Performance metrics for algorithmic traders

Rosenthal, Dale W.R.

University of Illinois at Chicago

22 June 2009

PERFORMANCE METRICS FOR ALGORITHMIC TRADERS

DALE W.R. ROSENTHAL

Abstract. Portfolio traders may split large orders into smaller orders scheduled over time to reduce price impact. Since handling many orders is cumbersome, these smaller orders are often traded in an automated ("algorithmic") manner. We propose metrics using these orders to help measure various trading-related skills with low noise. Managers may use these metrics to assess how separate parts of the trading process contribute execution, market timing, and order scheduling skills versus luck. These metrics could save 4 basis points in cost per trade yielding a 15% reduction in expenses and saving $7.3 billion annually for US-domiciled equity mutual funds alone. The metrics also allow recovery of parameters for a price impact model with lasting and ephemeral effects. Some metrics may help evaluate external intermediaries, test for possible front-running, and indicate sloppy or overly passive trading.

JEL: G12, G14, G23, G24

1. Introduction

Traders with a portfolio of orders often split and execute those orders across time to hide alpha or reduce execution costs in light of market liquidity. Berke (2010) estimates that over 30% of volume is the result of order splitting. McPartland (2010) estimated that $13.4 billion would be spent on trading infrastructure globally in 2010. Since the use of order splitting,

---

Date: June 22, 2009 (first version); January 4, 2012(this version).
Department of Finance, University of Illinois at Chicago, 601 S. Morgan St. (MC 168), Chicago, IL 60607, +1 312 996 7170, daler@uic.edu
Valuable feedback was provided by Bruce Lehmann, Albert Menkveld, Thierry Foucault, and other participants at the CREATES market microstructure symposium, the CQA Spring 2010 meeting, FERM 2010, and seminar participants at Spot Trading LLC..

smart order routing, and other algorithmic trading techniques is becoming widespread, we can expect that a significant portion of spending on trading infrastructure is for such tools. Indeed, Honoré (2009) estimated US financial firms would spend \$1 billion on smart order routing alone (to access superior prices) in 2010. For such a large amount of money spent, managers must know the value of high-speed execution systems, market timing models, and order schedulers.

In particular, we seek to distinguish if good trading performance is due to execution skill, short-term market timing ability, skillful scheduling of orders, or just luck. This article proposes metrics to help portfolio managers measure these skills and reduce the noise of these measurements.

Significant savings may be achieved by application of the metrics we propose. In 2010, the assets under management in equity and hybrid mutual funds was about \$5.6 trillion in the US with the US comprising 48% of the world market, according to the Investment Company Institute. Those funds had about 64% turnover annually and expense ratios of 0.87% for equity funds and 0.84% for hybrid funds. Simulation results with the proposed metrics show possible savings of 2 basis points per trade for both trading and order scheduling skills (a total of 8 bp from entry to exit of a position). Given these figures, a typical fund could save 13 basis points per year. Across US-domiciled equity mutual funds, expenses could thus be reduced by almost 15% leading to a savings of \$7.3 billion annually.[1] One could expect even greater savings for pension funds and insurance companies.

High-speed smart order routing and execution management systems (EMS) were ushered in by Reg NMS (in the US) and the MiFID directive (in the EU). Reg NMS's demand that each order be provided "best execution" could

[1]Bond funds and internationally-domiciled equity funds would also benefit from the metrics proposed here; however, the benefit to those funds is more difficult to assess.

be interpreted as the right to challenge any trade which did not achieve the best price aross all venues at that time. While Reg NMS allows for other interpretations of best execution, a best-price interpretation is the most likely for small orders where liquidity is not a concern. If we expand the time window of comparison, best execution would likely imply achieving a price close to that achieved by others trading at about the same time. This comparison is one of the proposed metrics. Heston *et al.* (2010) suggested such a measure relates to spreads and impact incurred only by an individual trader; that suggestion is shown to be correct.

Under Reg NMS, larger orders might be challenged as being too large to be executed at one time. This issue is typically handled by an order management system (OMS) or trading engine which schedules (*i.e.* splits) orders across time. There are many reasons to split the execution of large orders. Splitting an order to hide alpha was studied by Kyle (1985). Bertsimas and Lo (1998) examined order splitting to optimize execution costs. Almgren and Chriss (2001) extended this to reduce the mean-variance cost of trading and to create the idea of an efficient execution frontier. Engle and Ferstenberg (2007) showed that combining the portfolio and order scheduling optimization yields a better optimal portfolio. The failure to do so may explain why portfolios frequently underperform the Sharpe ratios suggested by standard portfolio optimizations. Therefore, the order scheduling optimization must be an integral part of portfolio optimization if investment managers are to get useful forecasts of Sharpe ratios. Metrics proposed here may help investment managers by offering methods to estimate parameters needed to combine the order splitting and investment optimizations.

When markets are stagnant and returns are low, small transactions costs may consume a large fraction of returns or, even worse, result in negative returns. For large endowments, pension funds, retirement plans, and mutual funds, these problems are particularly acute; the large portfolio trades they execute may greatly strain market liquidity. Given the amount of money involved and the liquidity strain, these portfolio trades must be handled with the utmost skill and care: poor executions, lousy market timing, or careless order scheduling may be very costly. Indeed, a cross-sectional comparison of mutual fund performance by Fama and French (2010) shows that costs tend to flow through to lower returns; however, they along with Puckett and Yan (2011) also find that some funds have persistent outperformance. Puckett and Yan relate this outperformance to trading skill.

We propose metrics which help portfolio managers measure various trading skills by isolating the effect of luck on a set of trades. Some metrics are resistant to gaming and may be used to evaluate external traders. Two of the metrics may even detect some basic forms of front-running or "trading alongside." Finally, many of the metrics here may be taken in isolation; this allows academics and policy makers to study various aspects of market quality across venues or time.

2. Developing the Metrics: Considering Counterfactuals

A rich literature exists on investment performance metrics. Yet despite the increasing focus on trading, less work has been done on trading performance metrics. We keep in mind the helpful guidelines set by Lehmann (2003) and develop the metrics by asking counterfactual questions. In this way, our reasoning is similar to that of another metric, the implementation shortfall of Perold (1988). The implementation shortfall compares the average price achieved plus an estimate of opportunity costs to the initial portfolio price. The metrics are therefore complementary to the implementation shortfall. The set of questions also results in many of the metrics decomposing overall trading performance.

2.1. **Terminology.** To discuss order splitting, we need clear terminology. I follow Engle and Ferstenberg (2007) and call the order which is split the *parent order*; the orders generated by splitting I call *child orders*. A set of parent orders executed together constitute a *portfolio order*.

I refer to metrics being suitable for *internal* or *external* use in the sense of Lehmann (2003). Internal use presumes a performance auditor who does not seek to distort or game metrics and who knows all investment and trading decisions including actions considered but not taken. External use requires that an auditor know only the investment decision and executions; however, metrics for external use must be resistant to gaming.

Since the number of child orders generated by order scheduling may be large, the trading process is often automated. *Algorithmic trading* is the automated, often research-driven, creation and management of orders. Our concern is mostly large algorithmically-traded portfolio orders with parent orders split into many child orders. We examine child orders and executions across instruments and days by splitting time into *bins*. While bins are contiguous, their lengths of time may vary. Thus we can choose shorter bins when volume or volatility is typically higher.

2.2. **Parent Order Metrics.** The first set of metrics are the *parent order metrics* which ask the following questions:

- What if we began trading when somebody external knew of our order?
- What was the marginal (or incremental) cost of our last trade?
- What would the profit be for providing liquidity to that last trade?
- What is the lasting effect our trading had on prices?
- How much worse did we do than that lasting effect on prices?

The answers to these questions give us the parent order metrics which decompose trading effects at an instrument level. We can answer the last three questions by examining a period after trading: a parent order traded over an entire day might use the following day for comparison.

Information leakage measures the instrument's value drift between when orders were revealed (perhaps via bidding for the portfolio of orders) and when those orders began to be traded. An order revealed, even partially, for pricing may be front run. This metric allows testing if order information is leaking to the market.

Incremental impact measures the difference between the price at trading end and the average execution price. That difference is the incremental effect of additional trading on the average price. If there were no price impact due to trading, this quantity would have expectation 0.

Decaying impact is the difference between the instrument price at the end of trading and some fair price in the next period; since some impact decays after trading, this measures the decay of non-temporary price impact (realized implementation shortfall and incremental impact) over the post-trading period. This is a proxy for what a market maker would earn by providing liquidity for the last scheduled orders.[2]

Permanent impact is the difference between the price at the start of trading and the next period's fair price; this measures the lasting price change. To the extent a parent order contains economic information, we should believe that $E(PI) > 0$. Noise traders and liquidity providers may have expectationally negative permanent impact.

Adverse selection is the difference between the implementation shortfall and the next period's fair price; this is the cost which dissipates over the

[2]A market maker providing liquidity effectively trades against their client at the end of trading. For example, a client making a large purchase might increase the price of a security. A market maker could then provide liquidity or "facilitate" the trade by selling short to the client for the last order. This might prevent the client from pushing prices higher and thus is sometimes referred to as "price improvement." Nonetheless, the market maker is then short a security which is likely to decline in price. Some clients might view this as an acceptable risk-shifting agreement; others are not so sanguine.

next period. Thus this includes both price impact and poor child order execution. Microstructure models suggest such costs are due in part to market maker concerns about adverse selection. Without adverse selection, we would expect a trading period fair price to be close to the next-period fair price.

2.3. **Intertemporal Metrics.** The second set of metrics use individual executions to build *intertemporal metrics* which answer the following questions:

- What if each child order had been filled at a fair price?
- What if each child order had been filled when it was scheduled?
- What if each child order were scheduled to match the typical distribution of volume over time?
- What if each child order were scheduled to match the actual distribution of volume over time?

These intertemporal metrics measure various skills and noise across time. They do so by decomposing parent order performance by looking at child orders and executions in the context of their respective time bins. The first three metrics (trading, fill time, and ordertiming shortfalls) are the most informative and correspond to separate decisions about how to trade, how patient to be with limit orders or unsent trades, and how to schedule child orders. These metrics also let us see if particular times of day are troublesome.

Trading shortfall compares our average execution price in each bin to a fair price for that bin. Thus the main determinant of trading shortfall should be execution skill. This should reduce much of the noise in trading performance and allow us to see execution quality more clearly. If we could consistently achieve superior executions, we would expect a negative trading shortfall. Algorithmic traders optimizing execution strategies should see the effects of those changes manifested in trading shortfall. The performance of an EMS configured to execute trades in a short window of time would be encapsulated by this metric.

Fill time shortfall measures the cost of foregoing immediate execution of child orders, *i.e.* the cost of diverging from our execution plan. An EMS or OMS configured to let orders sit passively in hopes of better prices is essentially timing the market over a short horizon; thus the fill time shortfall attempts to assess short-term market timing skill. This metric encapsulates the "strategy deviations" discussed by Kissell and Malamut (2005), *i.e.* the benefit of sacrificing execution immediacy for price sensitivity.

Order timing shortfall measures the cost of not trading according to the average distribution of volume over time. A negative shortfall suggests a strategy deviates from trading in line with expected volume when prices are advantageous. The performance of an OMS, trading engine, or algorithmic trader creating order schedules would be encapsulated by the order timing shortfall. This metric can also be seen as capturing the performance of a chosen point on the efficient frontier of trading (*i.e.* a chosen order schedule).

Volume shortfall measures the cost of volume distribution variation or noise. This is done by comparing the cost of executing orders following the average versus the actual volume distribution. Since the average volume distribution is an (unconditional) expectation, we expect this metric to have mean 0 (as is true for most noise terms).

Finally, the *perfect VWAP shortfall* measures the difference between the price at the start of trading and a basic volume-weighted average price (VWAP) strategy of splitting orders according to the average volume distribution. This does not answer one of the above questions — since it is needed merely for the definition of the metrics.

3. Mathematically Defining the Metrics

All decompositions assume we have no alpha or have pre-subtracted it.[3] We also assume that, over short to intermediate time frames, the VWAP is fair (*i.e.* unbeatable on average without alpha). From experience and discussions with practitioners, this seems reasonable: No practitioners consulted thought VWAP could be beat without alpha. Opiela (2006) quotes a practitioner who found various trading algorithms missed VWAP by 1/4 of a cent (about 1/2 bp) and that this is consistent with performance claimed by other trading algorithm vendors. We later prove that, absent alpha, VWAP is a fair price and use this to derive some basic asymptotic properties.

To rigorously define these ideas, we say that trading starts at time $t = 0$, ends at time $t = T$, and:

$$\begin{aligned} q, \tilde{q} &= \text{signed ordered, executed amount (\textit{e.g.} shares);} \\ p_-, p_t &= \text{price at first desire to trade, price at time } t; \\ \bar{p}_T, \check{p}_+ &= \text{average fill price, next-period fair price.} \end{aligned}$$

3.1. Defining the Parent Order Metrics.

The parent order metrics offer more insight into how well a parent order was traded beyond what the realized implementation shortfall measures. These metrics are defined for a parent order but examining them at a portfolio order level reduces noise.

To recap: Information leakage (IL) is the price change from order revelation to trading start. Incremental impact (II) compares the average price to the price at trading end. Decaying impact (DI) compares the price at trading end to the next-period fair price. Permanent impact (PI) compares the starting price to the next-day fair price. Adverse selection (AS) compares the average price to the next-period fair price. Table 1 gives the definitions

[3] If we expect prices to increase by 4 basis points over the period, a purchase with an implementation shortfall of 10 basis points should only explain 6 basis points of shortfall. However, a sale should explain 14 basis points of shortfall.

of these metrics while Figure 1 illustrates them. Note that they yield two equalities for the realized implementation shortfall $RIS = \tilde{q}(\bar{p}_T - p_0)$ of Perold (1988):

$$RIS = PI + AS = PI + DI - II. \tag{1}$$

The metrics are summarized in Table 1.

Metric	Definition	Concept
Information Leakage (IL)	$q(p_0 - p_-)$	Pre-trade price drift
Incremental Impact (II)	$\tilde{q}(p_T - \bar{p}_T)$	Marginal cost of last trade
Decaying Impact (DI)	$\tilde{q}(p_T - \check{p}_+)$	Profit of providing liquidity
Permanent Impact (PI)	$\tilde{q}(\check{p}_+ - p_0)$	Information dissemination cost
Adverse Selection (AS)	$\tilde{q}(\bar{p}_T - \check{p}_+)$	Average price vs post-trade.

TABLE 1. Definitions and concepts for parent order metrics. The realized implementation shortfall, $RIS = \tilde{q}(\bar{p}_T - p_0)$, may be composed of these metrics via $RIS = PI + AS = PI + DI - II$.

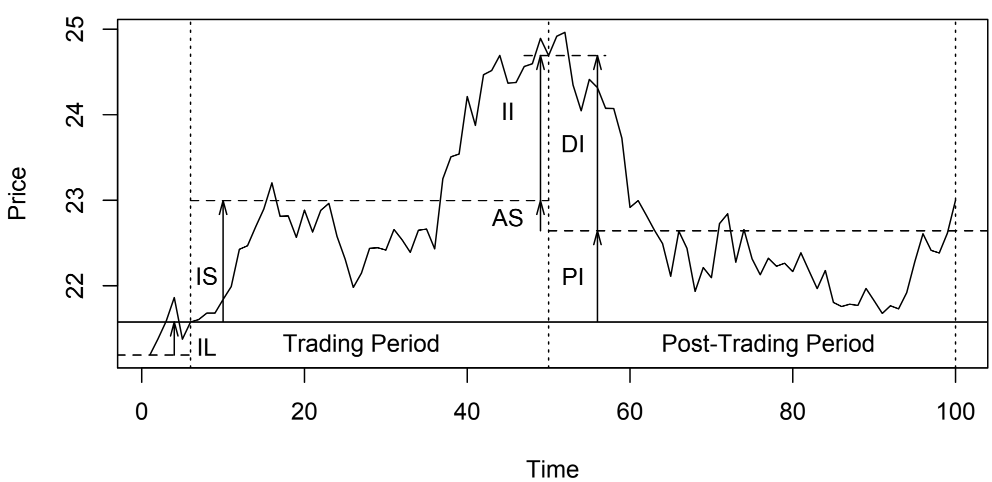

FIGURE 1. Price versus time with parent order metrics. Dashed lines show our average fill price and the next-period fair price.

3.2. **Defining the Intertemporal Metrics.** For more detail, we analyze child order executions. This lets us answer how much of a parent order's metrics (from the preceding subsection) are due to luck versus different skills. One *caveat* bears mentioning, however: While these metrics are defined on a parent order level, care should be taken if they are examined below the portfolio order level. Focusing on one securitiy in isolation may miss the interaction of how parent orders across the portfolio of orders are scheduled or traded. For example, a long-short portfolio might schedule its technology buys and sells to stay balanced throughout the trading period.

To do these analyses, the trading period is partitioned into "bins" to define the decompositions.[4] As in the preceding subsection, we use a "fair" price to assess execution quality — although we now do so for each time bin. We then compare how we would have done had we (i) achieved fair bin prices; (ii) executed all child orders when scheduled; (iii) traded in line with the expected volume distribution; and (iv) predicted volumes correctly.

This more detailed decomposition requires more notation:

$$\begin{aligned}
q_j &= \text{child order quantity in bin } j \quad (q = \textstyle\sum_j q_j);\\
\tilde{q}_j &= \text{child order quantity filled in bin } j \quad (\tilde{q} = \textstyle\sum_j \tilde{q}_j);\\
V_j, \bar{V}_j &= \text{realized, average volume in bin } j \quad (V = \textstyle\sum_j V_j);\\
D_j, \bar{D}_j &= \text{realized, average fraction of period volume in bin } j;\\
\check{p}_j &= \text{fair price in bin } j\text{; and,}\\
\bar{p}_j &= \text{realized average price in bin } j.
\end{aligned}$$

We again consider the realized part of the implementation shortfall of a security, $RIS = \sum_j \tilde{q}_j \bar{p}_j - \tilde{q} p_0$. We take no position on whether this is an appropriate execution benchmark nor does this choice affect decomposition definitions.[5] For a benchmark price of the starting price, p_0, this is the

[4] The decompositions could also be defined using kernel estimators.

[5] Other possible benchmarks include VWAP and auction (open/close) prices.

same as $\sum_s \kappa_{bs}|q_s|$ in Lehmann (2003). We then decompose the realized implementation shortfall:

$$
\begin{aligned}
RIS =& \underbrace{\sum_j \tilde{q}_j(\bar{p}_j - \check{p}_j)}_{\text{Trading Shortfall}} + \underbrace{\sum_j (\tilde{q}_j - \tilde{q}\frac{q_j}{q})\check{p}_j}_{\text{Fill Time Shortfall}} + \underbrace{\sum_j \tilde{q}(\frac{q_j}{q} - \bar{D}_j)\check{p}_j}_{\text{Order Timing Shortfall}} + \\
&+ \underbrace{\sum_j \tilde{q}(\bar{D}_j - D_j)\check{p}_j}_{\text{Volume Shortfall}} + \underbrace{\sum_j \tilde{q}(D_j\check{p}_j - p_0)}_{\text{Perfect VWAP Shortfall}}
\end{aligned}
\tag{2}
$$

Recapping these definitions: Trading Shortfall (TS) measures how we do in each bin versus fair prices for those bins. Fill time shortfall (FTS) measures the cost of the realized versus desired execution timing.[6] Order timing shortfall (OTS) measures the cost of executing following the desired order schedule versus the average volume distribution. Volume shortfall (VS) measures the cost due to volume distribution noise by comparing the cost of executing orders following the average versus the actual volume distribution. Perfect VWAP shortfall (PVS) measures the difference between using a VWAP benchmark and the initial price; this is needed to define the other metrics. Apart from the trading shortfall, the intertemporal metrics use the fair price in each time bin (since that price is independent of trading skill).

Since these metrics are part of a decomposition, they are likely to be inter-dependent (except for the volume shortfall which, by construction, is orthogonal to the other metrics). However, the trading shortfall, fill time shortfall, and order timing shortfall all correspond to products or decisions made on a trading desk. EMS products such as Morgan Stanley's SORT, and Goldman Sachs's Sigma X would affect the trading shortfall. They could also affect the fill time shortfall if they were allowed to let orders sit

[6]We use q_j/q for the desired execution timing distribution. This assumes orders are sent when we desire executions.

for more than a short period of time. OMS and trading engine products such as Morgan Stanley's Benchmark Execution Strategies (BXS) and Goldman Sachs's Algorithmic Trading (GSAT) would affect the order timing shortfall.

4. Analysis: Moments, Error, and Parameter Estimation

To explore the meaning and use of these metrics, we need a fair price for each bin and a price impact model. Until now, we have used VWAP as a fair price without justification. We now justify that usage so that we may check our intuition and see how the metrics perform.

4.1. **Fair Price for Comparison.**

Since a fair price should not be beatable on average, a fair guess would be to use VWAP. To show this qualifies as a fair price, we begin with a painless proof. We then show that attaining VWAP may be difficult.

Proposition 1 (VWAP is Fair)**.** *Assume the price impact of trading is arbitrage-free as in Huberman and Stanzl (2004); we have no alpha; and, VWAP is measured with pre-specified begin and end times which coincide with the period of trading. Then, a trader cannot expect to beat VWAP.*

Proof. By Huberman and Stanzl, one trader cannot beat their average price nor expect to make quasi-arbitrage profits.

Two traders active over the trading period cannot beat their average prices nor make arbitrage profits knowing only their own trades (implied by "no alpha"). Thus they cannot expect to beat the other's average price. Since VWAP is a trade-size-weighted average of unbeatable prices, neither trader can expect to beat VWAP.

Finally, assume k traders cannot beat VWAP without alpha. For $k+1$ traders, the $k+1$-st trader cannot make arbitrage profits nor beat the average price of the first k traders without alpha. None of the first k traders can make arbitrage profits nor beat the average price of the $k+1$-st trader without alpha. Since none of the traders can expect to beat another's average price, none can expect to beat VWAP. □

Corollary 1 (VWAP May Be Unattainable)**.** *Under the setup for Proposition 1: Suppose one or more traders have alpha while the others do not. In that case, the traders without alpha should expect to do worse than VWAP.*

The proof is obvious given the proposition. From here forward, we assume a fair price $\check{p}_j$ is the bin j VWAP and $\check{p}_+$ is the post-trading period VWAP.

4.2. **Dynamics.** We use a model with three types of price impact: permanent, decaying (to 0), and temporary (affecting only the generating trade). These are parameterized by Greek letters π, δ, and (τ, ϕ).

As in Almgren and Chriss (2001) and Obizhaeva and Wang (2006), we assume continuous trading at a rate $\tilde{q}_j/t_j$ within a bin j. Trading induces permanent impact (π) linear in quantity and decaying impact (δ) decreasing geometrically from an initial impact linear in quantity. Temporary impact (τ) is linear in the trading rate plus a fixed fee (ϕ) per trade.[7] We can then write the volume and price evolution equations:

$$V_j = \nu_j + |\tilde{q}_j| \qquad \nu_j \overset{indep}{\sim} (\mu_{\nu_j}, \sigma^2_{\nu_j}) \qquad \nu_j > 0 \tag{3}$$

$$p_{j+1,0} = p_0 + \sum_{k=1}^{j} \Big[\sigma_{p,k} Z_k + \underbrace{\pi \tilde{q}_k}_{\text{permanent}} + \underbrace{\delta^{j+1-k} \tilde{q}_k}_{\text{decaying}}\Big] \qquad Z_k \overset{iid}{\sim} (0,1) \tag{4}$$

$$\bar{p}_j = p_0 + \sum_{k=1}^{j} \Big[\sigma_{p,k} Z_k + \underbrace{\pi \tilde{q}_k}_{\text{permanent}} + \underbrace{\delta^{j+1-k} \tilde{q}_k}_{\text{decaying}}\Big] + \underbrace{\tau \frac{\tilde{q}_j}{t_j} + \phi \operatorname{sgn}(\tilde{q}_j)}_{\text{temporary}} \tag{5}$$

where $\nu_j \perp\!\!\!\perp Z_k$ for all j, k. No distributional assumptions are made on the ν's or Z's beyond their support, mean, and variance.

[7] The decaying term is nearly equivalent to that in the model of Obizhaeva and Wang (2006) and allows for price impact which dissipates after trading a parent order.

These allow us to compute the bin VWAP and post-trade period VWAP:

$$\check{p}_j = p_0 + \sum_{k=1}^{j} \left[\sigma_{p,k} Z_k + \pi \tilde{q}_k + \delta^{j+1-k} \tilde{q}_k \right] + \tilde{q}_j \frac{\tau \frac{|\tilde{q}_j|}{t_j} + \phi}{V_j}, \tag{6}$$

$$\check{p}_+ = p_0 + \pi \tilde{q} + \sum_{j=1}^{n} D_{j,+} (\sigma_{p,j} Z_j + \underbrace{\delta^j \sum_{k=1}^{n} \delta^{n+1-k} \tilde{q}_k}_{\text{decaying impact}}). \tag{7}$$

4.3. **Analyzing Parent Order Metrics.** We analyze the parent order metrics along with the realized implementation shortfall for comparison. To do this, we assume a probability triple at any time t $(\Omega, \mathcal{F}_t, \mathcal{P})$ with $\mathcal{F}_t$'s encapsulating all information known at time t. This yields the results in Table 2. We next make some assumptions about bin lengths and the volume distribution to simplify the full equations and gain insight into the price impact model parameters.

Metric	$E(\cdot)$	$\text{Var}(\cdot)$
$IL\vert\mathcal{F}_-$	$q(p_0 - p_-) = 0$	$q^2\sigma_p^2(t_0 - t_-)$
$RIS\vert\mathcal{F}_0$	$\sum_{j=1}^{n}\sum_{k=1}^{j}\left[\frac{\pi\tilde{q}_j\tilde{q}_k}{\tilde{q}} + \frac{\delta^{j+1-k}\tilde{q}_j\tilde{q}_k}{\tilde{q}}\right]$ $+\frac{\tau}{\tilde{q}}\sum_{j=1}^{n}\frac{\tilde{q}_j^2}{t_j} + \phi\,\text{sgn}(\tilde{q}_j)$	$\sum_{j=1}^{n}\left(\sum_{k=j}^{n}\sigma_{p,k}\tilde{q}_k\right)^2$
$II\vert\mathcal{F}_0$	$\pi\sum_{j=1}^{n}\sum_{k=j+1}^{n}\frac{\tilde{q}_j\tilde{q}_k}{\tilde{q}}$ $+\delta\sum_{j=1}^{n}\left[\delta^{n-j}\tilde{q}_j - \sum_{k=1}^{j}\frac{\delta^{j-k}\tilde{q}_j\tilde{q}_k}{\tilde{q}}\right]$ $-\frac{\tau}{\tilde{q}}\sum_{j=1}^{n}\tilde{q}_j^2/t_j - \phi\,\text{sgn}(\tilde{q}_j)$	$\sum_{j=1}^{n-1}\left(\sum_{k=1}^{j}\sigma_{p,k}\tilde{q}_k\right)^2$
$AS\vert\mathcal{F}_0$	$\sum_{j=1}^{n}\sum_{k=1}^{j}\left[\frac{\pi\tilde{q}_j\tilde{q}_k}{\tilde{q}} + \frac{\delta^{j+1-k}\tilde{q}_j\tilde{q}_k}{\tilde{q}}\right]$ $+\frac{\tau}{\tilde{q}}\sum_{j=1}^{n}\tilde{q}_j^2/t_j + \phi\,\text{sgn}(\tilde{q}_j) - \pi\tilde{q}$ $-\sum_{j=1}^{n}\delta^j\bar{D}_j\sum_{k=1}^{n}\delta^{n+1-k}\tilde{q}_k$	$\sum_{j=2}^{n}\left(\sum_{k=1}^{j-1}\sigma_{p,k}\tilde{q}_k\right)^2$ $+\tilde{q}^2\sum_{j=1}^{n}\left(\sum_{k=j}^{n}\sigma_{p,k}\bar{D}_k\right)^2$
$DI\vert\mathcal{F}_T$	$\sum_{k=1}^{n}\delta^{n+1-k}\tilde{q}_k\left(1 - \sum_{j=1}^{n}\delta^j\bar{D}_j\right)$	$\tilde{q}^2\sum_{j=1}^{n}\left(\sum_{k=j}^{n}\sigma_{p,k}\bar{D}_k\right)^2$
$PI\vert\mathcal{F}_0$	$\pi\tilde{q} + \sum_{k=1}^{n}\delta^{n+1-k}\tilde{q}_k\sum_{j=1}^{n}\delta^j\bar{D}_j$	$\tilde{q}^2\sum_{j=1}^{n}\left(\sum_{k=j}^{n}\sigma_{p,k}\bar{D}_k\right)^2$ $+\tilde{q}^2\sigma_p^2$

TABLE 2. Parent order metric expectations and variances. The summations are over n child orders.

We define time bins (via t_j's) so all $\tilde{q}_j = \tilde{q}/n$. (We assume this makes $\sigma_{p,j} = \sigma_p/\sqrt{n}$ to first order.) This simplifies the formulæ for realized implementation shortfall and incremental impact (RIS and II). If the expected volume distribution is non-degenerate, we can also simplify the AS, DI, and PI formulæ.[8] The results of these simplifications are shown in Table 3.

Metric	$E(\cdot)$	$\mathrm{Var}(\cdot)$
$RIS\|\mathcal{F}_0$	$\pi\tilde{q}\frac{n+1}{2n} + \frac{\tilde{q}\delta}{n(1-\delta)} + \tau\frac{\tilde{q}}{n^2}\sum_{j=1}^{n}\frac{1}{t_j} + \phi\,\mathrm{sgn}(\tilde{q}_j)$	$\sigma_p^2\tilde{q}^2\frac{5n^2+6n+7}{6n^2}$
$II\|\mathcal{F}_0$	$\pi\tilde{q}\frac{n-1}{2n} - \tau\frac{\tilde{q}}{n^2}\sum_{j=1}^{n}\frac{1}{t_j} - \phi\,\mathrm{sgn}(\tilde{q}_j)$	$\sigma_p^2\tilde{q}^2\frac{(n-1)(2n-1)}{6n^2}$
$DI\|\mathcal{F}_T$	$\frac{\tilde{q}\delta}{n(1-\delta)}$	$\sigma_p^2\tilde{q}^2\frac{5n^2+6n+7}{6n^2}$
$PI\|\mathcal{F}_0$	$\pi\tilde{q}$	$\sigma_p^2\tilde{q}^2\frac{11n^2+6n+7}{6n^2}$
$AS\|\mathcal{F}_0$	$\frac{\tilde{q}\delta}{n(1-\delta)} - \pi\tilde{q}\frac{n+1}{2n} + \frac{\tau\tilde{q}}{n^2}\sum_{j=1}^{n}\frac{1}{t_j} + \phi\,\mathrm{sgn}(\tilde{q}_j)$	$\sigma_p^2\tilde{q}^2\frac{2n^2+4}{3n^2}$

TABLE 3. Simplified parent order metric expectations and variances. Expectation approximations are of $O(1/n^2)$ where n is the number of child orders.

4.4. **Analyzing Intertemporal Metrics.** Intertemporal metrics depend heavily on the strategies for creating and executing orders. Some intertemporal metrics are, expectationally, covariances and relate expected changes in market conditions, order placement strategies, and order decisions as mentioned in Lehmann (2003). With that in mind, we examine the intertemporal metrics in light of the impact model.

Trading Shortfall (TS) is the quantity-weighted sum of the difference between our bin average prices and the bin VWAPs:

$$E(TS|\mathcal{F}_0) = \sum_{j=1}^{n} \tilde{q}_j \left(\tau\frac{|\tilde{q}_j|}{t_j} + \phi\right)\left(1 - \frac{|\tilde{q}_j|}{|\tilde{q}_j| + \mu_{\nu_j}}\right), \tag{8}$$

$$\mathrm{Var}(TS|\mathcal{F}_0) = \sum_{j=1}^{n} \tilde{q}_j^4 \left(\tau\frac{|\tilde{q}_j|}{t_j} + \phi\right)^2 \frac{\sigma_{\nu_j}^2}{(|\tilde{q}_j| + \mu_{\nu_j})^4}. \tag{9}$$

[8]Formally, we assume that $\bar{D}_j = O(1/n)$ where n is the number of child orders.

Trading shortfall is thus directly affected by temporary impact alone. This corroborates findings by Heston *et al.* (2010).

Fill Time Shortfall (FTS) measures the cost of our orders being filled at times different from when we send them:

$$E(FTS|\mathcal{F}_0) = \tilde{q}\,\mathrm{Cov}(\tilde{q}_\cdot/\tilde{q} - q_\cdot/q, \check{p}_\cdot), \tag{10}$$

$$\begin{aligned}\mathrm{Var}(FTS|\mathcal{F}_0) &= \sum_{k=1}^{n}\left(\sum_{j=k}^{n}\sigma_{p,k}(\tilde{q}_j - \tilde{q}\frac{q_j}{q})\right)^2 \\ &\quad + \sum_{j=1}^{n}(\tilde{q}_j - \tilde{q}\frac{q_j}{q})^2(\tau\frac{|\tilde{q}_j|}{t_j} + \phi)^2\tilde{q}_j^2\frac{\sigma_{\nu_j}^2}{(\mu_{\nu_j} + |\tilde{q}_j|)^4}.\end{aligned} \tag{11}$$

The expectation of fill time shortfall yields the relation between divergence from an execution schedule and better prices. A positive covariance would imply a tendency to have orders filled when prices are disadvantageous.

Order Timing Shortfall (OTS) measures the cost of sending orders according to our desired schedule versus the average volume distribution:

$$E(OTS|\mathcal{F}_0) = \tilde{q}\,\mathrm{Cov}(q_\cdot/q - \bar{D}_\cdot, \check{p}_\cdot), \tag{12}$$

$$\begin{aligned}\mathrm{Var}(OTS|\mathcal{F}_0) &= \tilde{q}^2\sum_{k=1}^{n}\left(\sum_{j=k}^{n}\sigma_{p,j}(\frac{q_j}{q} - \bar{D}_j)\right)^2 \\ &\quad + \tilde{q}^2\sum_{j=1}^{n}(\frac{q_j}{q} - \bar{D}_j)^2(\tau\frac{|\tilde{q}_j|}{t_j} + \phi)^2\tilde{q}_j^2\frac{\sigma_{\nu_j}^2}{(\mu_{\nu_j} + |\tilde{q}_j|)^4}.\end{aligned} \tag{13}$$

The expectation of order timing shortfall is the covariance between prices and the divergence of our planned execution schedule from the actual volume distribution. A positive covariance implies a tendency to send orders when prices are disadvantageous versus sending orders in line with trading volume.

The last two intertemporal metrics do not measure performance. The *Volume Shortfall* (VS) measures the cost due to divergence between the actual and average volume distribution. If we take expectations, we may write it as the covariance between prices and differences of the realized and average volume distribution. This term is random noise in the volume distribution (hence $\text{Var}(VS)$ is omitted). The *Perfect VWAP Shortfall* (PVS) measures the cost of a VWAP benchmark versus the benchmark we instead chose. This metric is merely needed for definition of the other decompositions.

4.5. **Recovering the Price Impact Parameters.** Using two parent order metrics and the trading shortfall, we can recover the price impact model parameters.

The permanent impact (PI) and decaying impact (DI) allow us to recover the permanent π and decaying δ parameters by linear and nonlinear regressions of PI and DI on executed quantities:

$$PI = \beta_{0,PI} + \pi\tilde{q} + \epsilon \tag{14}$$

$$DI = \beta_{0,DI} + \frac{\delta}{1-\delta}\frac{\tilde{q}}{n} + \epsilon. \tag{15}$$

The $O(1/n^2)$ bias terms in the PI and DI equations should be small for parent orders split into many child orders (*i.e.* n large). If we split orders into differing numbers of n child orders, we can add the $O(1/n^2)$ terms to the above regressions to correct for biases.[9]

We recover the temporary (τ, ϕ) parameters by linear regression of bin trading shortfalls on bin executed quantities scaled by the fractions of bin volume not due to our trading:

$$TS_j = \beta_{0,TS} + \tau\frac{\tilde{q}_j|\tilde{q}_j|}{t_j}\left(1 - \frac{|\tilde{q}_j|}{V_j}\right) + \phi\tilde{q}_j\left(1 - \frac{|\tilde{q}_j|}{V_j}\right) + \epsilon. \tag{16}$$

The β_0 intercepts are nuisance parameters to correct for imbalances in the data. No further intuition into the β_0's is warranted.

[9]Without varying n, we cannot identify the intercept versus the bias terms.

5. Interpretation: Beyond the Counterfactuals

The preceding analysis illustrates the value of these metrics. We can also answer questions beyond the counterfactuals which led to the metrics. Finally, we can also consider to what extent the various metrics may be gamed. This is critical if we are measuring external managers or traders.

The realized implementation shortfall is well-covered in Perold (1988) and others. Worth noting, however, is that RIS depends on all three types of price impact: permanent, decaying, and temporary. Since permanent price impact is inescapable (by definition), IS is not a clean optimizable measure of trading performance. Furthermore, the unrealized portion of the implementation shortfall values unfilled shares at a price from one point in time — often the end of trading. This makes the implementation shortfall especially easy to game.

5.1. Interpreting Parent Order Metrics.

Information leakage (IL) measures the cost of price drift from when we reveal order information to when trading starts. Since portfolio orders may be submitted to multiple agents to solicit execution pricing, the possibility for front-running exists.[10] If the IL were large, we might suspect front-running or "pre-hedging" of our order.

The expectation and variance under the null hypothesis of no information leakage yield a sensible t-test for unusual price drift:

$$t = \frac{\operatorname{sgn}(q)(p_0 - p_-)}{\sigma_p \sqrt{t_0 - t_-}}, \tag{17}$$

with rejection suggesting front-running. To increase the test power, we could look at this statistic on a portfolio order level. The test can be gamed by claiming that trading started earlier than it actually did. Test failure is

[10]Bid submissions for portfolio orders may list standard metrics along with a few of the instruments to be traded. Those characteristics may allow partial inference of the portfolio order. Lucchetti (2005) describes these issues in greater depth.

not evidence of front-running. However, it suggests further examination — especially if some order information is revealed prior to trading.

The incremental impact (II) metric measures how far prices were pushed beyond our average execution price. A high II might attract more traders providing liquidity (by taking the opposite side of our trading) — since a high II implies a rapid and recent change in price. A high II might also suggest the final orders were traded too aggressively and should have been spread over a longer time or that order completion should not have been mandated.

The decaying impact (DI) metric measures cumulated decaying price impact which depends on the order schedule. Consistently higher or lower DI metrics indicate poor or excellent scheduling of child orders. A high DI might also suggest a parent order was traded over too short a time period.

Since II and DI refer to prices at trading end, they can be gamed to look artificially small by ceasing trading prior to the reported stop time. This means II and DI are suitable for internal use but should not be used to measure external traders unless the manager sees all executions with timestamps and knows all were traded in the market. Even traders unaware of these metrics might distort them unknowingly. For example, broker-dealers who provide liquidity at the end of trading effectively stop trading early; they reduce II and DI in attempting to profit from price reversion.

The permanent impact (PI) metric measures total permanent price impact which is linear in the total quantity traded and is the change in equilibrium price due to economic information in the order. This suggests PI is unaffected by skill and should be fairly consistent over time.

The adverse selection (AS) metric measures some permanent as well as decaying and temporary impact. Since classic microstructure models impound fears of adverse selection into price impact, the composition of AS

should come as no surprise. Costs beyond permanent impact suggest how much execution suffered due to adverse selection. This could include costs for execution immediacy, front-running beyond the pricing (pre-trade) period, and trading against others with superior information.

Gaming the PI and AS metrics is more difficult since they do not refer to prices from one moment in time. However, as with II and DI, liquidity provision can skew PI and AS. For example, a broker providing liquidity at trading period end and exiting that position in the following period essentially shifts the last part of trading to the next period. Liquidity provision biases the next period's VWAP to make PI appear larger and AS appear smaller. (This may also slightly reduce IS.)

For a parent order metric focused on optimizable aspects of trading, we can examine $AS+\frac{PI}{2}(1+1/n)$. This yields a metric wholly based on decaying and temporary impact — factors we can control through effective order scheduling and trading acumen. However, the intuition behind this metric is not clear nor is it immune from gaming: liquidity provision by an external broker would reduce this metric, albeit less so than for either AS or PI alone.

5.2. **Interpreting Intertemporal Metrics.** Intertemporal metrics allow us to distinguish between skill at executing trades, being suitably patient, or scheduling child orders. If we can purchase these individual skills separately, these metrics let us evaluate different managers for each task.

Trading shortfall is a function of temporary impact and, therefore, the rate of trading $\tilde{q}_j/t_j$. A skilled trader probably adapts the rate of trading to market conditions and should have a consistently small trading shortfall. This small trading shortfall should be consistent across instruments, dates, and intraday. Similarly, a trader who is disciplined but poor at execution should have a consistently large trading shortfall.

A trader who is neither disciplined nor skilled at execution should have a noisy and inconsistent trading shortfall. This is crucial to discern: unpredictability yields more orders filled at both poor and excellent prices. Orders filled at excellent prices give inconsistent traders stories of seemingly superior execution. These stories may sound convincing; but, with these metrics, the lack of skill is apparent.

We can render TS uninformative by being all or none of the volume in a bin: that bin's trading shortfall is then 0. If we are most of the bin volume, the shortfall will be close to 0. Very poor traders can mask their lack of skill by only trading heavily at illiquid times. This would likely yield a poor implementation shortfall but a small trading shortfall. Barring such blatant manipulation, this metric should be valid for external use.

Trading shortfall also lets us extend the idea of information leakage to the period during trading. If a front-runner is averse to liquidity risk, a natural strategy is to front-run earlier in the period and exit the accumulated position later in the period. This realizes profits and reduces liquidity risk by trading against known liquidity (the remainder of the order).

The front-runner's inventory accumulation and liquidation distorts our executions versus fair prices. An order front-run in this manner should exhibit large positive trading shortfalls in the beginning of the trading period and small or negative trading shortfalls in the end of the trading period.

We can test if the shortfall pattern merits further attention by sorting shortfalls and counting the number b in the worst half of performance and the first half-trading period. The probability of b such "losing bins" is

$$P(b \text{ of worst half in first half-period}) = \binom{n/2}{b} \frac{1}{2^{n/2}}. \tag{18}$$

A trader who places limit orders provides liquidity to the market and sacrifices execution immediacy. A superior ability to be properly patient (a "cool hand") should yield a consistently negative fill time shortfall.

Overly passive traders are likely to receive superior prices for low amounts early in the trading period and inferior prices later as they (aggressively) "catch up" to the desired fill schedule. Thus overly passive traders should have negative fill time shortfalls in earlier time bins and positive fill time shortfalls in later bins. More sophisticated over-passive trader might "catch up" periodically. If the order schedule is specified before trading, the fill time shortfall should be valid for external use.

A trader who schedules child orders to achieve the lowest average execution cost should have a consistently small order timing shortfall. Traders with a consistently large order timing shortfall should examine their scheduling of child orders. If the child order schedule is designed to match the average volume $\bar{D}_j$, then the order timing shortfall is zero and cannot be taken as informative. This metric may be used externally so long as the child order schedule is not specified after trading.

A trader with a superior ability to predict the divergences from the average volume distribution should have a consistently small or negative volume shortfall. However, these divergences are often caused by unpredictable events such as news and large orders. This makes the volume shortfall noisy and suggests separating it from the other metrics. Lacking an ability to prediction these divergences, we should interpret this metric as noise.

6. Example Analysis

To clearly show the value of these metrics, we analyze two streams of orders and trades. We choose a ten-bin day with each bin encompassing, on average, 10% of the daily share volume. Prices are simulated with price impact as in Section 4.2 due to buying and selling noise traders. Price impact also occurs due to two strategic traders who split their net buy orders to execute them over the day.

The example day is shown in Table 4. We can then see how the two different traders perform across this day. While amounts that differ by one share are unusual, we can merely consider these to be the number of 100-share lots traded with results differing only by a common constant of proportionality.

Bin (j)	Day 1 Buy	Day 1 Sell	Day 1 VWAP	Day 2 Buy	Day 2 Sell	Day 2 VWAP
1	2,262	590	20.00288	2,290	4,387	19.97238
2	1,286	1,808	20.00508	1,165	1,149	19.97101
3	423	941	20.00323	5,084	3,811	19.97435
4	1,233	495	20.00503	5,468	7,697	19.97460
5	4,151	4,824	20.00462	2,432	1,756	19.97238
6	2,592	4,882	20.00091	3,919	4,131	19.97304
7	1,703	10,050	19.98758	270	2,604	19.97080
8	657	1,010	19.96862	1,574	3,368	19.96391
9	174	1,146	19.97139	2,215	1,137	19.96402
10	1,533	1,785	19.97230	3,682	1,306	19.97067

TABLE 4. Noise buying, noise selling, and volume-weighted average prices (VWAPs) for ten time bins/day of trading. On average, each bin contains $\bar{D} = 10\%$ of the day's volume.

Next we consider the actions of our two traders who trade on day 1 (Table 5). Trader A hews closely to the given order schedule with some mild "front-loading" (ordering more earlier). Trader B deviates aggressively in an attempt to achieve better execution prices. Both are constrained to fill the order in the ten periods.

We then calculate volumes (noise buy + noise sell + trader A buy + trader B buy), average prices, and the day's VWAP. We get that trader A's average price was \$20.00416 and trader B's average price was \$20.00234. The day 1 VWAP was \$19.99443 and the day 2 VWAP was \$19.97181. From these

	Trader A		Average	Trader B		Average
Bin (j)	Orders	Buys	Price A	Orders	Buys	Price B
1	130	130	20.01878	100	80	20.01444
2	125	120	20.00823	100	85	20.00678
3	120	125	20.01443	100	95	20.01219
4	115	105	20.01274	100	85	20.00669
5	110	115	20.01825	100	95	20.01744
6	100	100	20.00664	100	100	20.00247
7	90	95	19.99240	100	100	19.98409
8	80	75	19.97687	100	110	19.96689
9	70	75	19.97998	100	115	19.97636
10	60	60	19.97970	100	135	19.97726

Table 5. Trader A and B's orders, buys, and average prices for ten time bins on day 1. Trader A and B each purchase 1,000 shares over the ten periods. Trader B's orders are as per the average volume distribution (constant at $\bar{D} = 10\%$).

we compute parent order metrics. These metrics (Table 6) show that trader B has a lower implementation shortfall than trader A. Most analyses would conclude that trader B is more skilled than trader A. However, trader A has lower incremental and decaying impact than trader B, suggesting trader B was more aggressive at the end of day 1.

Trader	IS	II	DI	AS	PI
A	\$4.16	-\$24.46	\$7.89	\$32.35	-\$28.19
B	\$2.34	-\$17.08	\$13.45	\$30.53	-\$28.19

Table 6. Trader A and B's parent order metrics (per share) for ten time bins of trading in day 1 and ten time bins without their trading in day 2.

We can get more information if we use the trade data to compute the intertemporal metrics. In that case, we see where each trader is better and worse. Table 7 shows that trader A traded slightly better than trader B ($TS^A < TS^B$), that trader B was a better short-term market timer than trader A ($FTS^B < FTS^A$), and that trader B is a better order scheduler that trader A ($OTS^B < OTS^A$). Therefore, if we could split their duties, we might prefer to have trader B schedule orders, trader A trade the order split according to that schedule, and trader B to advise trader A on when to be patient with short-term market conditions.

	Trader A			Trader B			
Bin (j)	TS_j^A	FTS_j^A	OTS_j^A	TS_j^B	FTS_j^B	OTS_j^B	VS_j
1	1.42	0.00	600.09	0.92	-400.06	0.00	1427.85
2	-0.22	-100.03	500.13	0.14	-300.08	0.00	1383.62
3	0.77	100.02	400.06	0.85	-100.02	0.00	1704.12
4	0.29	-200.05	300.08	0.14	-300.08	0.00	1641.89
5	0.99	100.02	200.05	1.22	-100.02	0.00	283.14
6	0.07	0.00	0.00	0.16	0.00	0.00	565.48
7	-0.02	99.94	-199.88	-0.35	0.00	0.00	-233.47
8	0.24	-99.84	-399.37	-0.19	199.69	0.00	1651.32
9	0.27	99.86	-599.14	0.57	299.57	0.00	1715.42
10	0.14	0.00	-798.89	0.67	699.03	0.00	1341.44
Total	3.96	-0.08	3.12	4.14	-1.96	0.00	11480.80

TABLE 7. Trader A and B's trading shortfalls, fill time shortfalls, and order time shortfalls along with the volume shortfalls (common to all traders) for ten time bins of trading. Trader B's orders are identical to the average volumed distribution which is constant at $\bar{D} = 10\%$. Totals may differ due to rounding.

If we regress trader A's or trader B's trading shortfalls on time, the slope is not significant; however, the mean trading shortfalls are significantly different from zero for both traders. This suggests that they are not sloppy traders but are disciplined. In this example, however, a t-test cannot reject the null hypothesis that both traders possess similar skill at trading.

If we regress trader A's order timing shortfall on time, the slope is highly significant: -157.47 with t=8.8; trader A's scheduling does better as the day progresses. This suggests trader A might want to shift some trading from the start of the day to the end. If we regress trader B's fill time shortfall on time, the slope is highly significant: 100.53 with t=6.8; trader B's short-term market timing skill decreases as the day wears on. Thus trader B might want to limit market timing to the start of the trading day.

Finally, we can note that the volume shortfall numbers dominate the other metrics. This is because they are noise; and, their relative size shows the noise reducing powers of these intertemporal performance metrics.

7. Conclusion

We have introduced metrics to augment the implementation shortfall. These metrics offer more information for specialized applications, have sensible and intuitive interpretations, and are based on parent and child orders. Some metrics may be gamed and should only be used for internal performance evaluation; others resist gaming and may be used to evaluate the performance of external execution providers.

Gameable metrics can help in optimizing internal trading processes to reduce execution costs. Gaming-resistant metrics can help money managers reduce trading costs by choosing algorithmic trading software or external traders which perform best at various skills. These metrics could also reveal changes in those skills. In particular, intertemporal (child order) metrics let us measure skills at execution, short-term market timing, and child order scheduling. By examining the significance of these metrics, we can test if a trader is sloppy or undisciplined in one of these areas or if two traders have statistically discernable differences in skills. We can also measure how these metrics vary with market conditions and time. While we assume no model for such variation, we can still discern significant correlations between these skills and advantageous prices. These correlations (or linear regressions of shortfalls) may offer guidance as to how we can improve our overall performance. The time trends of these metrics can also help detect the times when a trader's skill is strongest and traders who are overly passive.

Some metrics have clear relations to a price impact model with lasting and various types of ephemeral price impact. Researchers may estimate these model parameters with the help of these metrics. Those parameters may help traders to lower execution costs and academics to compare execution quality across time or markets. Other metrics suggest basic tests of whether someone is front-running or "trading alongside" our order.

Simulations suggest use of these metrics could yield savings of about 2 basis points per trade for both trading and order scheduling skills. For typical turnover figures, a typical fund could save 13 basis points per year. That amounts to almost a 15% reduction in expenses and, for US-domiciled equity mutual funds alone, a savings of $7.3 billion annually. Were we to consider international equity funds, bond funds, pension funds, and insurance companies, the savings would be much larger.

Better execution also benefits the market overall. Hendershott *et al.* (2011) show that increasing use of algorithmic trading has narrowed spreads, reduced adverse selection, reduced trade-related price discovery and thus has improved liquidity and made quotes more informative. Easley and O'Hara (2010) note that the benefits of algorithmic trading may also increase investor participation in the market and lower the cost of capital for firms accessing the capital markets. We would expect this to also increase allocative efficiency since risk capital can seek out ventures with less frictions. Therefore, helping firms to trade better should benefit actors in both the primary and secondary markets.

References

Almgren, R. and N. Chriss. (2001)., "Optimal Execution of Portfolio Transactions." *Journal of Risk*, 3, 5–39.

Berke, L. J. (2010)., "US Institutional Equity Brokerage 2010: Assets, Commission Management and Concentration." Research report, TABB Group.

Bertsimas, D. and A. W. Lo. (1998)., "Optimal Control of Execution Costs." *Journal of Financial Markets*, 1, 1–50.

Easley, D. and M. O'Hara. (2010)., "Microstructure and Ambiguity." *Journal of Finance*, 65, 1817–1846.

Engle, R. F. and R. Ferstenberg. (2007)., "Execution Risk." *Journal of Portfolio Management*, 33, 34–44.

Fama, E. F. and K. R. French. (2010)., "Luck versus Skill in the Cross-Section of Mutual Fund Returns." *Journal of Finance*, 65, 1915–1947.

Hendershott, T., C. M. Jones, and A. J. Menkveld. (2011)., "Does Algorithmic Trading Improve Liquidity?" *Journal of Finance*, 66, 1–33.

Heston, S. L., R. A. Korajczyk, and R. Sadka. (2010)., "Intraday Patterns in the Cross-Section of Stock Returns." *Journal of Finance*, 65, 1369–1407.

Honoré, A. (2009)., "Smart Order Router Supply and Demand: Everything to Everyone." Research report, Aite Group.

Huberman, G. and W. Stanzl. (2004)., "Price Manipulation and Quasi-Arbitrage." *Econometrica*, 74, 1247–1276.

Investment Company Institute (2010)., *2010 Investment Company Fact Book*. Washington, DC.

Kissell, R. and R. Malamut (2005)., "Understanding the Profit and Loss Distribution of Trading Algorithms." In "Algorithmic Trading: Precision, Control, Execution," 41–49, New York: Institutional Investors Guides.

Kyle, A. S. (1985)., "Continuous Auctions and Insider Trading." *Econometrica*, 53, 1315–1336.

Lehmann, B. N. (2003)., "What We Measure in Execution Cost Measurement." *Journal of Financial Markets*, 6, 227–231.

Lucchetti, A. (2005)., "NYSE Probes Firms for Possible Improper Trading." *Wall Street Journal*.

McPartland, K. (2010)., "Data Center Networking: Redefining the Total Area Network." Research report, TABB Group.

Obizhaeva, A. and J. Wang (2006)., "Optimal Trading Strategy and Supply/Demand Dynamics." Working paper, Massachusetts Institute of Technology.

Opiela, N. (2006)., "Hype and Algorithms: Is Algorithmic Trading the Way of the Future or Just 'Okay'?" *CFA Magazine*, 46–47.

Perold, A. F. (1988)., "The Implementation Shortfall: Paper Versus Reality." *Journal of Portfolio Management*, 14, 4–9.

Puckett, A. and X. S. Yan. (2011)., "The Interim Trading Skills of Institutional Investors." *Journal of Finance*, 66, 601–633.

Department of Finance, University of Illinois at Chicago, daler@uic.edu