Implementation Shortfall – One Objective, Many Algorithms

VWAP (Volume Weighted Average Price) has ruled the algorithmic trading world for a long time, but there has been a significant move over the past year toward using decision price, or implementation shortfall, algorithms. ITG®'s Hitesh Mittal explains why it is high time for this change in focus.

Perold (1988)[1] defines *implementation shortfall* as the difference in return between a theoretical portfolio and the implemented portfolio. When deciding to buy or sell stocks during portfolio construction, a portfolio manager looks at the prevailing prices (decision prices). However, due to a number of factors, the execution prices may be different from the decision prices. This can result in returns that differ from the portfolio manager's expectations.

Given the availability of such a simple measure as implementation shortfall, why has VWAP (Volume Weighted Average Price) remained the primary benchmark in algorithmic trading for so many years? What are the properties of algorithms such as Target Percentage of Volume that make them desirable for portfolio managers and traders?

This article explores the intuition behind these algorithms, and then examines the following aspects of implementation shortfall benchmarking in algorithmic trading:

- Why implementation shortfall is the optimal benchmark.
- Components of an implementation shortfall algorithm.
- How some popular algorithms (VWAP, Target Percentage of Volume) can be used to reduce implementation shortfall, and the shortcomings of these algorithms.
- How risk control helps implementation shortfall algorithms in reducing market impact and opportunity cost.
- Automatic and explicit risk control techniques in Algorithmic trading.
- Negotiating the conflict between anonymous liquidity and risk control.

VWAP: Going with the flow

VWAP's enduring appeal lies in its ease of attainability -- it is a moving target, and hence a more forgiving benchmark than arrival prices. With VWAP benchmarks, a trader or an algorithm models the volume distribution and then slices and dices the trades within a certain time interval on that distribution. As long as an algorithm does that, it is likely to achieve the VWAP over a given time horizon. This is true even if there are significant stock price moves during the day, either due to market impacts of the trading, or due to the stock's volatility.

---

[1] Perold, Andre F. (1988), The implementation shortfall: paper vs. reality, *Journal of Portfolio Management*, 14 (Spring), 4–9.

But, with the arrival price benchmarking, if a trader or an algorithm executes trades in size quickly, the ensuing market impact is likely to result in average execution prices that are worse than the arrival price. On the other hand, if the algorithm/trader executes the trades slowly (to reduce market impact), the volatility of the stocks may still result in execution prices that are far from the arrival price.

In short, a trader/algorithm's ability to achieve execution prices close to arrival prices depends on the size of the order, the liquidity available for the stock, and the volatility of the stock.[2] It is then appropriate to develop an algorithm that anticipates execution costs by looking at all these factors (order size, volume, volatility, and correlation). In fact, there are many such models that try to anticipate the cost of trading by factoring in these properties of trade-lists, including the ITG ACE® (Agency Cost Estimator) model. However, the appeal of the VWAP benchmark in part lies in the fact that it does not require such cost predictions.

VWAP vs. Arrival Price

So how does arrival price based algorithms stack up against VWAP algorithms in execution? The following charts plot profit and loss (P&L) for VWAP benchmark (in VWAP algorithm-chart 1). It is apparent from the chart that the available liquidity (trade sizes/total volume) do not seem to have much impact on the algorithm's ability to meet VWAP

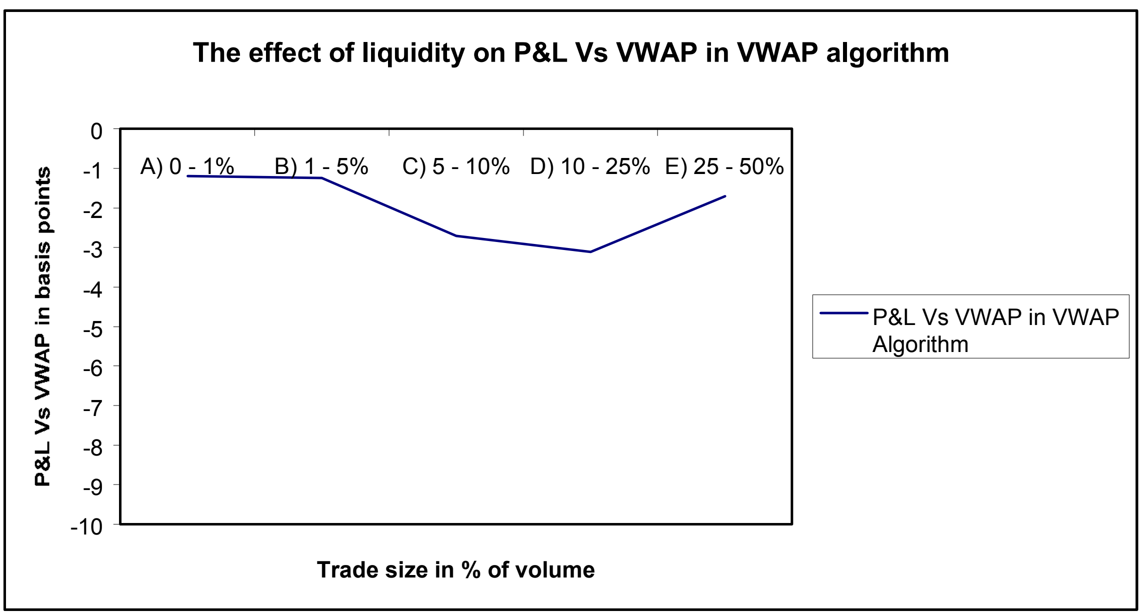

Chart 1: VWAP- P&L vs. percentage of volume
(Source: ITG)

---

[2] In addition, stock correlations also have an impact on the difficulty of achieving arrival price, as later explained in the article.

Implementation Shortfall algorithms do not have the same luxury. For obvious reasons there is a clear impact of liquidity on the performance Vs Arrival Price on Implementation Shortfall algorithms,i.e., the performance degrades as the trade sizes grow relative to the available liquidity (average trading volume).

VWAP as Implementation Shortfall

This article has discussed two important properties of VWAP algorithms so far--they are easily achievable and they are simple to measure. VWAP algorithms can also serve as good tools for arrival price or implementation shortfall (IS) algorithms, as long as the traders involved are not risk averse, as explained below.

One of the objectives of implementation shortfall algorithms is to reduce the market impact of the trade. This requires the algorithm to minimize the demand and supply imbalance at any given point during the trading time horizon. For example, assume that one has a period of three hours in which to trade. During that time horizon, the volume distribution is one million shares, as follows:

Average volume per hour:

Hour 1: 300,000
Hour 2: 400,000
Hour 3: 300,000

Suppose that one has a trade for 100,000 shares. To reduce the market impact, one intuitively wants to minimize the demand/supply impact at any point. To do that we will have to trade as follows:

Hour 1: 30,000 Demand/Supply: 10%
Hour 2: 40,000 Demand/Supply: 10%
Hour 3: 30,000 Demand/Supply: 10%

In addition to being intuitive, this is exactly how a VWAP algorithm will trade during this period. Thus, if a portfolio manager asks a trader to use the VWAP benchmark, not only can he or she measure the trader's performance easily (by checking VWAP), he or she can also make sure that the trades have very low market impact.

Below is a chart that plots the performance of VWAP and IS algorithms (in basis points away from arrival price) for trades with different liquidity requirements:

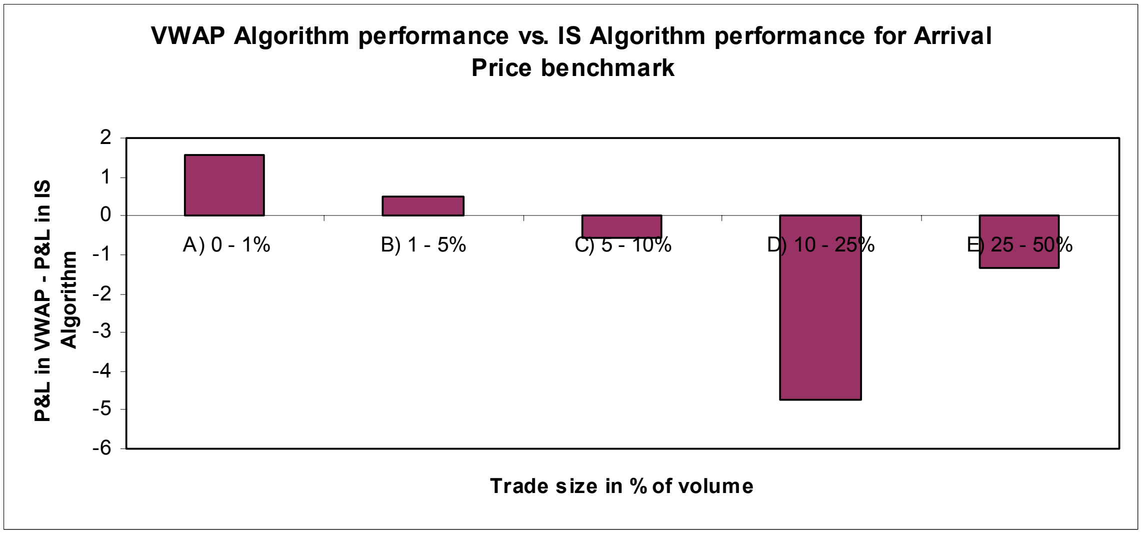

Chart 3: Comparison of P&L (vs. Arrival Price) in VWAP algorithm & IS Algorithm plotted against percentage of volume. (Source: ITG)

The chart explains that the performance in VWAP Algorithm is similar to the performance in Implementation Shortfall algorithm. One should also note that the demanded liquidity impacts both algorithms in a similar fashion. This explains in part why VWAP algorithms have remained so popular: they are easy to measure, easy to execute, and can produce similar results on average. However, this does not take into account the potential hidden costs of using a VWAP algorithm for implementation shortfall, as discussed below.

Problems with VWAP algorithms for Implementation Shortfall

1) Potentially high opportunity costs

Traders and Portfolio Managers care about average costs vs. arrival price, and they also care about the consistency of the cost. A passive algorithm such as VWAP can ensure a good average price vs. arrival price, but it cannot ensure consistency. To illustrate that, one can plot the standard deviation of the P&L vs. arrival price of trades against the percent of volume for VWAP and implementation shortfall algorithms.

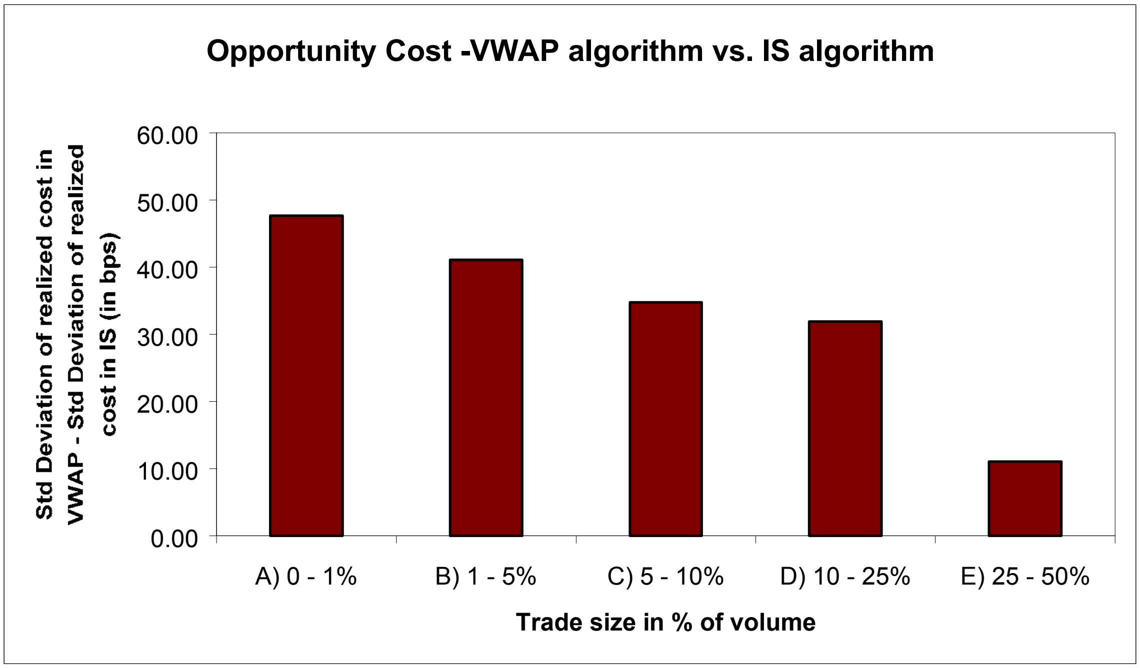

Chart 4:
Comparison of Standard deviation of cost (vs. Arrival Price) of VWAP algorithm & IS Algorithm plotted against percentage of volume. (Source: ITG)

One can clearly see that VWAP does a poor job compared to implementation shortfall algorithms in terms of consistency of performance vs. arrival price as the trade size/volume goes down. As the trade size/volume rises above roughly 20% of ADV, the differences are not very significant.

Example 1:

One can understand this behavior by way of an example: suppose MSFT trades 30 million shares on an average day. If a trader has three million MSFT shares to trade, a VWAP algorithm may be appropriate. However, if the trader gets 30,000 shares of MSFT to trade, then the savings of market impact (by spreading the trade over the whole day) is not significant compared against the opportunity cost the trader could save by trading the stock within the next few minutes.

2) VWAP can be influenced

It is possible to beat a VWAP benchmark, by conducting trades in a manner that may actually lead to increasing the trading impact. In general, any benchmark that has future price as a component can be influenced. Closing price and VWAP are examples of such benchmarks.

VWAP in its simplest form can be written as:

VWAP = Pt1.Vt1 + Pt2.Vt2 +….Pt(n).Vt(n) / Total volume
Where Pt(n) is the volume weighted average price of interval n and Vt(n) is the volume in interval n.

Average Price = APt1.St1 + APt2.St2 … Apt(n).St(n) / Total shares

If a trader overloads (buys/sells more than the volume distribution) in the first few bins (for example, t=1 and t=2), this will result in market impact and can affect prices in the later bins. Since the trader is frontloading the distribution, St1/Total shares and St2/Total shares will be higher than Vt1/Total shares and Vt2/Total shares. Similarly, Pt1 and Pt2 would be better than Pt3….Pt(n), and the average execution price would be better than the VWAP price. One can clearly see that even though a trader can beat the VWAP benchmark by influencing the VWAP itself, this will have a larger impact overall and will result in higher cost versus the arrival price.

In our opinion, it is not sufficient to examine performance vs. VWAP alone. In evaluating a VWAP algorithm's performance, a portfolio manager or trader should also compare the algorithm's execution price to the arrival price.

3) Risk aversion can focus on VWAP rather than on Arrival Price

Traders/VWAP algorithms may become risk averse to the VWAP benchmark and thus may not seize upon opportunities that could lead to a more optimal outcome, i.e. a reduction in implementation shortfall. For example, alternative trading systems such as POSIT® allow a trader to execute shares with no market impact. But by doing so, an algorithm might deviate from the VWAP price, and it may not take that chance, even if it would be the optimal thing to do from an arrival price perspective. Thus, one can see that VWAP is often not an optimal strategy from an average performance perspective, even if a portfolio manager or a trader has very low risk aversion with regard to the arrival price.

Beyond VWAP

Using Volume Participation Algorithms in Implementation Shortfall

As discussed, VWAP algorithms suffer from high opportunity cost (in terms of standard deviation of execution price vs. arrival price) for trades that represent a low percentage of ADV. Use of a participation algorithm can help solve that problem. A participation algorithm is similar to a VWAP algorithm except that it uses a constant participation rate. So, if a trader believes that he or she can live with the impact that will be caused by 10% participation, then he or she can simply use a 10% participation algorithm.

Returning to the MSFT example discussed above, if a 10% participation algorithm is used instead of VWAP, trading three million shares of MSFT (with 30 million shares average daily volume) will have results similar to using a VWAP algorithm. But, in contrast, 30000 shares of MSFT would be traded in a few minutes, with little deviation likely from the arrival price.

Implementation Shortfall – Average Cost and Opportunity Cost

As discussed, implementation shortfall algorithms are mainly concerned with balancing market impact and opportunity cost. This section will examine the factors that influence market impact and opportunity cost.

Average Cost: this article has discussed intuitive ways (Volume participation, VWAP) to reduce average costs. Volume participation is only a first order approximation of a cost-optimal strategy, and there are other factors that play a role in minimizing the demand and supply imbalance. Order type placement is one such factor. For instance, the spread and temporary price impact is higher at the beginning of the day because there is much more uncertainty of the future price of a security. Consequently, traders place more limit orders at the beginning of the day than towards the close. Liquidity suppliers are more careful early in the day, and often charge more or try to get a risk premium. An implementation shortfall algorithm should model these factors along with the volume distribution to determine an optimal trading distribution with minimum market impact.

Opportunity Cost: Opportunity cost can be defined as the standard deviation of the trading cost. This is a function of trade distribution, stock volatility, volatility distributions and the correlation among stocks on a trade list over a given time horizon.

By assuming normal distribution, traders can use opportunity cost and average cost numbers to determine various “if-then” scenarios. For example, traders can determine the probability of a strategy costing more than a certain number of basis points, or determine the worst-case scenario 95% of the time for a given strategy.

ITG Algorithms and Implementation Shortfall

ITG algorithmic trading products – ITG SmartServer and ITG HorizonPlus provide Implementation Shortfall algorithms that model all these factors and provide a trader with the solutions that have the least opportunity cost for a given trading cost and the least trading cost for a given opportunity cost. ITG HorizonPlus provides these options to a trader in the form of trading efficient frontier (described in the next section) so that a trader can choose his/her tradeoff easily. These algorithms adjust themselves by looking at real time conditions and thus make the best use of historical and real time information. These algorithms control risk implicitly during the optimization process. ITG HorizonPlus also provides many ways to explicity control the residual risk of the trading list during trading. Risk control becomes an important factor in reducing Implementation Shortfall as this article illustrates in the next few paragraphs.

Implementation Shortfall - Trading Efficient Frontier

If an algorithm trades on certain distributions over a given time horizon, then the average cost and standard deviation of that cost can be predicted. Many distributions can lead to the same execution cost. Each of these distributions will have a different standard deviation of the cost. For a given execution cost, an optimal distribution is the trading distribution that has the lowest standard deviation. Many such optimal distributions can yield pairs of costs and standard deviations and these can be plotted in a chart, (such as the one below):

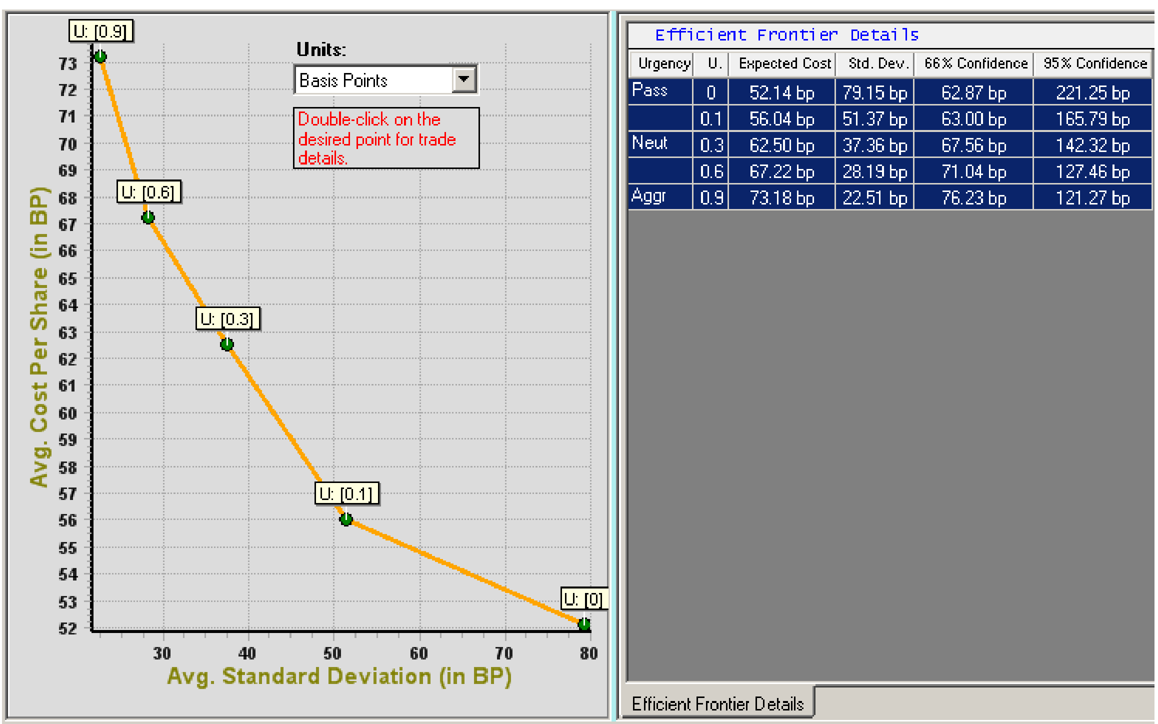

Chart 6: Efficient Frontier (Source: ITG)

Note that the trading efficient frontier is downward sloping -- the inverse of the efficient frontier, which one is accustomed to in modern portfolio theory. That is because this is not plotting risk vs. return, instead this is a plot of opportunity risk vs. average execution cost. The higher the opportunity cost, the lower the average execution cost.

Implementation Shortfall and Risk Control

In trading, risk control refers to keeping the risk of the unexecuted list low. Risk control has a very useful effect on implementation shortfall algorithms. This article has discussed the different components that affect the cost of trading and opportunity costs. Risk control generally reduces the opportunity cost of trading and thus provides a trader or an algorithm with an extended time horizon in which to trade. The extended time horizon is an opportunity to reduce the market impact of these trades, resulting in smaller overall trading costs. This is illustrated below for a two-sided (both buy and sell order) list as well as for a single-sided list.

If one has a two-sided list and the two sides are highly correlated, then each side acts as a hedge for the other. Thus, even though the volatility of each stock will increase the opportunity risk, a strong and negative correlation (positive, but on different sides) will reduce the risk of keeping (not trading) those names. An algorithm can take advantage of this effect to increase the time horizon in which that trade list can be traded. Increasing the time horizon will lead to reduction of the cost. Thus, by controlling the risk, an algorithm can reduce the trading cost as well as the opportunity cost. The higher the correlation between the buy- and sell-side baskets, the lower the risk of the overall list and the longer the time horizon can be extended.

To use another example: suppose one has a two-sided list where the buy side represents $30 million and the sell side basket represents $5 million, and both sides have very high correlation with each other. In this case, an implementation shortfall algorithm should front-load the buy-side basket and back load the sell-side basket. Front loading the buy-side basket will reduce the opportunity cost and back loading the sell-side basket will reduce the market impact as well as the opportunity cost (by hedging against the buys).

The same concept applies if the list is single-sided -- an algorithm will determine a different optimal distribution if it is aware of the correlations of the stocks. For example, if there are multiple sectors, it may be important to keep the sector imbalance low at any point in time, because the correlation between the stocks from different sectors will be low.

Implementation Shortfall and Hedging

With a single-sided list there are other ways to reduce the risk and thus increase the trading time horizon. One common way is through hedging, (if the trader can find a liquid instrument which tracks the trading list). The liquidity of this instrument is crucial, otherwise trading that instrument itself will result in undesirable market impact.

For example, if a buy-side trade list has very low tracking error to the S&P 500, the trader can sell S&P 500 futures to hedge that list (provided that trading S&P 500 futures will have much less impact than trading each stock in the list). Once hedged, a trader can extend the time horizon of the list and start closing out the future positions as the algorithm trades the underlying list. In this case, a trader would want to specify the objective of keeping the tracking error against the S&P 500 to a minimum during the whole trading horizon.

ITG algorithms provide mechanisms for traders to indicate such explicit risk objectives. Specifically, ITG ResRisk+SM provides traders with a way to run such optimization scenarios manually.

Risk Control: Explicit or Automatic

In the section "**Implementation Shortfall and Risk Control"** it was observed that an algorithm could control risk automatically with an optimal distribution by looking at the stocks' correlations and volatility. It was also observed in the section "**Implementation Shortfall and Hedging"** that a trader may want to provide explicit risk objectives when hedging with other instruments. There are other cases in which traders may want to specify explicit risk control objectives for an algorithm. Some common cases are as follows:

1) Instead of relying on the risk models used by an algorithm, some traders prefer more intuitive approaches to risk control. For example, traders might specify in their objectives that the algorithm should keep the entire list dollar-neutral. Traders might additionally want to specify that the algorithm should keep the list dollar-neutral for a specific sector or sectors in the list.

2) Traders/PMs might want to control the risk of new holdings as a result of new executions. Tools such as ITG ResRisk+SM provide traders and portfolio managers with the ability to control the risk of holdings while an algorithm is still executing the list.

Anonymous Liquidity and Risk Control:

Anonymous block liquidity pools such as POSIT® present some very useful options to traders and algorithms looking to reduce market impact. Unlike the liquidity available in the exchanges, POSIT®'s liquidity can be accessed without information leakage. Thus, it creates zero market impact, freeing the algorithm/trader to deal with trades that are a smaller percentage of daily volume. But if a trader or an algorithm is following one of the risk control methods discussed above, then they may question whether trading a particular stock will increase or reduce the overall risk of their list.

A simple rule of thumb is that for a single-sided list, trading a stock will always reduce the overall risk of the list (unless the trader is hedging the list with another instrument). If the list is two-sided, trading a stock may increase or reduce the risk of the overall list. Both of these points are explained in Appendix 1. An algorithm or a trader should be able to determine if the reduction in cost for trading that stock will compensate for the potential increase in risk. For example, POSIT® provides an interface that can be used to take advantage of such tradeoffs without letting an algorithm see the liquidity available inside.

Note that when increases/decreases in risk are discussed, absolute terms (dollars) are used rather than percentages. Risk could go up percentage-wise and at the same time drop in dollar terms, because of a reduction in the number of shares remaining on the trade list.

Risk Control and Number of Stocks:

All the risk control techniques defined in this section are useful for a large trade list. As a general rule, one starts seeing the positive effects of risk control as the number of stocks on the list climbs above 20.

Conclusion

Various algorithms can be used to reduce implementation shortfall, from traditional ones such as VWAP and volume participation, to more sophisticated algorithms such as list-based or risk-based implementation shortfall algorithms. While VWAP and volume participation algorithms provide intuitive solutions, they lack certain properties that can be used to reduce the market impact and the opportunity costs of trading.

An ideal implementation shortfall algorithm should model the optimal trade distribution by looking at the liquidity profile, trade sizes, volatility of stocks, volatility distributions of stocks, spread distribution of stocks, and especially the stock correlations. Risk control helps an algorithm increase the trade time horizon and thus reduces the market impact while lowering the opportunity cost. For a risk-controlled list, traders can also hedge with liquid instruments, (as long as the algorithms they use permit them to specify explicit risk control objectives). Traders may also prefer specifying explicit risk control objectives if they favor an intuitive risk control approach over automatic optimizations.

Finally, anonymous block liquidity sources such as POSIT® can be extremely helpful in reducing implementation shortfall. The use of risk control may be helpful while accessing such liquidity, particularly in the trading of two-sided lists.

Notes:

1) The data displayed in the above charts are random samples from the database of ITG execution algorithms.

Appendix 1.

If a trade list is single-sided, then even though a stock may reduce the risk of the overall list in terms of percentage, it will always increase the risk in terms of dollars. We can illustrate this with a two-stock example, which can further be generalized for more than two stocks:

Let us assume that in most cases, a correlation between two stocks is positive. Assume we have stocks S1 and S2 that have very low correlation of $\rho 12$. The volatility of S1 and S2 are $\sigma 1$ and $\sigma 2$. We have D1 dollars of S1 to buy and D2 dollars of S2 to buy.

The total risk of the list would be

$$(D1\ \sigma 1)^2 + (D2\ \sigma 2)^2 + 2\ D1\ D2\ \sigma 1\ \sigma 2\ \rho 12$$

Even if $\rho 12$ is very low, since it is always positive, reducing D2 or reducing D1 will always reduce the overall risk of the list.

If S2 is a stock to be sold then we can write the risk as

$$(D1\ \sigma\ 1)^2 + (D2\ \sigma\ 2)^2 - 2\ D1\ D2\ \sigma 1\ \sigma 2\ \rho 12$$

In this case reducing D2 can either reduce the risk or increase the risk depending on D1, D2, $\sigma 1$, $\sigma 2$ and $\rho 12$. For example if $\rho 12$ is very low then reducing the dollar amount of any stock will always reduce the total risk but if it is very high then reducing the dollar amount of one stock only may actually increase the overall risk.