I-Star Model
Description of Market Impact Model from OTS

Robert Kissell
Kissell Research Group

Draft
Revised January 2013

Summary:
This paper provides an overview of the market impact modeling methodology developed in Optimal Trading Strategies and further expanded in The Science of Algorithmic Trading. It is prepared as a discussion piece to share ideas. Included below is background pertaining to the market impact model, the underlying datasets, parameter estimate techniques, and ways to evaluate performance. All comments, suggestions, feedback is welcome.

Paper Outline

1. I-Star Impact Model (all units are in basis points)

\[ I^* = \hat{a}_1 \cdot \left( \frac{S}{ADV} \right)^{\hat{a}_2} \cdot \sigma^{\hat{a}_3} \]

\[ MI = \hat{b}_1 \cdot I^* \cdot POV + \left(1 - \hat{b}_1\right) \cdot I^* \]

\[ TR = \sigma \cdot \frac{1 \cdot 1 \cdot S \cdot 1 - POV}{250 \cdot 3 \cdot ADV \cdot POV} \cdot 10^4 bp \]

2. Derivation of the Model

Source:
3. Underlying Dataset
4. Estimating Parameters
5. Forecasting Costs
6. Back-testing the model
7. Advanced Modeling Techniques

I. Market Impact Model

The market impact model (I-Star) is based on a cost allocation approach where participants incur costs based on size of their order and the overall participation with market volumes. The model is broken down into two components: 1) Instantaneous Impact denoted as I-Star or $I^*$ and represents the theoretical cost of the order if the entire order was released to the market in its entirety, and 2) Market Impact Cost which represents the cost for the specified trading strategy. Market impact further consists of a temporary and permanent impact term.

These equations are:

$$I^* = a_1 \cdot \left( \frac{Q}{ADV} \right)^{a_2} \cdot \sigma^{a_3}$$

$$MI = b_1 \cdot I^* \cdot POV + (1 - b_1) \cdot I^*$$

1.1 Instantaneous Impact Equation (I-Star)

$$I^* = a_1 \cdot \left( \frac{Q}{ADV} \right)^{a_2} \cdot \sigma^{a_3}$$

Instantaneous impact cost (I-Star) represents the theoretical cost of an order if all shares were released to the market in its entirety at once. This component can also be thought of as the payment that is required to attract additional sellers or buyers into the marketplace. For example, the premium buyers must provide or the discount sellers must grant in order to complete the order within their specified timeframe.

The variables of the instantaneous impact equation are:

Source:
Market impact represents the cost that is expected to be borne by the trader based upon the underlying execution strategy, e.g., percentage of volume (POV), trade schedule, etc.

Market impact consists of a temporary and permanent component. Temporary impact represents the liquidity cost or urgency needs of the investor. Temporary impact cost is higher for the more urgent strategies (such as a high participation rate) and lower for the less urgent strategies (such as VWAP or a lower participation rate). Trading risk, however, is higher for the less urgent strategies and lower for the more urgent strategies.

Permanent impact represents the information leakage of the trade and is an unavoidable cost of business. Permanent impact will not change based on the underlying strategy or POV rate. Permanent impact can also serve as an indicator of market sentiment or market efficiency (or at least the perception of each).

The variables of the market impact equation are:

\[ I^* = \text{instantaneous impact (from eq. 1.1)} \]
\[ POV = \text{percentage of volume} \]
\[ b_1 = \text{percentage of temporary impact (} 0 \leq b_1 \leq 1 \) \]

II. Derivation of the Model

Consider a situation where buyers have \( V \) shares to buy and sellers have \( V \) shares to sell – both within the same time period and urgency needs. In this case we have an equilibrium condition and expect there to be \( V \) shares traded in the marketplace without any extraordinary price movement because the number of buyers and sellers is equal (but there may be some price movement due to market, natural alpha, or noise).

Now suppose that another participant (participant “A”) enters the market with an order to buy \( Q \) shares over the same time period and with the same urgency needs. This creates a buy market imbalance equal to the \( Q \) shares.

Source:
In order for these additional $Q$ shares to execute buyers will have to offer a premium to attract additional sellers. Let’s define this premium to be $I^*$. We can think of this quantity as the premium or mark-up required to attract enough sellers in order for all $Q + V$ shares to trade.

This amount, $I^*$, however, will not be borne only by participant “A” but rather shared (allocated) across all buyers. We can think of this approach as an average costing methodology.

Our temporary impact parameter ($b_i$ from equation 1.2) represents the percentage of total temporary market impact (e.g., the percentage of $I^*$ which is temporary). Thus, the permanent impact of the incremental imbalance of $Q$ shares is $(1 - b_i)I^*$.

The total temporary impact of these $Q$ shares is $b_iI^*$. But again, it is not fair to assume that this entire cost will be borne entirely by investor A. Rather, it is more appropriate to assume that all buyers will incur a portion of this quantity based on their overall trading activity.

Since we now expect there to be $Q + V$ shares traded, that is, the original $V$ shares plus the newly arrived $Q$ shares, the portion of total temporary impact expected to be borne by investor A is calculated in proportion to her total trade volume. This is:

$$\text{1.3 } \frac{Q}{Q + V}$$

This expression also represents the percentage of volume (POV) rate since $V$ is the number of shares expected to trade in the market without the order and $Q$ is the incremental number of shares from the order. Thus, $Q + V$ is the total number of shares expected to trade in the market.

Therefore we have,

$$\text{Permanent Impact} = (1 - b_i)I^*$$

$$\text{Temporary Impact} = b_iI^* \frac{Q}{Q + V}$$

Or alternatively,

$$MI = b_iI^*POV + (1 - b_i)I^*$$

Source:
We can also see from equation 1.3 that if participant A transacts more aggressively, say in a shorter period where only $\frac{1}{2} \cdot V$ shares are expected to trade, the market impact cost allocated to her will now be:

$$\frac{Q}{Q + \frac{1}{2} \cdot V}$$

which is a higher percentage than previous.

And if she trades over a longer period of time where $2V$ shares are expected to trade, market impact cost allocated to her will be:

$$\frac{Q}{Q + 2V}$$

which is a smaller percentage than previous.

This example helps illustrate that market impact cost is directly related to the urgency of the strategy (quicker trading will incur higher costs on average than slower trading which will incur lower costs on average). This also illustrates that the temporary impact function has a non-linear relationship with percentage of volume (plot temporary impact as a function of POV rate). (Trading risk, on the other hand, will be lower for the more urgent orders and higher for the more passive orders, e.g., trader’s dilemma).

Up to this point we have still not yet defined the functional form of $I^\ast$. Experience, empirical evidence, and observations have found that trading cost is dependent upon size, volatility, and strategy. Thus, our functional form needs to include at least these variables.

[Some alternative models have included market cap or other stock fundamental factors to differentiate between costs for different stocks even for the same relative size, e.g., 5% adv. Our analysis has found that volatility provides a better fit than variables such as market cap, log of market cap, etc. – give example.]

Our preferred functional form for $I^\ast$ is the following power function:

$$I^\ast = a_1 \left( \frac{Q}{ADV} \right)^{a_2} \sigma^{a_3}$$

Notice that the function form includes parameters $a_1, a_2, a_3$ so that we do not force any preconceived notions onto the model such as a square root function with size or a linear

Source:
relationship with volatility (give examples of each). These parameter values are derived from our underlying dataset.

Other functional forms of $I^*$ have been proposed and tested, but the power function above has been found to be the most robust, stable, and accurate over time. One important question that often arises regarding this model is that since the parameters $a_1, a_2, a_3, b_i$ are estimated across a dataset of stocks and are the same for all stocks, how do we differentiate trading cost across stocks (for the same size orders %adv)?

This is explained as follows. The instantaneous impact equation for a particular stock $k$ is:

$$ I^*_k = a_1 \left( \frac{Q}{\text{ADV}_k} \right)^{a_2} \sigma_k^{a_3} $$

Rewrite this expression as follows:

$$ I^*_k = a_1 \cdot \sigma_k^{a_3} \left( \frac{Q}{\text{ADV}_k} \right)^{a_2} $$

Now we have an expression with a stock-specific sensitivity component and a universal shape component. Therefore, the functional form of this model has a stock-specific sensitivity term determined from its volatility (e.g., as a proxy for price elasticity) and a universal shape term (which has been shown to be reasonable).

### III. Underlying Dataset

The underlying dataset for the I-star market impact model is comprised of derived tic data and calculated end of day data. Several trading intervals are used for the analysis such as a full day, half-hour or hour intervals, morning and afternoon, etc. [We can discuss this in more detail as needed]

We have tried fitting this model in the past using actual order data but have not had much success. This is believed to be due to 1) survivorship bias - investors allow orders that are trading well (inexpensive) to continue to trade and cancel those orders that are underperforming (expensive), 2) there is often a large concentration of small orders resulting in skewed/biased results, 3) lack of complete datasets - B/D’s are not usually privileged to the full intentions of the portfolio manager and often only observe day orders from the fund (the fund may give a large multi-day order to different brokers each day in order to disguise their trading intentions).

Source:
Some summary notes:

- It is definitely worth discussing and revisiting the underlying datasets to see where everyone has had successes and difficulties.
- Some funds have requested information on this model in order to fit the parameters to their own trade and order data.
- Actual order day is used to back-test and evaluate the model fitted from tic data.

The variables used in the model are defined as follows:

\[ Q = \text{trade imbalance} - \text{the difference between buy initiated and sell initiated trades in the trading interval. All trades are designated as either a buy-initiated or a sell-initiated trade (using a modified Lee & Ready tick rule) based on the perceived urgency needs of the trade. That is, up tic or zero-up tic is denoted as a buy-initiated trade and a down tic or zero-down tic is denoted as a sell-initiated trade. An alternative technique to the modified Lee & Ready tick rule is to map each trade to the market quote at the time of the trade to determine if the trade was buy or sell initiated.} \]

\[ Side = \begin{cases} 
1 & \text{if positive imbalance} \\
-1 & \text{if negative imbalance} 
\end{cases} \]

\[ V^* = \text{the actual volume that traded in the interval.} \]

\[ ADV = \text{the 30 day average daily volume (computed over exchange hours).} \]

\[ \sigma = \text{annualized volatility expressed as a decimal (e.g., 0.20 and not 20%) computed as the standard deviation of log returns (natural logs) over of the previous 30 days. Preference now is to incorporate an adjustment based on the current VIX index [give example of the technique to forecast volatility]} \]

\[ POV = \frac{Q}{V^*}, \text{the computed market imbalance divided by total volume that traded in the period (such as the day, hour or half-hour interval).} \]

\[ \overline{P} = \text{VWAP price over the interval} \]

\[ P_0 = \text{Price of first trade in the interval} \]
\[MI = \text{Side} \cdot \frac{P - P_0}{P_0}\]

IV. Estimating Model Parameters

Estimating the parameters of this model \((a_1, a_2, a_3, b_1)\), unfortunately, is not a straight forward process because the model is non-linear and parameters \(a_1\) and \(b_1\) are extremely sensitive.

For example, we have:

\[
I^* = a_1 \cdot \left(\frac{Q}{\text{ADV}}\right)^{a_2} \cdot \sigma^{a_3}
\]

\[
MI = b_1 \cdot I^* \cdot \text{POV} + (1 - b_1) \cdot I^*
\]

But after substitution we have:

\[
1.4 \quad MI = b_1 \cdot a_1 \cdot \left(\frac{Q}{\text{ADV}}\right)^{a_2} \cdot \sigma^{a_3} \cdot \text{POV} + (1 - b_1) \cdot a_1 \cdot \left(\frac{Q}{\text{ADV}}\right)^{a_2} \cdot \sigma^{a_3}
\]

From eq. 1.4 we see that it is difficult to isolate \(a_1\) from \(b_1\) and non-linear optimization techniques are very sensitivity to convergence technique, starting solution, and value of \(b_1\). The best we can do directly is to estimate the product \(k_1 = b_1 a_1\) but this will not allow us to determine the effect of \(b_1\) or \(a_1\) individually and would also require a non-linear constraint.

To resolve this issue, we estimate the parameters following a two step process. First, estimate the temporary impact parameter \(\hat{b}_1\). Second, substitute this value into equation 1.2 and estimate parameters \(a_1, a_2, a_3\) based on the estimated value \(\hat{b}_1\). (Other techniques to solve the parameters are welcome).

IV.1 Estimate \(\hat{b}_1\)

Parameter \(b_1\) has been estimated using two different techniques.

1) Run a regression on the following equation

Source:
\[ \text{MI} = \alpha_0 + \alpha_1 \cdot \text{POV} \]

Estimate \( b_1 \) from the estimated alphas:

\[ \hat{b}_1 = \frac{\hat{\alpha}_1}{\hat{\alpha}_0 + \hat{\alpha}_1} \]

[To do: show a graphical example of this technique]

Note: when running the above regression be sure to aggregate by order sizes. For example, run the regression for sizes 1-5%, 5-10%, etc. This will eliminate the correlation between size and POV rate.

2) Make an educated guess for \( b_1 \) and then solve for \( a_1, a_2, a_3 \) through an iterative process. We then determine which value of \( b_1 \) has the best fit. I have used the R2 and/or non-linear R2 statistic for this purpose. For example, solve for \( a_1, a_2, a_3 \) using values of \( b_1 = 0.90, \ldots, 0.99 \). For each iteration, record each parameter and the non-linear R2 statistic. After all the iterations are run determine which has the best fit and use those parameters.

IV.2 Estimate \( \hat{a}_1, \hat{a}_2, \hat{a}_3 \)

The process to estimating the parameters \( \hat{a}_1, \hat{a}_2, \hat{a}_3 \) is as follows.

Start with equation 1.2 and substitute \( \hat{b}_1 \) for \( b_1 \).

\[ \text{MI} = \hat{b}_1 \cdot I^* \cdot \text{POV} + (1 - \hat{b}_1) \cdot I^* \]

Next factor the equation as follows:

\[ \text{MI} = I^* \cdot \left( \hat{b}_1 \cdot \text{POV} + (1 - \hat{b}_1) \right) \]

Divide through by \( I^* = \left( \hat{b}_1 \cdot \text{POV} + (1 - \hat{b}_1) \right) \)

\[ \frac{\text{MI}}{I^*} = 1 \]

and rewriting \( I^* \) in its full functional form gives:

Source:
\[ \frac{MI}{\ast} = a_1 \left( \frac{Q}{ADV} \right)^{a_2} \sigma^{a_3} \]

Model parameters \( \hat{a}_1, \hat{a}_2, \hat{a}_3 \) can now be estimated via non-linear regressions techniques such as non-linear least squares or maximum likelihood estimates. In fitting the equation above be cautious of heteroscedasticity – it is often helpful to make a correction for this error term. (We can discuss these techniques used during the estimation of the data if needed). Since the market impact term on the LHS can be negative we are not able to do a log transformation at this step (but see below).

One method that has been used in the past to fit these parameters is to first group the data into size and volatility buckets and to compute the average market impact of each category. This grouping of data is helpful for analysis because it eliminates the noise (better shows a pattern) and also provides a balanced data set (e.g., the underlying sample data is not skewed to certain sizes). The last part is that when we do group the data, we should filter out those categories with only a small number of samples. For example, require at least 10 or 20 data point for each grouping bucket to be included in the estimation analysis.

After grouping the data, it is often the case that the LHS of the equation will have all positive values (or only a few categories with negative values). In these situations we can perform a log transformation of the data. If all LHS data points are positive then we can transform all data points. If there are a few grouping categories with negative values we can transform all the positive LHS records and either eliminate the groupings with negative values or give these records a transformed LHS value of say -3 or -5 (the exact value will depend upon the positive LHS values and units).

Another benefit of performing a log transformation is that this transformation will correct (approximately) for the heteroscedasticity effect – no other adjustment is required.

The log transformation of the model is as follows:

Start with:

\[ \frac{MI}{\ast} = a_1 \left( \frac{Q}{ADV} \right)^{a_2} \sigma^{a_3} \]

Take the natural log of each side:

\[ \ln \left( \frac{MI}{\ast} \right) = \ln(a_1) + a_2 \cdot \ln \left( \frac{Q}{ADV} \right)^{a_2} + a_3 \cdot \ln(\sigma) \]

Source:
This regression can now be solved via OLS – which is nice and direct and has easy statistics to interpret.

In effect, we are actually solving the following regression:

\[ mi = \alpha_1 + \alpha_2 \cdot size + \alpha_3 \cdot volatility \]

where,

\[ a_1 = e^{\alpha_1}, \ a_2 = \alpha_2, \ a_3 = \alpha_3 \]

V. Forecasting Costs

This section describes the process to forecast costs for a specified strategy such as POV rate or VWAP. These models can also be used to forecast market impact cost for any trade schedule such as a front-load (IS) or back-load strategy. Those forecasting techniques will be discussed at a later time.

It is important to mention that when we forecast trading costs it is important that both the expected cost and timing risk are jointly evaluated.

A simplified forecasting equation where we assume the imbalance \( Q \) will be equal to the size of the order \( S \) (e.g., \( E[Q] = S \)) and we have a specified trading rate \( POV \) is:

\[ I^* = \hat{a}_1 \cdot \left( \frac{S}{ADV} \right)^{\hat{a}_2} \cdot \sigma^{\hat{a}_3} \]

\[ MI = \hat{b}_1 \cdot I^* \cdot POV + (1 - \hat{b}_1) \cdot I^* \]

Timing Risk

\[ TR = \sigma \cdot \sqrt{\frac{1}{250} \cdot \frac{1}{3} \cdot \frac{S}{ADV} \cdot \frac{1 - POV}{POV}} \cdot 10^4 \cdot bp \]

This model is also robust enough to accommodate cost forecasting where the expected imbalance on the day will be different than the number of shares traded. For example, on an index reconstitution day, the imbalance is expected to be larger in size than the shares in a single order. This model can be formulated to account for various imbalances in presence of the order \( S \). These techniques will be discussed at a later time.

Source:
It is also important to mention that we need to forecast the POV rate based on the total number of shares that we expect to trade in the market. This includes estimates from the ADV and the number of incremental shares from the order. For example, suppose we expect 1 million shares to trade in the market over the course of the day without our order (e.g., the stocks ADV = 1 million shares) and our order is for 100,000 shares (10% of ADV). If we trade the order over the course of the day following a VWAP strategy, our POV rate will be 9.09% not 10%. This is computed as follows:

\[
E[POV] = \frac{100000}{100000 + 1000000} = 0.0909
\]

For higher POV rates, such as 20% or 25% this difference will be much larger and could potentially have a very large effect on the cost estimates.
VI. Back-testing the model

Since the parameters of this model were estimated using compiled tic data (and evaluated statistically from this tic data) we are comfortable up to this point that the model is appropriate to estimate price movement from derived imbalances. But we have not yet determined if it is appropriate for estimating costs for actual customer orders.

Since we do have available the details behind the order data we can easily calculate the trading cost (using arrival price), pov rate, ADV and volatility. We start by forecasting market impact and timing risk with the newly estimated parameters. These estimates are then compared to the actual costs in four (4) different ways.

**Technique #1 – Graphical Illustration**

Plot the estimated costs and actual costs for various order sizes (cost as y-axis and size as x-axis). It is best to compute the average costs for different order sizes in order to remove market noise. It is just as important to make sure there are enough observations in each size category to eliminate the effect of market noise. This is the most helpful performance analysis for clients – although the least helpful of the 3 from a statistical perspective.

**Technique #2 – Regression**

Run a regression between the actual and estimated costs using all data. If the forecasting model is accurate the results from the regressions should show an intercept statistically equal to zero and a slope statistically equal to 1. The R2 will probably be very low but the t-stat and f-value should be very high. Additionally, the correlation should be \( \rho > 0 \) but may still be low because of noise. This analysis is very helpful to highlight if the model is working well overall (e.g., all order sizes). This analysis is the second most helpful for clients to evaluate our model and the second most helpful statistical technique.

**Technique #3 – Z-Score**

This technique allows us to jointly evaluate both the accuracy of the market impact and timing risk models. The test consists of computing a statistical z-score to determine the number of standard deviations the actual cost was from the estimated cost. The z-score is calculated as follows:

\[
Z = \frac{Actual - Estimated MI}{TR}
\]

If the model is accurate we should find the average Z value to be close to zero and the standard deviation (or variance) to be close to 1. That is, and accurate model will have:

Source:
\( Z \sim (0,1) \)

It is important in this technique to evaluate the Z statistic for various order sizes and categories such as buy/sell, market cap, etc. (and the same with the other tests) to ensure the model is working well overall. The data results could also be evaluated based on the distribution of the Z statistic and the chi-square goodness of fit test.

I have found this technique to be the most important and useful in evaluating the models from a statistical basis as a very important feature in performing real-time TCA analysis (e.g., in the algorithms or from a reporting perspective). But it has generated the least amount of interest from clients even though it is the strongest proof of accuracy.

**Technique #4 – Error Analysis**

In this performance evaluation technique we analyze the error term (regression residual) in order to determine if there are any factors driving trading cost that we have not captured. We compute the error term \( \delta_i \) (difference between estimated and actual) as follows:

\[
\delta_i = \text{EstimatedMI}_i - \text{ActualCost}_i
\]

Then we regress \( \delta_i \) on factors such as market movement, sector movement, order size (to determine robustness of fit), market cap, side, etc. A statistically significant result would show that the “factor” is a consistent contributor to trading cost.

- In the past, I have found the S&P500 index (as a proxy for market movement) to be a statistically significant factor for explaining the regression error. But I have not included SP500 index in the forecasting equation, however, because we have never constructed an accurate market forecasting model. Expected market movement on the day is taken to be zero – but there may be some insight we can derive from the futures market at the open.

- I have used market movement and the stock’s adjusted beta as a post trade report – e.g., to decompose trading cost between performance, skill, and noise.

**VII. Advanced Modeling Techniques**

As we move forward we can also discuss how these equations can be used to gain insight into some of the more advanced modeling scenarios and questions. The hot topics seem to be:

- Model Expansion for Crossing - Internal & Dark (shows the benefit of Crossing)

Source:
• Trade Scheduling – Cost Allocation Approach (Front & Back-loading)
• Incremental Market Imbalance – Estimating Costs around the Russell and other index Reconstitution
• High Frequency Trading – Estimating HF Volumes

Source: