we can clearly see that pre- and post-trade returns are not properties of the venue, but rather artifacts of the order-placement behavior of the algorithm.

Randomized Experiments

Meaningful venue analysis requires full transparency and control of the algorithm's process for placing orders. A firm that controls its algorithmic tools can evaluate aspects of venue quality through controlled, randomized experiments. However, buy-side firms can't do this type of venue analysis using the fills returned by their broker's algorithms, since confounding biases can create an illusion of difference where none exists, as the graphs above illustrate. And quantitative comparisons of venues using the fills provided by multiple brokers are even more problematic, as different design choices made by the different brokers when building their algorithms create different confounding biases.

Venue Fee Structure

Is any accessible metric a quick proxy for venue quality? *In To Hop (The Queue) Or Not To Hop (The Queue) (Pragma Research Note 3, March 27, 2012)* we presented data from the kind of carefully controlled, randomized experiment described above (Figure 3).

The chart shows a clear difference in execution quality among several venues for passive limit orders. On average, a strategy that places a limit order on Exch III and then crosses the spread to "clean up" if the market ticks up before the order is filled will have an additional basis point of shortfall relative to the same order posted at Inverted Exch.

Exchanges that don't charge takers a big fee offer better execution quality on passive orders because they come earlier in the "inter-market queue." When two exchanges have offers at the same price, takers often very reasonably choose to take at the venue that charges less in exchange fees. This preference means that orders posted at these inverted venues effectively have priority, are less subject to adverse selection, and have less shortfall. But because takers at inverted venues aren't charged a big fee, providers at inverted venues don't receive a big rebate.

This is important because, as a recent Notre Dame

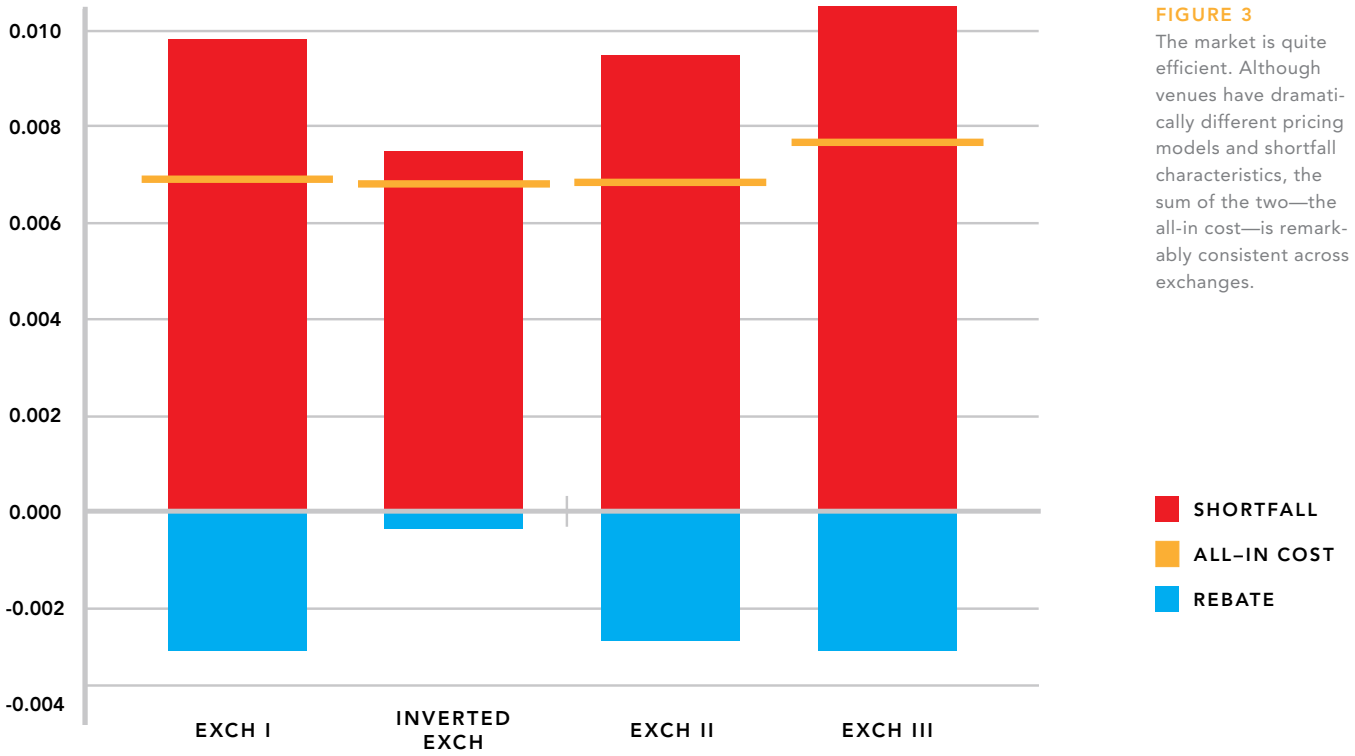

FIGURE 3
The market is quite efficient. Although venues have dramatically different pricing models and shortfall characteristics, the sum of the two—the all-in cost—is remarkably consistent across exchanges.

study[1] reported (and as is evident from inverted exchanges' small market share), most brokers make relatively little use of inverted exchanges like NASDAQ BX, presumably because they are too expensive. But as the research shows, brokers' efforts to maximize their rebates come at the expense of execution quality for their clients. Similarly, economic incentives govern brokers' use of their internalization pools, and these incentives may put them in conflict with their clients' best interests.

In other words, the most significant influences on venue quality are already well understood: venue fee structure and the patterns in routing behavior that those fee structures create.

Conclusion

The concept of venue toxicity has gained considerable traction with the buy side over recent years. However, many standard methodologies for analyzing venues are deeply flawed. Firms that understand and control their algorithms' behavior can engineer proper randomized experiments to get at venue quality, but such experiments are beyond the reach of most buy-side clients. Drawing conclusions about venues based on fills from broker algorithms without that understanding is a hopeless undertaking.

The biggest differences in quality among venues result from their economic models. Venues that charge takers the highest fees offer providers the worst execution quality. To serve its customers best, the buy side needs a thorough understanding of the relationship between fee structure and performance, and how economic incentives affect their brokers' routing decisions.

1 http://ssrn.com/abstract=2367462