Transaction Cost Analytics in DRIP

Lakshmi Krishnamurthy

v2.50 7 November 2016

Execution of Portfolio Transactions – Optimal Trajectory

Overview, Scope, and Key Results

1. Portfolio Transactions under Market Impact: Almgren and Chriss (2000) consider the execution of portfolio transactions with the aim of minimizing a combination of volatility risk and transaction costs arising from temporary and permanent market impact.
2. Efficient Frontier under Linear Cost: For a simple linear cost model, they explicitly construct an *efficient frontier* in the space of time-dependent liquidation strategies, which have the minimum expected cost for a given level of uncertainty.
3. Choice of the Utility Function: This enables one to select optimal strategies either by minimizing a quadratic utility function, or by minimizing the Value-at-Risk.
4. Liquidity Adjusted Value at Risk: The latter choice leads to the concept of liquidity-adjusted VaR, or L-VaR, that explicitly considers the best trade-off between the volatility risk and the liquidity costs.

Motivation Background, and Synopsis

1. Transactions Changing the Portfolio Composition: Almgren and Chriss (2000) consider the optimal execution of portfolio transactions that move a portfolio from a given staring composition to a specified final composition within a specified period of time.
2. The Bertsimas and Lo Approach: Bertsimas and Lo (1998) define the best execution as the dynamic trading strategy that provides the minimum cost of trading over a fixed period of time, and they also show that under a variety of circumstances one can

find such a strategy by employing a dynamic optimization procedure; but they ignore the volatility of revenues of different trading strategies.

3. Maximization of Expected Trading Revenue: Almgren and Chriss (2000) work in the more general framework of maximizing the *expected revenue* – or equivalently minimizing the costs – with a suitable penalty for the *uncertainty* of revenue (or cost).
4. Market Microstructure Framework: This general framework arises in the market microstructure theory, but with a different purpose in mind. The *uninformed discretionary trader* trades an exogenous endowment over an exogenously specified amount of time to maximize the profits (Admati and Pfleiderer (1988)); the informed strategic trader trades over multiple periods on information not widely available, again to maximize profits (Kyle (1985)). In both cases the literature focuses on the link between the trader and the market maker, and a theory is produced to predict the market clearing price of the security at each period. Thus a trader's optimal strategy is used as a means to study the price formation in the markets, not as an object of interest in itself.
5. Variance of the Trading Cost: Almgren and Chriss (2000) study the variance of the trading cost in optimal execution because it fits in with the intuition that the trader's utility should figure in the definition of *optimal* in "optimal execution".
6. Example: Trading Illiquid Volatile Securities: For example, in trading a highly illiquid, volatile security, there are two extreme outcomes; trade everything now at a known, but high, cost, or trade in equal sized packets over a fixed time at a relatively lower cost. The latter strategy has a lower expected, but this comes at the expense of greater uncertainty in the final revenue.
7. Estimation of the Trading Uncertainty: How to evaluate the above uncertainty is partly subjective, and is a function of the trader's tolerance for risk. All that can be done is to insist that for a given level of uncertainty that the cost be minimized. This idea extends to a complete theory of optimal execution that includes an efficient frontier of optimal execution strategies.
8. Consistency with Expectations from Intuition: The framework of risk in execution yields several results that are consistent with the intuition. For example, it is evident

that all else equal, a trader will choose to execute a block of illiquid security less rapidly than a liquid security.

9. Models Lacking Consistency with Intuition: While this seems obvious, Almgren and Chriss (2000) demonstrate that a model that ignores risk does not have this property; without enforcing a strictly positive penalty for risk one cannot produce models that trade differently across the spectrum of liquidity.
10. Arithmetic Brownian Motion Price Dynamics: The incorporation of risk into optimal execution does not come without cost. First, in order to be able to produce tractable analytical results, Almgren and Chriss (2000) are forced to work in largely in the framework of price dynamics that are an arithmetic walk with independent increments.
11. Use of Static Optimization Procedures: They obtain results using *static optimization* procedures which they show lead to globally optimal trading trajectories. That is, optimal trading paths may be determined in advance of trading. Only the composition of the portfolio and the trader's utility function figure on the trading path.
12. Why does Static Optimization Work? The fact that the static strategy can be optimal even when the trader has the option to dynamically change his trading mid-course is a direct result of the assumptions of independence of returns and symmetry for the penalty functions for risk.
13. Using Non-Symmetric Penalty Functions: An interesting deviation from the symmetric penalty function was communicated by Ferstenberg, Karchmer, and Malamut at ITG Inc. They argue that the opportunity is a subjective quantity and is measures differently by different traders. Using a trader defined cost function g, they define opportunity costs as the expected costs of g applied to the average execution price obtained by the trader relative a benchmark price. They assume that the risk-averse traders will use a convex function g that is not symmetric in the sense that there is a strictly greater penalty for underperformance than for the same level of outperformance. They show that in this setting, the optimal strategy relative to g not only depends on the time remaining, but also on the performance of the strategy up to the present time, and the present price of the security. In particular, this means that in their setting, optimal strategies are dynamic.

14. Serial Correlations among Price Movements: As it is well known that price movements exhibit some serial correlations across various time horizons (Lo and MacKinlay (1988)), that market conditions change, and that some participants possess private information (Bertsimas and Lo (1998)), one may question the usefulness of results that obtain strictly in an independent-increment framework.
15. The Dynamic Nature of Trading: Moreover, as trading is known to be a dynamic process, the conclusion that optimal trading strategies can be statically determined calls for critical examination. Almgren and Chriss (2000) examine what quantitative gains are available that incorporate all the relevant information.
16. Impact of the Serial Correlations: First they consider short term serial correlations in price movements. They demonstrate that the marginal improvements available by explicitly incorporating this information into trading strategies is small, and more importantly, independent of the portfolio sizes; as portfolio sizes increase, the percentage gains possible decrease proportionately.
17. Combining "Correlated" and "Shifting" Strategies: The above is precisely true for linear transaction cost models, and is approximately true for more general models. The results of Bertsimas and Lo (1998) suggest that trading a strategy built to take advantage of serial correlation will essentially be a combination of a "correlation free" strategy and a "shifting strategy" that moves from one trade period to the next based on the information available in the last period's return. Therefore Almgren and Chriss (2000) argue that by ignoring serial correlation, they a) preserve the main interesting features of their analysis, and b) introduce virtually no bias away from "truly optimal" solutions.
18. Impact of Scheduled News Events: Second, Almgren and Chriss (2000) examine the impact of scheduled new events on optimal execution strategies. There is ample evidence that anticipated news announcements, depending on their outcome, can have a significant temporary impact on the parameters governing price movements.
19. Scheduled News Events - Literature Review: For a theoretical treatment see Brown, Harlow, and Tinic (1988), Kim and Verrecchia (1991), Easterwood and Nutt (1999), and Ramaswami (1999). For empirical studies concerning earnings announcements, see Patell and Wolfson (1984) for changes in mean and variance of intra-day prices,

and Lee, Mucklow, and Ready (1993) and Krinsky and Lee (1996) for changes in the bid-ask spread. For additional studies concerning news announcements, see Charest (1978), Morse (1981), and Kalay and Loewentstein (1985).

20. Model Incorporation of Scheduled Events: Almgren and Chriss (2000) work in a simple extension of their static framework by assuming that the security again follows an arithmetic random walk, but at a time known at the beginning of trading, an uncorrelated event will cause a material shift in price dynamics, e.g., an increase or decrease of volatility.
21. Combining Piece-Wise Static Strategies: In this context they show that optimal strategies are piece-wise static. To be precise, they show that an optimal strategy entails following a static strategy up to the moment of the event, followed by another static strategy that can only be determined once the outcome of the event in known.
22. Variation from the Original Static Strategy: It is interesting to note that the static strategy that one follows in the first leg is in general not the same strategy one would follow in the absence of information concerning the event.
23. Accommodating Unanticipated External "Sudden' Events: Finally Almgren and Chriss (2000) note that any optimal execution strategy is vulnerable to *unanticipated events*. If such an event occurs during the course of trading and causes a material shift in the parameters of the price dynamics, then indeed a shift in the optimal trading trajectory must also occur.
24. Adaptation at Parameter Shift Edges: However if one makes a simplifying assumption that all events are either "scheduled" or "unanticipated" one then concludes that optimal execution is always a game of static trading punctuated by shifts in the trading strategies that adapt to material changes in the price dynamics.
25. Pre-determined vs. Active Approaches: If shifts are caused by events that are known ahead of time, then optimal execution benefits from a precise knowledge of the possible outcomes of the event. If not, the best approach is to be actively "watching" the market for such changes and react swiftly should they occur.
26. Simple Proxy for Unexpected Uncertainty: One approximate way to include such completely unexpected uncertainty into the model is to artificially raise the value of the volatility parameter.

27. Risk Averse Optimal Trading Strategies: As a first step, Almgren and Chriss (2000) obtain closed form solutions for trading strategies for any level of risk aversion.
28. Efficient Frontier of Optimal Strategies: They then show that this leads to an efficient frontier of optimal strategies, where an element of the frontier is represented by a strategy with a minimal level of cost for its level of variance of the cost.
29. Graphical Structure of the Frontier: The structure of the frontier is of some interest. It is a smooth convex function differentiable at its minimal point. The minimal point is what Bertsimas and Lo (1998) call the naïve strategy because is corresponds to trading equally sized packets using all available trading time equally.
30. Differential at the Minimum Point: The differentiability of the frontier at its minimum point indicates that one can obtain a first order reduction in the variance of the trading cost at the expense of only a second order in cost by trading a strategy slightly away from the globally minimal strategy.
31. Curvature at the Minimal Point: The curvature of the frontier at its minimum point is a measure of the liquidity of the security.
32. Half-Life of Optimal Execution: Another ramification of the Almgren and Chriss (2000) study is that for all levels of risk aversion except risk neutrality, optimal execution trades have a "half-life" that fall out of the calculations.
33. Independence form the Time to Complete Execution: A trade's half-life is independent of the actual specified time to liquidation, and is a function of the security's liquidity and volatility, and the trader's level of risk aversion.
34. Half-Life as Execution Time: As such Almgren and Chriss (2000) regard the half-life as an idealized time to execution, and perhaps a guide to the proper amount of time over which to execute a transaction.
35. Time Lesser than Half Life: If the specified time to liquidation is short relative to the trade's half-life, once can expect the cost of trading to be dominated by transaction costs.
36. Time Greater than Half Life: If the time to trade is long relative to the half-life, one can then expect most of the liquidation to take place well in advance of the limiting time.

The Definition of a Trading Strategy

1. Price Dynamics and Trade Execution: As a starting point, Almgren and Chriss (2000) define s trading strategy, and lay out the dynamics that they study. They start with a formal definition of a strategy for a sell program consisting of liquidating a single security. The definitions and results are analogous for a buy program.
2. Problem Setup - Security Liquidation: Suppose that the seller holds a block of X units of a security that they want to completely liquidate before time T. To keep the discussion, Almgren and Chriss (2000) speak of *units* of a security. Specifically they have in mind shares of stock, futures contract, and units of a foreign currency.
3. Trading Strategy - Price/Unit Strategy: The seller divides T into N units of length

   $$\tau = \frac{T}{N}$$

   and defines the discrete times

   $$t_k = k\tau$$

   for

   $$k = 0, \cdots, N$$

   The *trading trajectory* is defined to be the list $x_0, \cdots, x_N$ where x_k is the number of units that the seller plans to hold at time t_k.
4. Outright/Re-balanced Trajectories: The initial holdings is

   $$x_0 = X$$

   and liquidation at time T requires

$$x_N = 0$$

A trading trajectory can thought of as either the ex-post realized trades resulting from some process, or as a plan concerning how to trade a block of securities. In either case one may also consider *re-balancing* trajectories by requiring

$$x_0 = X$$

the initial position, and

$$x_1 = Y$$

the new position, but this is formally equivalent to studying trajectories of the form

$$x_0 = X - Y$$

and

$$x_N = 0$$

5. Outstanding Holdings/Incremental Trade Lists: Equivalently, a strategy may be specified using the “trade list” $n_1, \cdots, n_N$ where

$$n_k = x_{k-1} - x_k$$

is the number of units that the seller will sell between times t_{k-1} and t_k. Clearly, x_k and n_k are related by

$$x_k = X - \sum_{j=1}^{k} n_j = \sum_{j=k+1}^{N} n_j$$

$$k = 0, \cdots, N$$

6. Simultaneous Portfolio Buying and Selling: Almgren and Chriss (2000) also consider more general programs of buying and selling simultaneously several securities.
7. Inter-Execution Time Interval Specification: For notational simplicity they consider all the time intervals to be of equal length τ, but this restriction is not essential.
8. Behavior at N/τ Limits: Although they do not discuss it, in all their results it is easy to take the continuous-time limit of

$$N \to \infty$$

and

$$\tau \to 0$$

9. Definition of a Trading Strategy: Almgren and Chriss (2000) define a "trading strategy" to be a rule for determining n_k in terms of the information available at t_{k-1}. Broadly speaking they distinguish between two types of trading strategies – static and dynamic.
10. Static vs. Dynamic Trading Strategy: Static strategies are determined in advance of trading, that is the rule for determining each n_k depends only on information available at t_0. Dynamic strategies, conversely, depend on all information up to, and including, time t_{k-1}

Price Dynamics

1. Exogenous/Endogenous Price Move Factors: Suppose that the initial security price is S_0 so that the initial market value of the position is XS_0. The securities' price evolves

according to two exogenous factors – volatility and drift, and one endogenous factor – market impact.

2. <u>Market Forces vs. Trading Impact</u>: Volatility and drift are assumed to the result of market forces that occur randomly and independent of the trading.
3. <u>Earlier Literature on Market Impact</u>: Almgren and Chriss (2000) discussion s largely reflect the work of Kraus and Stoll (1972), and the subsequent works of Holthausen, Leftwich, and Mayers (1987, 1990) and Chan and Lakonishok (1993, 1995). See also Keim and Madhavan (1995, 1997).
4. <u>Origin of the Market Impact</u>: As the market participants begin to detect the volume that the seller (buyer) is selling (buying), they naturally adjust their bids (offers) downward (upward). Almgren and Chriss (2000) distinguish two kinds of market impact.
5. <u>Definition of Temporary Market Impact</u>: *Temporary* impact refers to the temporary imbalances in supply and demand caused by the seller's trading leading to temporary price movements away from equilibrium.
6. <u>Definition of Permanent Market Impact</u>: *Permanent* impact refers to the changes in the "equilibrium" price due to the seller's trading, which remain at least for the life of the liquidation.
7. <u>Price Evolution Stochastic Difference Equation</u>: Almgren and Chriss (2000) assume that the security price evolves according to the discrete random walk

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k - \tau g\left(\frac{n_k}{\tau}\right)$$

for

$$k = 1, \cdots, N$$

8. <u>Glossary of the Equation Terms</u>: Here σ represents the volatility of the asset, ξ_k's are draws from independent random variables each with zero mean and unit variance, and the permanent impact function $g(v)$ is a function of the *average rate* of trading

$$v = \frac{n_k}{\tau}$$

during the interval t_{k-1} to t_k.

9. Lack of Explicit Drift Term: In the above equation there is no drift term. Almgren and Chriss (2000) indicate that this is due to the assumption that they have no information about the direction of the future price movements.
10. Trading Term Horizons under Consideration: Over long term investment time scales, or in extremely volatile markets, it is important to consider *geometric* rather than arithmetic Brownian motion – this corresponds to letting σ in

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k - \tau g\left(\frac{n_k}{\tau}\right)$$

scale with S. But over short term "trading" horizons of interest, the total fractional price changes are small, and the differences between arithmetic and geometric Brownian motions are negligible.

Temporary Market Impact

1. Intuition behind the Temporary Market Impact: The intuition behind the temporary market impact is that a trader plans to sell a certain number of units n_k between times t_k and t_{k-1}, but may work the order in several smaller sizes to locate optimal points of liquidity.
2. Liquidity Reduction Impact on Price: If the total number of units n_k is sufficiently large, the execution price may steadily decrease between t_{k-1} and t_k in part due to the exhaustion of the supply of liquidity at each successive price level. This effect is assumed to be short-lived, and in particular, liquidity is assumed to return back after each period, and a new equilibrium price is established.

3. The Temporary Price Impact Function: This effect is modeled by introducing a temporary price impact function $h(v)$, the temporary drop in the average price per share caused by trading at an average rate v during one time interval.
4. Net Price Received at Execution: Give this, the actual price per share received on sale k is

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau}\right)$$

but the effect of $h(v)$ does not appear in the next "market" price S_k.
5. Choice of Market Microstructure: The functions $g(v)$ in

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k - \tau g\left(\frac{n_k}{\tau}\right)$$

and $h(v)$ in

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau}\right)$$

may be chosen to reflect any preferred model of market microstructure, subject only to certain natural convexity conditions.

Capture and Cost of Trading Trajectories

1. Capture across a Trading Trajectory: Almgren and Chriss (2000) then discuss the profits resulting from trading along a certain trajectory. They define the *capture* of a trajectory to be the full trading revenue upon completion of all trades. Due to the short term horizons that they consider, they do not include any notion of carry or time value of money in their discussions.

2. Full Trading Revenue across Execution: Thus, the capture is the sum of the product of the number of units n_k sold in each time interval times the effective price per share $\tilde{S}_k$ received on that sale. It is readily computed as

$$\sum_{k=1}^{N} n_k \tilde{S}_k = XS_0 + \sum_{k=1}^{N} \left[\sigma\sqrt{\tau}\xi_k - \tau g\left(\frac{n_k}{\tau}\right)\right] x_k + \sum_{k=1}^{N} n_k h\left(\frac{n_k}{\tau}\right)$$

3. Decomposition of the Capture Components: The first term on the RHS above is the initial market value of the position; each additional term represents a gain or a loss due to a specific market factor.
4. The Volatility Price Impact Term: The first term $\sigma\sqrt{\tau}\xi_k x_k$ represents the total impact from the volatility.
5. The Permanent Market Impact Term: The permanent market impact term $-\tau x_k g\left(\frac{n_k}{\tau}\right)$ represents the loss in the value of the position caused by a permanent price drop associated with selling a small piece of the position.
6. The Temporary Market Impact Term: And the temporary market impact term $n_k h\left(\frac{n_k}{\tau}\right)$ is the price drop due to selling, acting only on the units sold during the k^{th} period.
7. The Total Cost of Trading: The *total cost of trading* is the difference $XS_0 - \sum_{k=1}^{N} n_k \tilde{S}_k$ between the initial book value and the capture. This is the standard *ex-post* measure of the performance costs used in performance evaluations, and is essentially what Perold (1988) calls *implementation shortfall*.
8. Estimation of Implementation Short-fall: In this model, prior to trading, the implementation short-fall is a random variable. Write $\mathbb{E}[X]$ for the expected short-fall and $\mathbb{V}[X]$ for the variance of the short-fall.
9. Implementation Short-fall Mean/Variance: Given the simple nature of price dynamics, Almgren and Chriss (2000) readily compute

$$\mathbb{E}[X] = \sum_{k=1}^{N} \tau x_k g\left(\frac{n_k}{\tau}\right) + \sum_{k=1}^{N} n_k h\left(\frac{n_k}{\tau}\right)$$

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

The units of $\mathbb{E}[X]$ are in dollars, and the units of $\mathbb{V}[X]$ are dollars squared.

10. Distribution of Implementation Short-fall: The distribution of the short-fall is Gaussian if ξ_k is Gaussian, in any case if N is large, it is very nearly Gaussian.
11. Almgren and Chriss Minimizer Utility: Almgren and Chriss (2000) devote much of their paper to finding trajectories that minimize $\mathbb{E}[X] + \lambda\mathbb{V}[X]$ for various values of λ. They demonstrate that for each value of λ there corresponds a unique trading trajectory x such that $\mathbb{E}[X] + \lambda\mathbb{V}[X]$ is minimal.

Linear Impact Functions

1. Linear Temporary/Permanent Market Impact: Although Almgren and Chriss (2000) formulation does not require it, computing optimal trajectories is significantly easier if one takes the permanent and temporary impact functions to be *linear* in the rate of trading.
2. Linear Permanent Impact Market Function: For linear permanent impact, $g(v)$ has the form

$$g(v) = \gamma v$$

   in which the constant γ has units of ($/share)/share.
3. Corresponding Execution Time Security Price: With this form, each n units sold depresses the price per share by γn regardless of the time taken to sell n units.

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k - \gamma g\left(\frac{n_k}{\tau}\right)$$

readily yields

$$S_k = S_0 + \sigma \sum_{j=1}^{k} \sqrt{\tau_j}\xi_j - \tau\gamma(X - x_k)$$

4. Permanent Implementation Short-fall Mean: Then summing by parts, the permanent impact term in

$$\mathbb{E}[X] = \sum_{k=1}^{N} \tau x_k g\left(\frac{n_k}{\tau}\right) + \sum_{k=1}^{N} n_k h\left(\frac{n_k}{\tau}\right)$$

becomes

$$\sum_{k=1}^{N} \tau x_k g\left(\frac{n_k}{\tau}\right) = \gamma \sum_{k=1}^{N} x_k n_k = \gamma \sum_{k=1}^{N} x_k (x_k - x_{k-1})$$
$$= \frac{1}{2}\gamma^2 \sum_{k=1}^{N} [{x_{k-1}}^2 - {x_k}^2 - (x_k - x_{k-1})^2] = \frac{1}{2}\gamma X^2 - \frac{1}{2}\gamma \sum_{k=1}^{N} {n_k}^2$$

5. Linear Temporary Impact Market Function: Similarly, for the temporary impact we take

$$h\left(\frac{n_k}{\tau}\right) = \epsilon sgn(n_k) + \frac{\eta}{\tau} n_k$$

where sgn is the sign function.

6. Estimating the Fixed Costs of Execution: The units of ϵ are $/share, and those of η are ($/share)/(share/time). A reasonable estimate for ϵ is the fixed cost of selling, such as half of bid-ask spread plus premium.
7. Estimating the Linear Impact Coefficient: It is more difficult to estimate η since it depends on the internal and the transient aspects of the market microstructure. It is in

this term that one would expect the on-linear terms to be most important, and the approximation

$$h\left(\frac{n_k}{\tau}\right) = \epsilon sgn(n_k) + \frac{\eta}{\tau} n_k$$

to be most doubtful.

8. Total Temporary Impact Function: The linear model

$$h\left(\frac{n_k}{\tau}\right) = \epsilon sgn(n_k) + \frac{\eta}{\tau} n_k$$

is often called a *quadratic* cost because the total costs incurred by buying or selling n units in a single unit of time is

$$nh\left(\frac{n}{\tau}\right) = \epsilon|n| + \frac{\eta}{\tau} n^2$$

9. Temporary Implementation Short-fall Mean: With both linear cost models

$$g(v) = \gamma v$$

and

$$h\left(\frac{n_k}{\tau}\right) = \epsilon sgn(n_k) + \frac{\eta}{\tau} n_k$$

the expectation of the impact costs

$$\mathbb{E}[X] = \sum_{k=1}^{N} \tau x_k g\left(\frac{n_k}{\tau}\right) + \sum_{k=1}^{N} n_k h\left(\frac{n_k}{\tau}\right)$$

becomes

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} {n_k}^2$$

in which

$$\tilde{\eta} = \eta - \frac{1}{2}\gamma\tau$$

10. Strictly Convex Nature of $\mathbb{E}[X]$: Clearly $\mathbb{E}[X]$ is a strictly convex function as long as

$$\tilde{\eta} > 0$$

Note that if n_k all have the same sign, as would be the case for a pure sell program or a pure buy program, then

$$\sum_{k=1}^{N} |n_k| = |X|$$

11. $\mathbb{E}[X]$ and $\mathbb{V}[X]$ Computation Illustration: To illustrate, Almgren and Chriss (2000) compute $\mathbb{E}[X]$ and $\mathbb{V}[X]$ for linear impact functions for two of trajectory schemes at the opposite extremes: sell at a constant rate, and sell to maximize variance without regard to transaction costs.
12. Minimum Impact: Constant Execution Rate: The most obvious trajectory is to sell at a constant rate over the entire liquidation period. Thus one takes each

$$n_k = \frac{X}{N}$$

and

$$x_k = (N-k)\frac{X}{N}$$

$$k = 1, \cdots, N$$

13. Minimum Impact $\mathbb{E}[X]$ and $\mathbb{V}[X]$: From

$$\mathbb{E}[X] = \sum_{k=1}^{N} \tau x_k g\left(\frac{n_k}{\tau}\right) + \sum_{k=1}^{N} n_k h\left(\frac{n_k}{\tau}\right)$$

and

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau}\sum_{k=1}^{N} {n_k}^2$$

one has

$$\mathbb{E}[X] = \frac{1}{2}XTg\left(\frac{X}{T}\right)\left(1-\frac{1}{N}\right) + Xh\left(\frac{X}{T}\right) = \frac{1}{2}\gamma X^2 + \epsilon X + \tilde{\eta}\frac{X^2}{T}$$

and from

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

$$\mathbb{V}[X] = \frac{1}{3}\sigma^2 X^2 T\left(1-\frac{1}{N}\right)\left(1-\frac{1}{2N}\right)$$

14. Minimum Impact N/T Limits: The trajectory minimizes total expected costs, but the variance may be large if the period T is long. As the number of trading periods

$$N \to \infty$$

$$v = \frac{X}{T}$$

remains finite, and $\mathbb{E}[X]$ and $\mathbb{V}[X]$ have finite limits.

15. Minimum Variance: One Step Execution: The other extreme is to execute the entire position in the first time step. One then takes

$$n_1 = X$$

$$n_2 = \cdots = n_N = 0$$

$$x_1 = \cdots = x_N = 0$$

which results in

$$\mathbb{E}[X] = Xh\left(\frac{X}{\tau}\right) = \epsilon X + \eta\frac{X^2}{\tau}$$

and

$$\mathbb{V}[X] = 0$$

16. Minimum Variance N/T Limits: The trajectory has the smallest possible variance – equal to zero – because of the way time has been discretized in the model above. If N is large and hence τ is short, then on the full initial portfolio, one takes a price hit that can be arbitrarily large.
17. Trajectory between the Two Extremes: Almgren and Chriss (2000) show how to effectively compute trajectories that lie between the two extremes.

The Efficient Frontier of Optimal Execution

1. Computing the Optimal Execution Trajectories: Almgren and Chriss (2000) define and compute optimal execution trajectories and use that to later demonstrate a precise relationship between risk aversion and the definition of optimality.
2. Uniqueness of Optimal Execution Strategy: In particular, they show that each level of risk aversion there is a uniquely determined optimal execution strategy.

The Definition of the Frontier

1. Minimization of Expected Short-fall: The rational trader will always seek to minimize the expectation of short-fall for a given level of variance of the short-fall. Naturally a trader will prefer a strategy that provides minimum error in its estimate of expected costs.
2. Efficient Optimal Trading Strategy Definition: Thus a strategy is *efficient* or *optimal* if there is no other strategy that has lower variance for the same or a lower variance of the expected transaction costs, or, equivalently, no strategy which has no lower expected transaction costs for the same or lower level of variance.
3. Static vs. Dynamic Strategy Optimality: This definition of optimality of a strategy is the same whether the strategy is static or dynamic. It will be established later that under this definition and the price dynamics already stated, optimal strategies are in fact static.
4. Efficient Strategies - Constrained Optimization Formulation: One may construct efficient strategies by solving the constrained optimization problem

$$\min_{x:\, \mathbb{V}[x] \le V_*} \mathbb{E}[x]$$

That is, for a given maximum level of variance

$$V_* \geq 0$$

one finds a strategy that has the minimum expected levels of transaction costs.

5. Convex Objective Function and Domain: Since $\mathbb{V}[x]$ is convex, the set

$$\{\mathbb{V}[x] \leq V_*\}$$

is convex – it is a sphere – and since $\mathbb{E}[x]$ is strictly convex, there is a unique minimizer $x_*(V_*)$.

6. Sub-Optimal Trajectory Variance Cost: Regardless of the preferred balance of risk and return, every other solution x which has

$$\mathbb{V}[x] \leq V_*$$

has higher expected costs than $x_*(V_*)$ for the same or lower variance, and can never be more efficient.

7. Efficient Frontier of Optimal Strategies: Thus the family of all possible efficient (optimal) strategies is parametrized by a single variable V_* representing all possible maximum levels of variance in transaction costs. This family is referred to as *the efficient frontier of optimal trading strategies*.

8. Introducing KKT Type Constraint Multipliers: The constrained optimization problem

$$\underset{x:\, \mathbb{V}[x] \leq V_*}{min} \mathbb{E}[x]$$

is solved by introducing a constraint multiplier λ, thereby solving the unconstrained problem

$$\underset{x}{min} \left(\mathbb{E}[x] + \lambda \mathbb{V}[x]\right)$$

9. Frontier as a Function of λ: If

$$\lambda > 0$$

$\mathbb{E}[x] + \lambda\mathbb{V}[x]$ is strictly convex, and the above minimizer has a unique solution $x^*(\lambda)$. As λ varies, $x^*(\lambda)$ sweeps out the same one parameter family, and thus traces out an efficient frontier.

10. <u>λ as a Risk Aversion Parameter</u>: The Parameter λ has a direct financial interpretation. It is already apparent from

$$\min_{x} \left(\mathbb{E}[x] + \lambda\mathbb{V}[x]\right)$$

that λ is a measure of risk aversion, that is, how much the variance is penalized relative to the cost.

11. <u>λ as an Efficient Frontier Curvature</u>: In fact, λ is the curvature – second derivative – of a smooth utility function, as will be made more precise eventually.

12. <u>Solution given $h(v)$ and $g(v)$</u>: For given values of the parameters, problem

$$\min_{x} \left(\mathbb{E}[x] + \lambda\mathbb{V}[x]\right)$$

can be solved by various numerical techniques depending on the functional forms chosen for $h(v)$ and $g(v)$. In the special case that these are *linear* functions, we may write the solution explicitly and gain a great deal of insight into the trading strategies.

Explicit Construction of Optimal Strategies

1. <u>Optimal Solution in Trajectory Space</u>: With $\mathbb{E}[x]$ from

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 + \epsilon\sum_{k=1}^{N}|n_k| + \frac{\tilde{\eta}}{\tau}\sum_{k=1}^{N}{n_k}^2$$

and $\mathbb{V}[x]$ from

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

and assuming that n_j does not change sign, the combination

$$\mathbb{U}[x] = \mathbb{E}[x] + \lambda \mathbb{V}[x]$$

is a quadratic function of the control parameters $x_1, \cdots, x_{N-1}$; it is strictly convex for

$$\lambda \geq 0$$

2. Finding the Unique Global Minima: Therefore one determines the unique global minimum by setting its partial derivatives to zero. One readily calculates

$$\frac{\partial \mathbb{U}[x]}{\partial x_j} = 2\tau \left(\lambda \sigma^2 x_j - \tilde{\eta} \frac{x_{j-1} - 2x_j + x_{j+1}}{\tau^2} \right)$$

for

$$j = 1, \cdots, N - 1$$

3. Combinations of Linear Difference Equations: Then

$$\frac{\partial \mathbb{U}[x]}{\partial x_j} = 0$$

is equivalent to

$$\frac{x_{j-1} - 2x_j + x_{j+1}}{\tau^2} = \tilde{\kappa}^2 x_j$$

with

$$\tilde{\kappa}^2 = \frac{\lambda\sigma^2}{\tilde{\eta}} = \frac{\lambda\sigma^2}{\eta\left(1 - \frac{\gamma\tau}{2\eta}\right)}$$

4. <u>τ Abstracted and Re-factored Parameter Set</u>: Note that

$$\frac{x_{j-1} - 2x_j + x_{j+1}}{\tau^2} = \tilde{\kappa}^2 x_j$$

is a linear difference equation whose solution may be written as a combination of the exponentials $e^{\pm\kappa t_j}$ where κ satisfies

$$\frac{2}{\tau^2}[\cosh(\kappa\tau) - 1] = \tilde{\kappa}^2$$

The tilde's on $\tilde{\eta}$ and $\tilde{\kappa}$ denote an $\mathcal{O}(\tau)$ correction; as

$$\tau \to 0$$

one has

$$\tilde{\eta} \to \eta$$

and

$$\tilde{\kappa} \to \kappa$$

5. <u>Trading Trajectory/Trade List Solutions</u>: The specific solution with

$$x_0 = X$$

and

$$x_N = 0$$

is a trading trajectory of the form

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X$$

$$j = 0, \cdots, N$$

and the associated trade list is

$$n_j = \frac{2 \sinh\left(\frac{1}{2}\kappa T\right)}{\sinh(\kappa T)} \cosh\left(\kappa\left(T - t_{j-\frac{1}{2}}\right)\right) X$$

$$j = 1, \cdots, N$$

where sinh and cosh are the hyperbolic sine and cosine functions, and

$$t_{j-\frac{1}{2}} = \left(j - \frac{1}{2}\right)\tau$$

These solutions – although not the efficient frontier – have been constructed previously by Grinold and Kahn (1999).

6. Monotonicity of the Trading Trajectory: One has

$$n_j > 0$$

as long as

$$X > 0$$

Thus for a program of selling a large initial long position, the solution decreases *monotonically* from its initial value to zero at the rate determined by the parameter κ.

7. Consequence of Monotonic Trading Trajectories: For example, the optimal execution of a sell program never involves buying of securities – although this ceases to be true if there is drift or serial correlation in price movements.
8. Approximation under Small Time Step: For a small time step τ one has the approximate expression

$$\kappa \sim \tilde{\kappa} + \mathcal{O}(\tau^2) \sim \sqrt{\frac{\lambda \sigma^2}{\eta \left(1 - \frac{\gamma \tau}{2\eta}\right)}} + \mathcal{O}(\tau)$$

$$\tau \rightarrow 0$$

Thus if the trading intervals are short κ^2 is essentially the ratio of the product of volatility and the risk-intolerance to the temporary transaction cost parameter.

9. Optimal Strategy Expected Cost/Variance: The expectation and the variance of the optimal strategy for a given initial portfolio size X are then

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 + \epsilon X + \tilde{\eta} X^2 \frac{\tanh\left(\frac{1}{2}\kappa\tau\right)[\tau \sinh(2\kappa T) + 2T \sinh(\kappa\tau)]}{2\tau^2 [\sinh(\kappa\tau)]^2}$$

and

$$\mathbb{V}[X] = \frac{1}{2}\sigma^2 X^2 \frac{\tau \sinh(\kappa T) \cosh\big(\kappa(T - \tau)\big) - T \sinh(\kappa\tau)}{[\sinh(\kappa T)]^2 \sinh(\kappa\tau)}$$

which reduce to

$$\mathbb{E}[X] = \frac{1}{2}XTg\left(\frac{X}{T}\right)\left(1 - \frac{1}{N}\right) + Xh\left(\frac{X}{T}\right) = \frac{1}{2}\gamma X^2 + \epsilon X + \tilde{\eta}\frac{X^2}{T}$$

$$\mathbb{V}[X] = \frac{1}{3}\sigma^2 X^2 T\left(1 - \frac{1}{N}\right)\left(1 - \frac{1}{2N}\right)$$

$$n_1 = X$$

$$n_2 = \cdots = n_N = 0$$

$$x_1 = \cdots = x_N = 0$$

$$\mathbb{E}[X] = Xh\left(\frac{X}{\tau}\right) = \epsilon X + \eta\frac{X^2}{\tau}$$

$$\mathbb{V}[X] = 0$$

in the limits

$$\kappa \to 0, \infty$$

The Half-Life of a Trade

1. <u>Definition of the Half-Life</u>: Defining

$$\theta = \frac{1}{\kappa}$$

the trade's "half-life", and using the discussion above, it can be seen that the larger the value of κ and smaller the value of θ, the more rapidly the trade list will be depleted. The value θ is exactly the amount of time it takes to deplete the holdings by a factor of e.

2. Half-Life Different from T: The definition of θ is independent of the exogenously specified execution time T; it is determined only by the security price dynamics and the market impact factors. If the risk aversion λ is greater than zero, i.e., if the trader is risk-averse, then θ is finite and independent of T.
3. Timeless Initial Portfolio Liquidation Rate: Thus, in the absence of any external time constraint, i.e.

$$T \to \infty$$

the trader will still liquidate his position on a time scale θ. The half-life θ is the intrinsic time scale of the trade.
4. Half Life Smaller than T: For a given T the ratio

$$\kappa T = \frac{T}{\theta}$$

tells us what factors constrain the trade. If

$$T \gg \theta$$

then the intrinsic half-life θ of the trade is small compared to the imposed time T; this happens because temporary costs are very small, because volatility is very large, or because of high risk aversion.
5. Impact of Small Half-Life: In this case the bulk of the trading will be done well in advance of the time T. Viewed on a time scale T the trajectory will look like a minimum variance solution

$$n_1 = X$$

$$n_2 = \cdots = n_N = 0$$

$$x_1 = \cdots = x_N = 0$$

6. Very High Half Life Limit: Conversely if

$$T \ll \theta$$

then the trade is highly constrained, and is dominated by temporary market impact costs. In the limit

$$\frac{T}{\theta} \to 0$$

one approaches the straight line minimum cost strategy

$$n_k = \frac{X}{N}$$

$$x_k = (N - k)\frac{X}{N}$$

$$k = 1, \cdots, N$$

7. Trade Size Independent Execution Strategy: A consequence of this analysis is that different sized baskets of the same security will be liquidated in exactly the same fashion, on the same scale, provided the risk aversion parameter λ is held constant.

8. Basket Size Based Liquidity Dependence: This may seem contrary to the expectation that large baskets are effectively less liquid, and should hence be liquidated less rapidly than smaller baskets.
9. Reasons for the Counter-Intuitiveness: This is a consequence of the linear market impact assumption which has the *mathematical* consequence that both variance and market impact scale quadratically with respect to the portfolio size.
10. Higher Order Temporary Impact Function: For large portfolios it may be more reasonable to assume that the temporary impact cost function has higher-order terms, so that such costs increase *super-linearly* with the trade size. With non-linear impact functions, the general framework used here still applies, but one does not obtain explicit exponential solutions as in the linear impact case.
11. Size Dependent Temporary Impact Parameter: A simple practical solution to this problem is to choose different values of η - the temporary impact parameter – depending up on the overall problem size being considered, recognizing that the model is at best only approximate.

Structure of the Frontier

1. Efficient Frontier and the Corresponding Trajectories: Using a specific choice for the parameters explained below, Almgren and Chriss (2000) produce a sample plot of the efficient frontier – each point on the frontier represents a distinct strategy for optimally liquidating the same basket. Their tangent line represents the optimal solution for a specified risk parameter

$$\lambda = 10^{-6}$$

   They also illustrate the trajectories corresponding to a few sample points on the frontier.
2. Trajectory corresponding to Positive λ: Their first trajectory has

$$\lambda = 2 \times 10^{-6}$$

– this would be chosen by a risk-averse trader who wishes to sell quickly to reduce exposure to volatility risk, despite the trading costs incurred in doing so.

3. Trajectory corresponding to Zero λ: Their second trajectory has

$$\lambda = 0$$

They refer to this as the naïve strategy since this represents an optimal strategy corresponding to simply minimizing expected transaction costs without regard to variance.

4. Linear Reduction of the Holdings: For a security with zero drift and linear transaction costs as defined above

$$\lambda = 0$$

corresponds to a simple linear reduction of holdings over the trading period. Since drift is generally not significant over short trading horizons, the naïve strategy is very close to the linear strategy.

5. Sub Optimality of the Strategy: As Almgren and Chriss (2000) demonstrate later, in a certain sense this is *never* an optimal strategy because one can obtain substantial reductions in variance for a relatively small increase in transaction costs.

6. Trajectory corresponding to Negative λ: Finally their trajectory C has

$$\lambda = -2 \times 10^{-6}$$

it would only be chosen by a trader who likes risk. He postpones execution, thus incurring higher costs both due to rapid sales at the end, and higher variance during the extended period that he holds the security for.

The Utility Function

1. The Risk-Reward Trade-off: Almgren and Chriss (2000) offer an interpretation of the efficient frontier of optimal strategies in terms of the utility function of the seller. They do this in two ways – by direct analogy with modern portfolio theory employing a utility function, and by a novel approach: Value-at-risk. This eventually leads to some general observations regarding the importance of utility in forming execution strategies.
2. Utility of Risk-Averse Functions: Suppose on measure utility by a smooth convex function $u(w)$ where w is the total wealth. This function may be characterized by its risk-aversion coefficient

$$\lambda_u = -\frac{u''(w)}{u'(w)}$$

3. Approximation in Estimating the λ: If the initial portfolio is fully owned, then as the transfer of assets happens from the risky stock into the alternative riskless investment, w remains roughly constant, and one may take λ_u to be a constant throughout the trading period. If the initial portfolio is highly leveraged, then the assumption of constant λ is an approximate one.
4. Formulation of the Optimal Execution Strategy: For short time horizons and small changes in w the higher derivatives of $u(w)$ may be neglected. Thus choosing an optimal execution strategy is equivalent to minimizing the scalar function

$$\mathbb{U}_{UTIL}[x] = \lambda_u \mathbb{V}[x] + \mathbb{E}[x]$$

   The units of λ_u are $\$^{-1}$; one is willing to accept an extra square \$ of variance if it reduces the expected cost by $\$\lambda_u$.
5. Constructing Family of Optimal Paths: The combination $\lambda\mathbb{V}[x] + \mathbb{E}[x]$ is precisely the one used to construct the efficient frontier seen earlier; the parameter λ, introduced as a Lagrange multiplier, has a precise definition as a measure of aversion

to risk. Thus, the methodology above used to construct the efficient frontier likewise produces a family of optimal paths, one for each level of risk aversion.

6. Static Nature of Optimal Path: Returning now to an important point raised earlier, the computation of optimal strategies by minimizing $\lambda\mathbb{V}[x] + \mathbb{E}[x]$ as measured at the initial trading time is equivalent to maximizing the utility at the outset of trading. As one trades, information arrives that could potentially alter the optimal path. The following theorem eliminates that possibility.
7. Time Homogenous Quadratic Utility Theorem: For a fixed quadratic utility function, the static strategies computed above are "time homogenous". More precisely given a strategy that begins at a time

$$t = 0$$

and ends at a time

$$t = T$$

the optimal strategy computed at

$$t = t_k$$

is simply a continuation from

$$t = t_k$$

to

$$t = T$$

of the optimal strategy computed at time

$$t = 0$$

8. <u>Proof Steps: General/Specific Functions</u>: The proof may be seen in two ways – by the algebraic computations based on the specific solutions above, and by general valid for generic non-linear impact functions.
9. <u>Proof Steps: Function Time Shift</u>: First suppose that at time k, where

$$k = 0, \cdots, N - 1$$

one were to compute a new optimal strategy. The new strategy would precisely be

$$x_j = \frac{\sinh\left(\kappa(T - t_j)\right)}{\sinh(\kappa T)} X$$

$$j = 0, \cdots, N$$

with X replaced by x_k, T replaced by $T - t_k$, and t_j replaced by $t_j - t_k$. Using the subscript (k) to denote the strategy computed at time k one would have

$$x_j^{(k)} = \frac{\sinh\left(\kappa(T - t_j)\right)}{\sinh\left(\kappa(T - t_k)\right)} x_k$$

$$j = k, \cdots, N$$

and the trade lists

$$n_j^{(k)} = \frac{2 \sinh\left(\frac{1}{2}\kappa\tau\right)}{\sinh\left(\kappa(T - t_k)\right)} \cosh\left(\kappa\left(T - t_{j-\frac{1}{2}}\right)\right) X$$

$$j = k + 1, \cdots, N$$

10. <u>Proof Step: Recovering Optimal Solutions</u>: It is then apparent that if x_k is the optimal solution from

$$x_j = \frac{\sinh\left(\kappa(T - t_j)\right)}{\sinh(\kappa T)} X$$

$$j = 0, \cdots, N$$

with

$$j \longmapsto k$$

then

$$x_j^{(k)} = x_j^{0}$$

and

$$n_j^{(k)} = n_j^{0}$$

where

$$x_j^{0} = x_j$$

and

$$n_j^{0} = n_j$$

are the strategies from

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X$$

$$j = 0, \cdots, N$$

and

$$n_j = \frac{2\sinh\left(\frac{1}{2}\kappa T\right)}{\sinh(\kappa T)} \cosh\left(\kappa\left(T - t_{j-\frac{1}{2}}\right)\right) X$$

$$j = 1, \cdots, N$$

11. Proof Step: Non-linear Impact: For general non-linear impact functions $g(v)$ and $h(v)$ the optimality condition

$$\frac{x_{j-1} - 2x_j + x_{j+1}}{\tau^2} = \tilde{\kappa}^2 x_j$$

is replaced by a second-order *non-linear* difference relation. The solution ${x_j}^{(k)}$ beginning at a given time is determined by the two boundary values x_k and

$$x_N = 0$$

It is then apparent that the solution does not change if we re-evaluate it at later times.

12. Origin of Time Stable Solutions: More fundamentally, the solutions are time stable because in the absence of serial correlations in the asset price movements, there is no more information about the price changes at later times than there is at the initial time.
13. Optimality over each Sub-interval: Thus, the solution which was determined to be optimal over the entire time interval is optimal as a solution over each sub-interval.

This general phenomenon is well known in the theory of optimal control (Bertsekas (1976)).

Value at Risk

1. Motivation behind Value at Risk: The concept of value at risk is traditionally used to measure the greatest amount of money – maximum profit or loss - a portfolio will sustain over a given period of time under "normal circumstances", where "normal" is defined by a confidence level.
2. Trading Value at Risk Definition: Given a trading strategy

$$x = (x_1, \cdots, x_N)$$

   the value-at-risk of x defined $Var_p[x]$ is defined to be the level of transaction costs by the trading strategy x that will not be exceeded p percent of the time. Put another way, it is the p^{th} percentile level of transaction costs for total costs of trading x.
3. Trading Value at Risk Expression: Under the arithmetic Brownian motion assumption, the total costs – the market value minus capture – are normally distributed with known mean and variance. Thus the confidence level is determined by the number of standard deviations λ_v from the mean of the inverse of the cumulative normal distribution function, and the value at risk for the strategy x is given by

$$Var_p[x] = \lambda_v \sqrt{\mathbb{V}[x]} + \mathbb{E}[x]$$

4. Relation to Implementation Short-fall: That is, with a probability p the trading strategy will not lose more than $Var_p[x]$ of its market value in trading. Borrowing from the language of Period (1988), the implementation shortfall of execution will

not exceed $Var_p[x]$ more than a fraction p of the time. A strategy x is efficient if it has the minimum possible value at risk for the confidence level p.

5. Execution Trajectory Optimized for VaR: Note that $Var_p[x]$ is a complicated non-linear function of x_j composing x; it can be easily evaluated for any given trajectory, but finding the minimizing trajectory directly is difficult.
6. Single Parameter Efficient Frontier Solution: But once the one-parameter family of solutions that form the efficient frontier is obtained, one only needs to solve a one-dimensional problem to find the optimal solutions for the value at risk model, that is, to fund the value of λ_u corresponding to a given value of λ_v. Alternatively on may characterize the solutions by a simple graphical procedure, or may read off the confidence levels corresponding to any particular point on the curve.
7. Almgren-Chriss Optimal VaR Illustration: Almgren and Chriss (2000) produce an illustration of the above, using the square root of variance in the x-axis as opposed to the variance in itself. In this co-ordinate system lines of optimal VaR have a constant slope, and for a given value of λ_v they simply find a tangent to the curve where the slope is λ_v.
8. Interim Optimal Execution Re-evaluation: The question of re-evaluation of the strategy is more complicated and subtle. If one re-evaluates the strategy half-way through the execution process, they will choose a new optimal strategy that is not the same as the original optimal one. The reason is that since λ_v is now held constant, λ_u necessarily changes.
9. General Challenges with the VaR Approach: Value at risk has many flaws from a mathematical point of view, as recognized by Artzner, Delbaen, Eber, and heath (1997). The particular issue encountered her would occur in any problem in which the time of measurement is a fixed date, rather than maintained at a fixed distance in the future. It is an open issue to formulate suitable measures of risk for general time-dependent problems.
10. Liquidity Adjusted Value at Risk: Despite this shortcoming, Almgren and Chriss (2000) the smallest possible value of $Var_p[x]$ as an informative measure of the possible loss associated with the initial position, in the presence of liquidity effects. This value, which they call L-VaR for Liquidity Adjusted Value at Risk, depends on

the time to liquidation and the confidence level chosen, in addition to the market parameters such as the impact coefficient (Almgren and Chriss (1999)).

11. Advantages of the L-VaR Approach: The optimal trajectories determined by minimizing the value at risk do *not* have the counter-intuitive scaling behavior seen earlier; even for linear impact functions, large portfolios will be traded closer to the straight line trajectory.
12. Using L-VaR for Large Portfolios: This is because the cost assigned to uncertainty scales *linearly* with the portfolio size, while the temporary impact cost scales *quadratically* as before. Thus the latter is more important for large portfolios.

The Role of Utility in Execution

1. General Observations on Optimal Execution: Almgren and Chriss (2000) use the structure of the efficient frontier in the framework that they have developed to make some general observations concerning optimal executions.
2. The Naïve Strategy Benchmark: They first restrict themselves to the situation where the trader has no directional view on the security being traded. Recall that in this case, the naïve strategy is the simple straight line strategy in which the trader breaks the blocks being executed into equal sized blocks to be sold over equal time intervals. They use this strategy as a benchmark for comparison with the other strategies used throughout here.
3. Convex $\mathbb{E}[x]$ to $\mathbb{V}[x]$ Mapping: A crucial insight is that the curve defining the efficient frontier is a smooth convex function $\mathbb{E}[\mathbb{V}]$ mapping the levels of variance $\mathbb{V}$ to the corresponding minimum mean transaction cost levels.
4. Region around the Naïve Strategy: Write $(\mathbb{E}_0, \mathbb{V}_0)$ for the mean and variance around the naïve strategy. Regarding $(\mathbb{E}_0, \mathbb{V}_0)$ as a point on the smooth curve $\mathbb{E}[\mathbb{V}]$ defined by the frontier, $\frac{\partial \mathbb{E}}{\partial \mathbb{V}}$ evaluated at $(\mathbb{E}_0, \mathbb{V}_0)$ is equal to zero. Thus for $(\mathbb{E}, \mathbb{V})$ near $(\mathbb{E}_0, \mathbb{V}_0)$ one has

$$\mathbb{E} - \mathbb{E}_0 = \frac{1}{2}(\mathbb{V} - \mathbb{V}_0)^2 \left.\frac{\partial^2 \mathbb{E}}{\partial \mathbb{V}^2}\right|_{\mathbb{V}=\mathbb{V}_0}$$

where

$$\left.\frac{\partial^2 \mathbb{E}}{\partial \mathbb{V}^2}\right|_{\mathbb{V}=\mathbb{V}_0}$$

is positive is positive by the convexity of the frontier at the naïve strategy.

5. Special Feature of the Naïve Strategy: By definition, the naïve strategy has the property that any strategy with lower variance in cost has a greater expected cost. However a special feature of the naïve strategy is that a first-order decrease in variance can be obtained – in the sense of finding a strategy with a lower variance – while only incurring a second order increase in cost.
6. Disadvantages of Risk Neutral Strategy: From the above it follows that for small increases in variance, one can obtain much larger reductions in cost. Thus unless the trader is risk-neutral it is always advantageous to execute a strategy that is at least to some degree "to the left" of the naïve strategy. Thus one concludes that, in this framework, from a theoretical standpoint, it never makes sense to trade a strictly risk-neutral strategy.
7. The Role of a Security's Liquidity: An intuitive proposition is that with all things being equal, a trader will execute a more liquid basket more rapidly than a less liquid one. In the extreme this is particularly clear. A broker given a small order to execute over the course of the day will execute the entire order almost immediately.
8. Executing the Highly Liquid Security: How does one explain this? The answer is that the market impact cost attributable to rapid trading is negligible compared with the opportunity cost incurred in breaking up the order over an entire day. Thus, even if the expected return on a security over the day is zero, the perception is that the risk of waiting is overweighed by any small cost of immediacy.
9. Absence of Risk Reduction Premium: Now if the trader were truly risk neutral, in the absence of any views, he would always use the naïve strategy and employ the allotted

time fully. This would make sense because any price to pay for trading immediately is worthless if one places no premium on risk reduction.

10. Limitation of Risk Neutral Approach: It follows that any model that proposes optimal trading behavior should predict that more liquid baskets are traded more rapidly than less liquid ones. A model that only considers the minimization of transaction costs, like that of Bertsimas and Lo (1998), is essentially a model that excludes utility.
11. Optimal Execution Independent of Liquidity: In such a model, and under Almgren and Chriss (2000) basic assumptions, traders will trade all baskets at the same rate irrespective of the liquidity, that is unless they have an explicit directional view on the security, or the security possesses extreme serial correlation in its price movements.
12. Super Linear Market Impact Functions: Almgren and Chriss (2000) do note that their model in the case of linear transaction costs does not predict a more rapid trading for smaller versus larger baskets of the same security. However, this is a consequence of choosing linear temporary impact functions and the problem goes away when one considers more realistic super-linear functions.
13. Risk Neutral Execution Half Life: Another way of looking at this is that the half-life of all black executions, under the assumption of risk-neutral preferences, is infinite.

Choice of Parameters

1. The Asset Intrinsic Dynamics Parameters: Almgren and Chriss (2000) compute some numerical examples for the purposes of exploring the qualitative properties of the efficient frontier. Throughout the examples they consider a single stock with the current market price of

$$S_0 = 50$$

and that they initially have one million shares, for an initial portfolio size of \$50 million. The stock will have 30% annual volatility, 10% expected annual rate of return, a bid-ask spread of $\frac{1}{8}$, and a median daily trading volume of 5 million shares.

2. Stock Asset Daily Return/Volatility: With a trading year of 250 days this gives a daily volatility of

$$\frac{0.3}{250} = 0.019$$

and expected fractional return of

$$\frac{0.1}{250} = 4 \times 10^{-4}$$

To obtain our absolute parameters σ and α one must scale it by the price, so

$$\sigma = 0.019 \times 50 = 0.95$$

and

$$\alpha = (4 \times 10^{-4}) \times 50 = 0.02$$

The table below summarizes the information.

3. Parameter Values for the Test Case:

Parameter Description	**Parameter Symbol**	**Parameter Value**
Initial Stock Price	S_0	\$50/share
Initial Holdings	X	10^6 shares
Liquidation Time	T	5 days
Number of Time Periods	N	5

30% Annual Volatility	σ	$0.95\ (\$/share)/day^{\frac{1}{2}}$
10% Annual Growth	α	$0.02\ (\$/share)/day$
Bid Ask Spread $\frac{1}{8}$	ϵ	$\$0.0625/share$
Daily Volume 5 million shares	γ	$2.5 \times 10^{-7}\$/share^2$
Impact at 1% of market	η	$2.5 \times 10^{-6}(\$/share)/(share/day)$
Static Holdings 11,000 shares	λ_u	$10^{-6}/\$$
VaR Confidence $p = 95\%$	λ_v	1.645

4. Incremental and Total Execution Times: Suppose that one wants to liquidate this position in one week so that

$$T = 5\ days$$

This is divided into daily trades such that τ is 1 day and

$$N = 5$$

5. Standard Deviation of the Trajectory: Over this period, if one holds the original position with no trading, the fluctuations in the stock value will be Gaussian with a standard deviation of

$$\sigma\sqrt{T} = 2.12\ (\$/share)$$

and the fluctuations in this value will have an absolute standard deviation of

$$\sqrt{V} = \$2.12M$$

As expected this is precisely the value of $\sqrt{V}$ for the lowest point in the efficient frontier, since that point corresponds selling along a linear trajectory rather than holding a constant amount.

6. Temporary Cost Function Parameter - ϵ: One then chooses the parameters for the temporary cost function

$$h\left(\frac{n_k}{\tau}\right) = \epsilon sgn(n_k) + \frac{\eta}{\tau} n_k$$

Almgren and Chriss (2000) set

$$\epsilon = \frac{1}{16}$$

that is, the fixed part of the temporary costs will be one-half the bid-ask spread.

7. Temporary Cost Function Parameter - η: For η they suppose that for each 1% of the daily volume traded they incur a price impact equal to one bid-ask spread. For example trading at a rate of 5% daily volume incurs a one-time cost on each trade of $\frac{5}{8}$. Under this assumption

$$\eta = \frac{\frac{1}{8}}{0.01 \times 5 \times 10^6} = 2.5 \times 10^{-6}$$

8. Permanent Cost Function Parameter - γ: For permanent costs, the common rule of thumb is that price effects become significant when 10% of the daily volume is sold. Assuming that "significant" means that the price depression is one bid-ask spread, and that the effect is linear for both smaller and larger trading rates, one has

$$\gamma = \frac{\frac{1}{8}}{0.1 \times 5 \times 10^6} = 2.5 \times 10^{-7}$$

Recall that this parameter gives a fixed cost independent of the path.

9. The Risk Aversion Parameter - λ: Almgren and Chriss (2000) have chosen

$$\lambda = \lambda_u = 10^{-6}$$

For these parameters, from

$$\kappa \sim \tilde{\kappa} + \mathcal{O}(\tau^2) \sim \sqrt{\frac{\lambda \sigma^2}{\eta \left(1 - \frac{\gamma \tau}{2\eta}\right)}} + \mathcal{O}(\tau)$$

$$\tau \to 0$$

one has for the optimal strategy that

$$\kappa \approx 0.61\ day$$

so that

$$\kappa T \approx 3$$

Since this value is near 1 in magnitude, the behavior is an interesting intermediate in-between the naïve extremes.

10. λ_v at 95% Confidence Level: For the value at risk representation, as assumed 95% confidence level gives

$$\lambda_v = 1.645$$

The Value of Information

1. Zero Drift Random Walk Assumption: The discussion carried out so far assumed that the price dynamics followed an arithmetic random walk with zero drift. Since past price paths provide no extra information on future price movements, the conclusion was that the optimal trajectories can be statically determined. There are three ways by which a random walk with zero drift may fail to represent the price process.
2. Non-zero Drift in Dynamics: First the price process may have drift. For example, if the trader has a strong directional view, the trader may want to incorporate this view into the liquidation strategy.
3. Cross Period Serial Correlation Impact: Second, the price process may exhibit serial correlation. The presence of first order serial correlation for example, implies that the price moves in a given period provide non-trivial information concerning the next period movement of the asset.
4. Incorporation of the Investor’s Private Information: Bertsimas and Lo (1998) study a general form of this assumption, wherein an investor possesses possibly private information of a serially correlated information vector that acts as a linear factor in the asset returns.
5. Exogenously Induced Material Parameter Shift: Lastly, at the start of trading, it may be known that at some specific point in time, an event will take place whose outcome will cause a material shift in the parameters governing the price process.
6. Literature Survey on Exogenous Events: Such event induced parameter shifts include quarterly and annual earnings announcements, dividend announcements, and share re-purchases. Event studies documenting these parameter shifts and providing theoretical grounding for their existence include Beaver (1968), Fama, Fisher, Jensen, and Roll (1969), Dann (1981), Patell, and Wolfson (1984), Kalay and Loewentstein (1985), Kim and Verrecchia (1991), Campbell, Lo, and MacKinlay (1997), Easterwood and Nutt (1999), and Ramaswami (1999).
7. Temporary Shifts on Dynamic Parameters: For example, Brown, Harlow, and Tinic (1988) show that events cause temporary shifts in both the risk and returns of individual securities, and the extent of these shifts depends on the outcome of the event. In general, securities react more strongly to bad news than good news.

8. Probabilistic Event Outcomes/Parameter Shifts: Almgren and Chriss (2000) study a stylized version of the events in which a known event at a known time – e.g., an earnings announcement – has several possible outcomes. The probability of each outcome is known, and the impact that a given outcome will have on the parameters of the price is also known. Clearly, optimal strategies must explicitly use this information, and Almgren and Chriss (2000) develop methods to incorporate event-specific information into their risk-reward framework.
9. Back-to-Back Static Strategies: The upshot is a piece-wise strategy that trades statically up to the event, and then reacts explicitly to the outcome of the event. Thus the burden is on the trader to determine which of the possible outcomes occurred and then trade accordingly.

Drift

1. Drift as a Directional View: It is convenient to regard the drift parameter in the price process as a directional view of price movements. For example, the trader charged with liquidating a single security may believe that this security is likely to rise. Intuitively it makes more sense to trade this issue more slowly to take advantage of this view.
2. Incorporating Drift into Price Dynamics: To incorporate drift into the price dynamics Almgren and Chriss (2000) modify

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k - \gamma g\left(\frac{n_k}{\tau}\right)$$

to

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k + \alpha\tau - \gamma g\left(\frac{n_k}{\tau}\right)$$

where α is an expected drift term. If the trading proceeds are invested in an interest bearing account, then α should be taken as the *excess* rate of return of the risky asset.

3. Price Expectation over Time Period: One can readily write the modified version of

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} {n_k}^2$$

as

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 - \alpha \sum_{k=1}^{N} \tau x_k + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} {n_k}^2$$

4. Updated Objective Function Optimality Condition: The variance is still given by

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

The optimality condition

$$\frac{x_{j-1} - 2x_j + x_{j+1}}{\tau^2} = \tilde{\kappa}^2 x_j$$

becomes

$$\frac{x_{j-1} - 2x_j + x_{j+1}}{\tau^2} = \tilde{\kappa}^2 \left(x_j - \bar{x}\right)$$

in which the new parameter

$$\bar{x} = \frac{\alpha}{2\lambda\sigma^2}$$

is the optimal level of security holding for a time independent portfolio optimization problem.

5. Drift Based Updated Execution Slice: For example, the parameters used in the example above give approximately

$$\bar{x} = 1{,}100\ shares$$

or 0.11% of our initial portfolio. One expects this fraction to be very small, since, by hypothesis, the eventual aim is complete liquidation.

6. Drift Based Updated Optimal Solution: The optimal solution

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X$$

$$j = 0, \cdots, N$$

becomes

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X + \left\{1 - \frac{\sinh\left(\kappa\left(T - t_j\right)\right) + \sinh\left(\kappa t_j\right)}{\sinh(\kappa T)}\right\} \bar{x}$$

for

$$j = 0, \cdots, N$$

with the associated trades

$$n_j = \frac{2\sinh\left(\frac{1}{2}\kappa\tau\right)}{\sinh(\kappa T)}\cosh\left(\kappa\left(T - t_{j-\frac{1}{2}}\right)\right)X + \frac{2\sinh\left(\frac{1}{2}\kappa\tau\right)}{\sinh(\kappa T)}\left[\cosh\left(\kappa t_{j-\frac{1}{2}}\right) - \cosh\left(\kappa\left(T - t_{j-\frac{1}{2}}\right)\right)\right]\bar{x}$$

7. Initial Position Independent Trajectory Correction: This trading trajectory is a sum of two distinct trajectories – the zero-drift solution as computed before, plus a "correction" which profits by capturing a piece of the predictable drift component. The size of this correction term is proportional to $\bar{x}$, and thus to α: it is independent of the initial portfolio size X.
8. Practical Incorporation into Program Trading: To place this in an institutional framework, consider a program trading desk that sits in front of customer flow. If this desk were to explicitly generate alphas on all securities that flow through the desk in an attempt to, say, hold securities with high alphas and sell securities with low alphas more rapidly, the profit would not scale in proportion to the average size of the programs. Rather it would only scale with the number of securities that flow through the desk. An even stronger conclusion is that since the optimal strategy disconnects into a static strategy unrelated to the drift term, and a second strategy related to the drift term, there is no particular advantage to restricting trading in securities which the desk currently holds the positions in.
9. Comparison: Highly Liquid Markets Scenario: The difference between this solution and the no-drift solution in

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)}X$$

$$j = 0, \cdots, N$$

may be understood by considering the case

$$\kappa T \gg 1$$

corresponding to highly liquid markets. Whereas the previous one relaxed from X to $\frac{X}{e}$ in a time scale of

$$\theta = \frac{1}{\kappa}$$

this one relaxes instead to the optimal static portfolio size $\bar{x}$. Near the end of the trading period the trader sells the remaining holdings to achieve

$$x_N = 0$$

at

$$t = T$$

10. <u>Caveat: Buy-Sell Symmetry Breaking</u>: In this case, one requires

$$0 \leq \bar{x} \leq X$$

in order for all trades to be in the same direction. This breaks the symmetry between a buy program and a sell program, if one wanted to consider buy programs it would be more logical to set

$$\alpha = 0$$

Gain due to Drift

1. Gain from Drift – Calculation Motivation: Now suppose that the price dynamics is given by

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k + \alpha\tau - \gamma g\left(\frac{n_k}{\tau}\right)$$

with

$$\alpha > 0$$

but one chooses to determine the solution as though

$$\alpha = 0$$

The situation may arise, for example, in case where the trader is trading a security with non-zero drift, but *unknowingly* assumes that the security has no drift. Almgren and Chriss (2000) explicitly calculate the loss associated with ignoring the drift term.

2. Gain adjusted $\mathbb{E}[x]$ and $\mathbb{V}[x]$: Write ${x_j}^*$ for the optimal solution

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X + \left\{1 - \frac{\sinh\left(\kappa\left(T - t_j\right)\right) + \sinh\left(\kappa t_j\right)}{\sinh(\kappa T)}\right\}\bar{x}$$

with

$$\alpha > 0$$

${x_j}^0$ for the sub-optimal solution

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X$$

$$j = 0, \cdots, N$$

or

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X + \left\{1 - \frac{\sinh\left(\kappa\left(T - t_j\right)\right) + \sinh\left(\kappa t_j\right)}{\sinh(\kappa T)}\right\} \bar{x}$$

with

$$\alpha = 0$$

Also write $\mathbb{E}^*[x]$ and $\mathbb{V}^*[x]$ for the optimal expected cost and its variance measured by

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 - \alpha \sum_{k=1}^{N} \tau x_k + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} {n_k}^2$$

and

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

with

$$x_j = {x_j}^*$$

and write $\mathbb{E}^0[x]$ and $\mathbb{V}^0[x]$ for the sub-optimal values of

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 - \alpha \sum_{k=1}^{N} \tau x_k + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} {n_k}^2$$

and

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

evaluated with

$$x_j = {x_j}^0$$

3. <u>Objective Function Gain from Drift</u>: The corresponding objective functions are

$$\mathbb{U}^*[X] = \mathbb{E}^*[X] + \lambda \mathbb{V}^*[X]$$

and

$$\mathbb{U}^0[X] = \mathbb{E}^0[X] + \lambda \mathbb{V}^0[X]$$

One can then define the *gain due to drift* to be the difference $\mathbb{U}^0[X] - \mathbb{U}^*[X]$; this is the reduction in the cost and the variance by being aware of and taking into account of the drift term. Clearly

$$\mathbb{U}^0[X] - \mathbb{U}^*[X] \geq 0$$

since x^* is the unique optimal strategy for the model with

$$\alpha > 0$$

4. <u>Upper Bound for the Gain</u>: Now the value of the terms in $\mathbb{U}^0[X]$ that come from

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 - \alpha \sum_{k=1}^{N} \tau x_k + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} {n_k}^2$$

and

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau {x_k}^2$$

is only *increased* by going from x^0 to x^* since x^0 and not x^* was the optimum strategy with

$$\alpha = 0$$

Therefore an *upper bound* for the gain is

$$\mathbb{U}^0[X] - \mathbb{U}^*[X] \geq \alpha\tau \sum_{k=1}^{N} ({x_k}^* - {x_k}^0)$$

5. <u>Adjustment Applied to the Holdings</u>: That is, in response to positive drift, one should increase the holdings throughout the trading. This reduces the net cost by the amount of the increase in the asset price one captures, at the expense of slightly increasing the transaction costs and the volatility exposure. An upper bound for the possible benefit is the amount of increase one captures.
6. <u>Explicit Expression for the Bound</u>: But ${x_k}^* - {x_k}^0$ is just the term in the square brackets in

$$x_j = \frac{\sinh\left(\kappa\left(T - t_j\right)\right)}{\sinh(\kappa T)} X + \left\{1 - \frac{\sinh\left(\kappa\left(T - t_j\right)\right) + \sinh\left(\kappa t_j\right)}{\sinh(\kappa T)}\right\} \bar{x}$$

times $\bar{x}$, which is clearly independent of X. Indeed this can be explicitly evaluated to get

$$\alpha\tau \sum_{k=1}^{N} ({x_k}^* - {x_k}^0) = \alpha\bar{x}T \left[1 - \frac{\tau}{T} \frac{\tanh\left(\frac{1}{2}\kappa T\right)}{\tanh\left(\frac{1}{2}\kappa\tau\right)}\right]$$

7. Gain Comparison against Execution Cost: Since $\frac{\tanh x}{x}$ is a positive decreasingly function, this quantity is positive and bounded above by $\alpha\bar{x}T$, the amount one would gain by holding $\bar{x}$ for a time T. Any reasonable estimates for the parameters show that this quantity is negligible compared to the impact costs incurred in liquidating an institutional sized portfolio over a short period.

Serial Correlation

1. Prior Period Price Increment Component: Now one supposes that the asset prices exhibit serial correlation, so that at each period one discovers a component of predictability of the asset price in the next period.
2. Methodology behind the Price Increment Estimation: In the model

$$S_k = S_{k-1} + \sigma\sqrt{\tau}\xi_k - \gamma g\left(\frac{n_k}{\tau}\right)$$

with a drift

$$\alpha = 0$$

one now supposes that the ξ_k are serially correlated with period-to-period correlation ρ

$$|\rho| < 1$$

One can determine ξ_k at time k based on the obtained $S_k - S_{k-1}$ and sale n_k

3. Optimal Strategy no more Static: With serial correlation the optimal trajectory is no longer a static trajectory determined in advance of trading; since each price movement gives some information about the immediate future price movements, the optimal trade list can be determined only one period at a time.
4. Estimation of the Realized Gain: Thus a full optimal solution requires the use of dynamic programming methods. However since the information is still roughly local in time, one can estimate the optimal gain attainable by an optimal strategy.
5. Almgren and Chriss (2000) Conclusions: Almgren and Chriss (2000) state their conclusion in advance of their estimation. The value of information contained in pure movements due to serial correlations is independent of the size of the portfolios being traded. The calculation demonstrated below lends intuition to this counter-intuitive statement.
6. Per Period Price Change Impact: Consider two consecutive periods during which the base strategy has the trader trading the same number of shares n in each period. With a linear impact price model, in each period price changes by $\left[\epsilon + \eta \frac{n}{\tau}\right]$ dollars/share. The trader pays this cost in each of the n shares, so the total cost due of market impact per period is $\left[\epsilon + \eta \frac{n}{\tau}\right] n$
7. Price Change from Serial Correlation: Suppose on has some price information due to serial correlations. If one knows ξ_k at the previous period, then the predictable component of the price change is roughly $\rho\sigma\sqrt{\tau}\Delta n$.
8. Incremental Cost of the Adapted Strategy: But this adaptation increases the impact costs. After the shift in the first period the price change is $\epsilon + \eta \frac{n-\Delta n}{\tau}$ while in the second period $\epsilon + \eta \frac{n+\Delta n}{\tau}$. These costs are paid on $n - \Delta n$ and $n + \Delta n$ shares respectively, so the market impact per period is now

$$\left[\frac{1}{2}\left(\epsilon + \eta \frac{n - \Delta n}{\tau}\right)(n - \Delta n) + \frac{1}{2}\left(\epsilon + \eta \frac{n + \Delta n}{\tau}\right)(n + \Delta n)\right] = \left[\epsilon + \eta \frac{n}{\tau}\right] n + \frac{n}{\tau} \Delta n^2$$

9. Optimal Per-Period Execution Shift: To determine has many shares one should shift, one solves the quadratic optimization problem

$$\max_{\Delta n} \left[\rho\sigma\sqrt{\tau}\Delta n - \frac{n}{\tau}\Delta n^2\right]$$

The optimal Δn is readily found as

$$\Delta n^* = \frac{\rho\sigma\tau^{\frac{3}{2}}}{2\eta}$$

and the maximum possible gain per period is $\frac{\rho^2\sigma^2\tau^2}{4\eta}$. This heuristic can be confirmed by a detailed dynamic programming computation that accounts for optimal shifts across multiple periods.

10. Limitation of Optimal Gain Execution: Almgren and Chriss (2000) also explain briefly the limitation of the above approximation. When ρ is close to zero, clearly this approximation is extremely close to accurate, because the persistence of the serial correlation effect dies down very quickly after the first period. When $|\rho|$ is too large to ignore, the approximation is too small for

$$\rho > 0$$

That is, $\frac{\rho^2\sigma^2\tau^2}{4\eta}$ understates the possible gains over ignoring serial correlation. Conversely when

$$\rho < 0$$

$\frac{\rho^2\sigma^2\tau^2}{4\eta}$ overstates the possible gains due to serial correlation. As

$$\rho > 0$$

is more frequently the case Almgren and Chriss (2000) assert that $\frac{\rho^2\sigma^2\tau^2}{4\eta}$ is useful for bounding the possible gains in most situations available from serial correlations.

11. Position Independence of Gain/Cost: Note that both the size of the adaptation, and the resulting gain, are independent of the amount of shares n that would be sold under an unadapted strategy. That is they are also independent of the size of the initial portfolio.
12. Gain/Cost Liquidity/Correlation Dependence: Instead the binding constraint is the liquidity of the security being traded, and the magnitude of the correlation coefficient. The more information available due to correlation and the more liquid the security, the more overall gain that is available due to adapting the strategy to the correlations.
13. Higher Order Impact Function Optimality: The results above are especially simple because of the assumption of linear impact functions. Almgren and Chriss (2000) also show briefly what happens in the more general case of nonlinear market impact functions

$$h(v) = h\left(\frac{n}{\tau}\right)$$

The cost per period due to market impact is

$$\left[\frac{1}{2}h\left(\frac{n-\Delta n}{\tau}\right)(n-\Delta n) + \frac{1}{2}h\left(\frac{n+\Delta n}{\tau}\right)(n+\Delta n)\right]$$
$$\approx\ h\left(\frac{n}{\tau}\right)n + \left[\frac{1}{2}h''\left(\frac{n}{\tau}\right)\frac{n}{\tau} + h'\left(\frac{n}{\tau}\right)\right]\frac{\Delta n^2}{\tau}$$

for small Δn. Now the optimal shift and the maximal gain are given by

$$\Delta n^* = \frac{\rho\sigma\tau^{\frac{3}{2}}}{vh'' + 2h'}$$

and $\frac{\rho^2\sigma^2\tau^2}{2(vh''+2h')}$ respectively, where h' and h'' are evaluated at the base execution rate of

$$v = \frac{n}{\tau}$$

The linear case is recovered by setting

$$h(v) = \epsilon + \eta v$$

This has the special property that h' is independent of v and

$$h'' = 0$$

14. <u>Optimality Dependence on Impact Exponent</u>: In general suppose

$$h(v) \sim \mathcal{O}(v^{\varpi})$$

as

$$v \to \infty$$

$$\varpi > 0$$

is required so that $h(v)$ is increasing; selling the share always pushes the price down more. The marginal cost is

$$h'(v) \sim \mathcal{O}(v^{\varpi-1})$$

$$\varpi > 1$$

corresponds to an increasing marginal impact, and

$$\varpi < 1$$

corresponds to a decreasing marginal impact. Then the per-period cost one pays on the base strategy is

$$\sim \mathcal{O}(v^{\varpi+1})$$

for large initial portfolios, and hence large rates of execution. The marginal gain from adapting to evolution is

$$\sim \mathcal{O}(v^{\varpi-1})$$

in the same limit.

Parameter Shifts

1. Price Dynamics Parameter Set Shift: Almgren and Chriss (2000) discuss the impact on optimal execution of scheduled news earnings such as earnings and dividend announcements. Such events have two features that make them an important object of study. First the outcome of the event determines the shift in the parameters governing the price dynamics – see Brown, Harlow, and Tinic (1988), Easterwood and Nutt (1999), and Ramaswami (1999).
2. Determining an Event' Full Impact: Second, the fact that they are scheduled increases the likelihood that one can detect what the true outcome of the event is. This situation is formalized below, and explicit formulas are given for price trajectories before and after the event takes place.

3. Scheduled Event Occurrence Time T_*: Suppose at some time T_* between now and the specified final time T an event will occur, the outcome of which may or may not cause a shift in the parameters of price dynamics.
4. New Regime Shifted Parameter Set: The term *regime set* or *parameter set* refers to the collection

$$R = \{\sigma, \eta, \cdots\}$$

of the parameters that govern the dynamics at any particular time, and the events of interest are those that have the possibility of causing *parameter shifts*.
5. Initial to Final Regime Shift: Let

$$R_0 = \{\sigma_0, \eta_0, \cdots\}$$

be the parameters of price dynamics at the time the execution begins. Suppose the market can shift to one of possible new sets of parameters p so that $R_1, \cdots, R_p$ is characterized by parameters $\sigma_j, \eta_j, \cdots$ for

$$j = 1, \cdots, p$$

6. Probability of a Regime Switch: One also supposes that probabilities can be assigned to these possible new states, so that p_j is the probability that regime R_j occurs. The probabilities are *independent* of the short term market fluctuations represented by ξ_k. Of course it is possible that some R_j has the same values as R_0 in which case p_j is the probability that no change occurs.
7. Globally Optimal Dynamic Trading Strategy: Almgren and Chriss (2000) consider a dynamic trading strategy the yields globally optimal strategies in the presence of a parameter shift at time T_*. Taking

$$T_* = t_s = s\tau$$

one pre-computes an initial trajectory

$$x^0 = \{{x_0}^0, \cdots, {x_s}^0\}$$

with

$${x_0}^0 = X$$

Denote

$$X_* = {x_s}^0$$

8. Landscape of Switchable Trajectories: They also compute a family of trajectories

$$x^j = \left\{{x_0}^j, \cdots, {x_s}^j\right\}$$

for

$$j = 1, \cdots, p$$

all of which have

$${x_s}^j = X_*$$

and

$${x_N}^j = 0$$

They follow the trajectory x^0 until the time of the shift. Once the shift occurs they assume they can quickly identify the outcome of the event and the new set of parameters governing the price dynamics. With this settled, the complete trading using the corresponding trajectory x^j is determined.

9. Key Almgren and Chriss (2000) Results: Almgren and Chriss (2000) show that it is possible to determine each trajectory using static optimization; although one cannot choose which one to use until the event occurs. Also the starting trajectory x^0 will *not be the same* as the trajectory one would use if they believed the regime R_0 would hold through the entire time T.
10. Trajectory Conditional on Fixed X_*: To determine the trajectories $x^0, x^1, \cdots, x^p$ they reason as follows. Suppose that the common value of

$$X_* = x_s{}^0 = x_s{}^j$$

is fixed. Then by virtue of the independence of the regime shift in itself from the security motions, the optimal trajectories conditional on the values of X_* are simply those that have already been computed with a small modification to include the given non-zero final value.
11. Sequential Pair of Static Strategies: One can immediately write

$$x_k{}^0 = \frac{\sinh\left(\kappa_0(T_* - t_k)\right)}{\sinh(\kappa_0 T_*)} X + \frac{\sinh(\kappa_0 t_k)}{\sinh(\kappa_0 T_*)} X_*$$

$$k = 0, \cdots, s$$

where κ_0 is determined from $\sigma_0, \eta_0, \cdots$. The trajectory is determined the same way as seen before; it is the unique combination of the exponentials $x^{\pm\kappa_0 t}$ that has

$$x_0{}^0 = X$$

and

$$x_s{}^0 = X_*$$

Similarly

$$x_k{}^j = \frac{\sinh\left(\kappa_j(T - t_k)\right)}{\sinh\left(\kappa_j(T - T_*)\right)} X_*$$

$$k = s, \cdots, N$$

$$j = 1, \cdots, p$$

Thus one only needs to determine X_*.

12. Principle behind the Estimation of X_*: To determine X_* one needs to determine the expected loss and the variance of the combined strategy. Let $\mathbb{E}_0$ and $\mathbb{V}_0$ denote the expectation and the loss incurred by the trajectory x^0 on the first segment

$$k = 0, \cdots, s$$

The quantities can be determined readily using

$$\mathbb{E}[X] = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^{N} |n_k| + \frac{\tilde{\eta}}{\tau} \sum_{k=1}^{N} n_k{}^2$$

and

$$\mathbb{V}[X] = \sigma^2 \sum_{k=1}^{N} \tau x_k{}^2$$

13. Mean Variance of the Compound Strategy: Then by virtue of the regime shift and the security motion's independence, the expected loss of the compound strategy is

$$\mathbb{E} = \mathbb{E}_0 + \mathbb{P}_1\mathbb{E}_1 + \cdots + \mathbb{P}_p\mathbb{E}_P$$

and its variance is

$$\mathbb{V} = \mathbb{V}_0 + \mathbb{P}_1\mathbb{V}_1 + \cdots + \mathbb{P}_p\mathbb{V}_P + \frac{1}{2}\sum_{i,j=1}^{p0} \mathbb{P}_i\mathbb{P}_j\left(\mathbb{E}_i - \mathbb{E}_j\right)^2$$

One can now do a one-variable optimization in X_* to maximize $\mathbb{E} + \lambda\mathbb{V}$. Almgren and Chriss (2000) provide a pictorial representation of the above.

Conclusions and Further Extensions

1. Efficient Frontier of Transaction Costs: The central feature of the Almgren and Chriss (2000) analysis has been to construct an *efficient frontier* in a two-dimensional plane whose axes are the expectation of the total cost and its variance.
2. Linear Impact Functions Analytical Solutions: Regardless of an individual's tolerance to risk, the only strategies which are candidates for being optimal solutions are found in this one-parameter set. For linear impact functions, they give complete analytical solutions for the strategies in this set.
3. Efficient Frontier Optimal Operating Characteristic: Then, considering the details of risk aversion, they have shown how to select an optimal point on the frontier either by classical mean-variance optimization, or by the concept of value at risk. These solutions are easily constructed numerically, and interpreted graphically by examining the frontier.
4. First Conclusion: Sub-optimal Strategies: Because the set of attainable strategies, and hence the efficient frontier, are generally *smooth* and *convex,* a trader who is at all risk-averse should never trade according to the naïve strategy of minimizing expected cost. This is because in the neighborhood of that strategy, the first order reduction in

the variance is attained at the expense of only a second order increase in the expected cost.

5. Second Conclusion: Custom Risk Optimization: Almgren and Chriss (2000) also observe that this careful analysis of the costs and risks of liquidation can be used to give a more precise characterization of the risk of holding the initial portfolio. As an example, they define a Liquidity-Adjusted VaR (L-VaR) to be, for as given time horizon, the minimum VaR of any static liquidation strategy.
6. Actual Gains of Dynamic Trading: Although it may seem counter-intuitive that the optimal strategies can be determined in advance of trading, Almgren and Chriss (2000) argue that only very small gains can be realized by adapting the strategy to the information as it is needed.
7. First Extension: Continuous Time Trading: The limit

$$\tau \to 0$$

is immediate in all of their solutions. Their trading strategy is characterized by a holdings function $x(t)$ and a *trading rate*

$$x(t) = \underset{\tau \to 0}{Limit} \frac{n_k}{\tau}$$

Almgren and Chriss (2000) minimum variance strategy has infinite cost, but the optimal strategies for finite λ have finite cost and variance. However, this limit is at best a mathematical convenience, as the market model is implicitly a "coarse-grained" description of the real dynamics.
8. Second Extension: Nonlinear Cost: The conceptual framework outlined by Almgren and Chriss (2000) is not limited to the linear permanent and temporary impact functions

$$g(v) = \gamma v$$

and

$$h(v) = \epsilon sgn(n_k) + \frac{\eta}{\tau} n_k$$

though the exact exponential/hyperbolic solutions are specific to that case. For nonlinear functions $g(v)$ and $h(v)$ that satisfy suitable convexity conditions, optimal risk-averse trajectories are found by solving a non-quadratic optimization problem; the difficulty of the problem depends on the specific functional forms chosen.

9. Third Extension Time Varying Coefficients: Almgren and Chriss (2000) framework also covers the case in which the volatility, the market impact parameters, and perhaps the expected drift are all time-dependent; finding the optimal strategy entails solving a linear system of size equal to the number of time periods (times the number of assets, for a portfolio problem). One example in which this is useful is if the price is expected to jump up or down on a known future date – say, an earnings announcement – as long as one has a good estimate of the expected *size* of this jump.

Numerical Optimal Trajectory Generation

1. Varying Time Interval Cost Distribution:

$$\mathbb{E}[x] = \sum_{k=1}^{N} \tau_k x_k g\left(\frac{n_k}{\tau_k}\right) + \sum_{k=1}^{N} n_k h\left(\frac{n_k}{\tau_k}\right)$$

$$\mathbb{V}[x] = \sum_{k=1}^{N} \tau_k {\sigma_k}^2 {x_k}^2$$

$$\tau_k = t_k - t_{k-1}$$

$$n_k = x_k - x_{k-1}$$

2. Time Varying Interval Linear Impact:

$$\mathbb{E}[x] = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^{N} |n_k| + \eta \sum_{k=1}^{N} \frac{{n_k}^2}{\tau_k} - \frac{1}{2}\gamma \sum_{k=1}^{N} {n_k}^2$$

$$\mathbb{V}[x] = \sum_{k=1}^{N} \tau_k {\sigma_k}^2 {x_k}^2$$

$$\tau_k = t_k - t_{k-1}$$

$$n_k = x_k - x_{k-1}$$

3. Varying Interval Linear Impact Objective:

$$\mathbb{U}[x] = \mathbb{E}[x] + \lambda\, \mathbb{V}[x]$$

implies

$$\mathbb{U}[x] = \frac{1}{2}\gamma X^2 + \epsilon \sum_{k=1}^{N} |n_k| + \eta \sum_{k=1}^{N} \frac{{n_k}^2}{\tau_k} - \frac{1}{2}\gamma \sum_{k=1}^{N} {n_k}^2 + \lambda \sum_{k=1}^{N} \tau_k {\sigma_k}^2 {x_k}^2$$

4. Varying Time Interval Linear Impact Jacobian:

$$\frac{\partial \mathbb{U}[x]}{\partial x_j} = 2\left\{\lambda \tau_j {\sigma_j}^2 x_j - \tilde{\eta}\left[\frac{x_{j+1}}{\tau_{j+1}} - \frac{x_{j-1}}{\tau_j} - x_j\left(\frac{1}{\tau_j} + \frac{1}{\tau_{j+1}}\right)\right]\right\} + \frac{1}{2}\gamma\left[x_{j-1} - 2x_j + x_{j+1}\right]$$

Extrema requires that

$$\frac{\partial \mathbb{U}[x]}{\partial x_j} = 0 \;\forall\; j = 1, \cdots, N-1$$

5. Estimation Quantities for Numerical Optimization: To carry out the numerical optimization, one needs the Jacobian (i.e., gradient) and the Hessian of the optimizer objective function in terms of

$$x_j \; \forall \; j = 1, \cdots, N - 1$$

This has to be computed for each of the following quantities:

a. Trajectory Slice Permanent Impact Function Expectation Left Holdings
b. Trajectory Slice Permanent Impact Function Expectation Right Holdings
c. Trajectory Slice Permanent Impact Function Expectation Cross Holdings Jacobian
d. Trajectory Slice Temporary Impact Function Expectation Left Holdings
e. Trajectory Slice Temporary Impact Function Expectation Right Holdings
f. Trajectory Slice Temporary Impact Function Expectation Cross Holdings Jacobian
g. Trajectory Slice Permanent Impact Function Variance Left Holdings
h. Trajectory Slice Permanent Impact Function Variance Right Holdings
i. Trajectory Slice Permanent Impact Function Variance Cross Holdings Jacobian
j. Trajectory Slice Temporary Impact Function Variance Left Holdings
k. Trajectory Slice Temporary Impact Function Variance Right Holdings
l. Trajectory Slice Temporary Impact Function Variance Cross Holdings Jacobian
m. Trajectory Slice Core Market Function Expectation Left Holdings
n. Trajectory Slice Core Market Function Expectation Right Holdings
o. Trajectory Slice Core Market Function Expectation Cross Holdings Jacobian
p. Trajectory Slice Core Market Function Variance Left Holdings
q. Trajectory Slice Core Market Function Variance Right Holdings
r. Trajectory Slice Core Market Function Variance Cross Holdings Jacobian
s. Trajectory Permanent Impact Function Expectation Left Holdings

t. Trajectory Permanent Impact Function Expectation Right Holdings
u. Trajectory Permanent Impact Function Expectation Cross Holdings Jacobian
v. Trajectory Temporary Impact Function Expectation Left Holdings
w. Trajectory Temporary Impact Function Expectation Right Holdings
x. Trajectory Temporary Impact Function Expectation Cross Holdings Jacobian
y. Trajectory Permanent Impact Function Variance Left Holdings
z. Trajectory Permanent Impact Function Variance Right Holdings
aa. Trajectory Permanent Impact Function Variance Cross Holdings Jacobian
bb. Trajectory Temporary Impact Function Variance Left Holdings
cc. Trajectory Temporary Impact Function Variance Right Holdings
dd. Trajectory Temporary Impact Function Variance Cross Holdings Jacobian
ee. Trajectory Core Market Function Expectation Left Holdings
ff. Trajectory Core Market Function Expectation Right Holdings
gg. Trajectory Core Market Function Expectation Cross Holdings Jacobian
hh. Trajectory Core Market Function Variance Left Holdings
ii. Trajectory Core Market Function Variance Right Holdings
jj. Trajectory Core Market Function Variance Cross Holdings Jacobian
kk. Objective Utility Function Permanent Impact Function Expectation Left Holdings
ll. Objective Utility Function Permanent Impact Function Expectation Right Holdings
mm. Objective Utility Function Permanent Impact Function Expectation Cross Holdings Jacobian
nn. Objective Utility Function Temporary Impact Function Expectation Left Holdings
oo. Objective Utility Function Temporary Impact Function Expectation Right Holdings
pp. Objective Utility Function Temporary Impact Function Expectation Cross Holdings Jacobian
qq. Objective Utility Function Permanent Impact Function Variance Left Holdings

rr. Objective Utility Function Permanent Impact Function Variance Right Holdings
ss. Objective Utility Function Permanent Impact Function Variance Cross Holdings Jacobian
tt. Objective Utility Function Temporary Impact Function Variance Left Holdings
uu. Objective Utility Function Temporary Impact Function Variance Right Holdings
vv. Objective Utility Function Temporary Impact Function Variance Cross Holdings Jacobian
ww. Objective Utility Function Core Market Function Expectation Left Holdings
xx. Objective Utility Function Core Market Function Expectation Right Holdings
yy. Objective Utility Function Core Market Function Expectation Cross Holdings Jacobian
zz. Objective Utility Function Core Market Function Variance Left Holdings
aaa. Objective Utility Function Core Market Function Variance Right Holdings
bbb. Trajectory Slice Core Market Function Variance Cross Holdings Jacobian

6. <u>Permanent Impact Expectation Left Sensitivity</u>:

$$\mathbb{E}_{P,k} = s\tau_k x_k g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

$$\frac{\partial \mathbb{E}_{P,k}}{\partial x_{k-1}} = s\tau_k x_k \frac{\partial g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_{k-1}}$$

$$\frac{\partial^2 \mathbb{E}_{P,k}}{\partial {x_{k-1}}^2} = s\tau_k x_k \frac{\partial^2 g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial {x_{k-1}}^2}$$

$$s = sign\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

7. Permanent Impact Expectation Right Sensitivity:

$$\mathbb{E}_{P,k} = s\tau_k x_k g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

$$\frac{\partial \mathbb{E}_{P,k}}{\partial x_k} = s\tau_k g\left(\frac{x_k - x_{k-1}}{\tau_k}\right) + s\tau_k x_k \frac{\partial g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_k}$$

$$\frac{\partial^2 \mathbb{E}_{P,k}}{\partial {x_k}^2} = 2s\tau_k \frac{\partial g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_k} + s\tau_k x_k \frac{\partial^2 g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial {x_k}^2}$$

8. Permanent Impact Expectation Cross Jacobian:

$$\mathbb{E}_{P,k} = s\tau_k x_k g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

$$\frac{\partial^2 \mathbb{E}_{P,k}}{\partial x_{k-1} \partial x_k} = s\tau_k x_k \frac{\partial^2 g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_{k-1} \partial x_k} + s\tau_k \frac{\partial g\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_{k-1}}$$

9. Temporary Impact Expectation Left Sensitivity:

$$\mathbb{E}_{T,k} = (x_k - x_{k-1}) h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

$$\frac{\partial \mathbb{E}_{T,k}}{\partial x_{k-1}} = -h\left(\frac{x_k - x_{k-1}}{\tau_k}\right) + (x_k - x_{k-1}) \frac{\partial h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_{k-1}}$$

$$\frac{\partial^2 \mathbb{E}_{T,k}}{\partial {x_{k-1}}^2} = -2 \frac{\partial h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_{k-1}} + (x_k - x_{k-1}) \frac{\partial^2 h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial {x_{k-1}}^2}$$

10. Temporary Impact Expectation Right Sensitivity:

$$\mathbb{E}_{T,k} = (x_k - x_{k-1})h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

$$\frac{\partial \mathbb{E}_{T,k}}{\partial x_k} = h\left(\frac{x_k - x_{k-1}}{\tau_k}\right) + (x_k - x_{k-1})\frac{\partial h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_k}$$

$$\frac{\partial^2 \mathbb{E}_{T,k}}{\partial {x_k}^2} = 2\frac{\partial h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_k} + (x_k - x_{k-1})\frac{\partial^2 h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial {x_k}^2}$$

11. Temporary Impact Expectation Cross Jacobian:

$$\mathbb{E}_{T,k} = (x_k - x_{k-1})h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)$$

$$\frac{\partial^2 \mathbb{E}_{T,k}}{\partial x_{k-1}\partial x_k} = -\frac{\partial h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_k} + \frac{\partial h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_{k-1}} + (x_k - x_{k-1})\frac{\partial^2 h\left(\frac{x_k - x_{k-1}}{\tau_k}\right)}{\partial x_k \partial x_{k-1}}$$

12. Trajectory Jacobian and Hessian Computation: In general, the trajectory Jacobian's and the Hessian's may be computed as a sequential, aggregate accumulations over the corresponding slices, with one very critical caveat. In the automated sensitivity generation schemes, all sensitivities to the left-most and the right-most nodes must be excluded, since these do not constitute the control nodes.
13. Power Objective Function Rationale/Formulation: A generalization of the mean-variance optimization and the Value-at-risk schemes is the power objective function formulation

$$\mathbb{U}[x] = \mathbb{E}[x] + \lambda(\mathbb{V}[x])^p$$

$$p = 1$$

corresponds to the regular mean-variance optimization scheme, and

$$p = 0.5$$

corresponds to the liquidity based VaR formulation.

14. Liquidity VaR Control Jacobian/Hessian:

$$\mathbb{U}[x] = \mathbb{E}[x] + \lambda(\mathbb{V}[x])^{p}$$

$$\frac{\partial \mathbb{U}[x]}{\partial x_i} = \frac{\partial \mathbb{E}[x]}{\partial x_i} + \lambda p(\mathbb{V}[x])^{p-1}\frac{\partial \mathbb{V}[x]}{\partial x_i}$$

$$\frac{\partial^2 \mathbb{U}[x]}{\partial x_i \partial x_j} = \frac{\partial^2 \mathbb{E}[x]}{\partial x_i \partial x_j} + \lambda p(p-2)(\mathbb{V}[x])^{p-2}\frac{\partial \mathbb{V}[x]}{\partial x_i}\frac{\partial \mathbb{V}[x]}{\partial x_j} + \lambda p(\mathbb{V}[x])^{p-1}\frac{\partial^2 \mathbb{V}[x]}{\partial x_i \partial x_j}$$

References

- Admati, A., and P. Pfleiderer (1988): A Theory of Intra-day Patterns: Volume and Price Variability *Review of Financial Studies* **1** 3-40.
- Almgren, R., and N. Chriss (1999): Value under Liquidation *Risk* **12 (12)** 61-63.
- Almgren, R., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.
- Artzner, P., F. Delbaen, J. M. Eber, and D. Heath (1997): Thinking Coherently *Risk* **10 (11)** 68-71.
- Beaver, W. (1968): The Information Content of Annual Earnings Announcements, in: *Empirical Research in Accounting; Selected Studies*. Supplement to *Journal of Accounting Research* 67-92.
- Bertsekas, D. P. (1976): *Dynamic Programming and Stochastic Control* **Academic Press**.

- Bertsimas, D., and A. W. Lo (1998): Optimal Control of Execution Costs *Journal of Financial Markets* **1** 1-50.
- Brown, K., W. Harlow, and S. Tinic (1988): Risk Aversion, Uncertain Information, and Market Efficiency *Journal of Financial Economics* **22** 355-385.
- Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997): *Econometrics of Financial Markets* **Princeton University Press**.
- Chan, L. K. C., and J. Lakonishok (1993): Institutional Trades and Intra-day Stock Price Behavior *Journal of Financial Economy* **33** 173-199.
- Chan, L. K. C., and J. Lakonishok (1995): The Behavior of Stock Prices around Institutional Trades *Journal of Finance* **50** 1147-1174.
- Charest, G. (1978): Dividend Information, Stock Returns, and Market Efficiency – II *Journal of Financial Economics* **6** 297-330.
- Dann, L. (1981): Common Stock Re-purchases: An Analysis of Returns to Bond-holders and Stock-holders *Journal of Financial Economics* **9** 113-138.
- Easterwood, J. C., and S. R. Nutt (1999): Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism? *Journal of Finance* **54 (5)** 1777-1797.
- Fama, E., L. Fisher, M. Jensen, and R. Roll (1969): The Adjustment of Stock Prices to new Information *International Economic Review* **10** 1-21.
- Grinold, R. C., and R. N. Kahn (1999): *Active Portfolio Management 2nd Edition* **McGraw-Hill**.
- Holthausen, R. W., R. W. Leftwich, and D. Mayers (1987): The Effects of Large Block Transactions on Security Prices: A Cross-Sectional Analysis *Journal of Financial Economy* **19** 237-267.
- Holthausen, R. W., R. W. Leftwich, and D. Mayers (1987): Large Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Price Effects *Journal of Financial Economy* **26** 71-95.
- Kalay, A., and U. Loewentstein (1985): Predictable Events and Excess Returns: The case of Dividend Announcements *Journal of Financial Economics* **14** 423-449.

- Keim, D. B., and A. Madhavan (1995): Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Traders *Journal of Financial Economy* **37** 371-398.
- Keim, D. B., and A. Madhavan (1997): Transaction Costs and Investment Style: An Inter-exchange Analysis of Institutional Equity Trades *Journal of Financial Economy* **46** 265-292.
- Kim, O., and R. Verrecchia (1991): Market Reaction to Anticipated Announcements *Journal of Financial Economics* **30** 273-310.
- Kraus, A., and H. R. Stoll (1972): Price Impacts of Block Trading at the New York Stock Exchange *Journal of Finance* **27** 569-588.
- Krinsky, I., and J. Lee (1996): Earnings Announcements and the Components of the Bid-Ask Spread *Journal of Finance* **51** 1523-1555.
- Kyle, A. S. (1985): Continuous Auctions and Insider Trading *Econometrica* **53** 1315-1336.
- Lee, C. M. C., B. Mucklow, and M. J. Ready (1993): Spreads, Depths, and the Impact of Earnings Information: An Intra-day Analysis *Review of Financial Studies* **6** 345-374.
- Lo, A. W., and A. C. MacKinlay (1988): Stock Market Prices do not follow Random Walks: Evidence from a Simple Specification Test *Review of Financial Studies* **1** 41-66.
- Morse, D. (1981): Price and Trading Volume Reaction surrounding Earnings Announcements: A Close Examination *Journal of Accounting Research* **19** 374-383.
- Patell, J. M., and M. A. Wolfson (1984): The Intra-day Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements *Journal of Financial Economics* **13** 223-252.
- Perold, A. F. (1988): The Implementation Short-fall: Paper versus Reality *Journal of Portfolio Management* **14** 4-9.
- Ramaswami, M. (1999): Stock Volatility declines after Earnings are announced – Honest! *Global Weekly Investment Strategy* **Lehman Brothers**.

Non-linear Impact and Trading-Enhanced Risk

Abstract

1. Price and Market Impact Volatility: Almgren (2003) determines optimal trading strategies for the liquidation of a large single-asset portfolio to minimize a combination of volatility risks and market impact costs.
2. Power Law Market Impact Function: The market impact cost is taken to be a power law of the trading rate with an arbitrary positive exponent. This includes, for example, the square root law that has been proposed based on market microstructure theory.
3. Holdings Size Dependent Characteristic Time: In analogy with the linear model, a *characteristic time* is defined for optimal trading, which now depends on the initial portfolio size and decreases as the execution proceeds.
4. Trade Size Dependent Liquidity Volatility: Also considered is a model in which the uncertainty of the realized price is increased by demanding rapid execution; it is shown that the optimal trajectories are defined by a *critical portfolio size* above which this effect is dominant and below which this effect may be neglected.

Introduction

1. Active vs. Passive Execution Strategy: In the execution of large portfolio transactions, a trading strategy must be determined that balances the risk of delayed execution against the cost of rapid execution; the choice is roughly between an *active* and *passive* trading strategy (Hasbrouck and Schwartz (1988), Wagner and Banks (1992)).
2. Construction of Optimal Execution Strategies: Several papers have constructed optimal strategies for the problem (Almgren and Chriss (1999), Grinold and Kahn

(1999), Almgren and Chriss (2000), Konishi and Makimoto (2001) under the assumption that the liquidity costs per share traded are a linear function of trading rate or block size, and that the only source of volatility in execution is the price volatility of the underlying asset.

3. Price Effect of Block Trades: There is an extensive literature studying the effects of block trades on prices – see Kraus and Stoll (1972), Holthausen, Leftwich, and Mayers (1987, 1990), Chan and Lakonishok (1993, 1995), Keim and Madhavan (1995, 1997), and Koski and Michaely (2000).
4. Assumption of Linear Trading Costs: In practice linearity of trading costs is an unrealistic assumption. Perold and Salomon Jr. (1991) have argued that the liquidity premium per share demanded by the market will either be a convex or a concave function of the block size depending on whether the market's perception is that the trader is information driven or liquidity driven, respectively.
5. Barra Market Impact Liquidity Premium: In the Barra Market Impact Model (Loeb (1983), Kahn (1993), Barra (1997), Grinold and Kahn (1999)) it is argued, based upon the detailed analysis of the risk-reward choice faced by the equity market maker that the liquidity premium per share should grow as the square root of the block size traded.
6. Block Size Dependent Liquidity Premium: Electronic trading systems such as Optimark (Rickard and Torre (1999)) have been constructed to allow traders specify precisely what liquidity premium they are willing to pay as a function of the block size, and search for clearing opportunities in the mismatch between profiles of different market participants. These effects can be captured by introducing *nonlinear impact functions* into the cost function which is minimized to determine the optimal trading strategies.
7. Approaches of Nonlinear Models: Although linear models are commonly used in empirical regression analyses for simplicity, nonlinear models can often be emulated by dividing trades into categories by size (Bessembinder and Kaufmann (1997), Huang and Stoll (1997)). In fact, Chakravarthy (2001) argues medium-sized trades have a disproportionately large effect on prices.

8. <u>Handling Non-Deterministic Liquidity Premiums</u>: An additional effect not considered in the theoretical strategies considered in the previous work is that the liquidity premium demanded by the market is not deterministic. In fact the premium will depend on the presence in the market at that instant of participants who are willing to take the other side of the trade.
9. <u>Motivation for Trading Enhanced Risk</u>: Since the presence of these counterparties cannot be predicted in advance, it represents an additional source of risk incurred by the trading profile. That is, a more complete model should include *trading-enhanced risk* representing the increased uncertainty in the execution price incurred by demanding rapid execution of large blocks.
10. <u>Manifestation of Stochastic Liquidity Premiums</u>: Trading-enhanced risk is an implicit feature in the model described in Rickard and Torre (1999). Chordia, Subrahmanyam, and Anshuman (2001) and Hasbrouck and Seppi (2001) argue that liquidity fluctuates due to intrinsic variations in the market activity independent of the trade size. This effect is included in the model by Almgren (2003) via the constant term $f(0)$ but additional interest is in the *increase* in the execution price uncertainty due to larger block sizes.
11. <u>Nonlinear Block Size Dependence</u>: Thus Almgren (2003) extends the models of Grinold and Kahn (1999) and Almgren and Chriss (2000) in a few important ways. First, the liquidity premium, expressed as an unfavorable motion of the price per share, may be an increasing nonlinear function of the trading rate and the block size – one is considered to be a proxy for the other.
12. <u>Power Law Premium - Closed Form</u>: This cost is reduced by trading slowly, but it must be balanced against the volatility risk incurred by holding the initial portfolio longer than is necessary. In particular, exact solutions are provided in the case this function is a power law with an arbitrary positive exponent, which covers the range of behavior outlined above.
13. <u>Dependence on Initial Portfolio Size</u>: Whereas in the linear case optimal trajectories are characterized by a single *characteristic time* independent of the initial portfolio size, in the nonlinear case the characteristic time depends upon the initial portfolio size, and scales appropriately as the remaining portfolio diminishes during trading.

14. Comparison with Price Volatility Risk: The realized price per share itself is a random variable, whose variance increases ith the increased rate of trading. This introduces an additional source of risk in addition to the volatility. In contrast to the effect of market volatility, this additional risk is *decreased* by trading slowly, submitting small blocks for execution at each time.
15. Volatile Liquidity - Closed Form Solutions: Nearly explicit optimal solutions including this effect can be constructed, and an asymptotic analysis can be used to show that the effect of trading-enhanced risk is most important for large initial portfolios. Indeed for any given set of parameters there is a characteristic portfolio size above which the optimal strategy is determined by the need to reduce trading-enhanced risk, and below which this effect may be ignored.

The Model

1. Holdings Trade and Liquidation Time: The general framework followed is that from Almgren and Chriss (2000). At time

$$t = 0$$

X shares of an asset are held, which are to be completely liquidated by the time

$$t = T$$

The initial size X is positive for a sell program and negative for a buy program; in the former case, there is a long exposure to the market until all the holdings have been eliminated, while in the latter case there is short exposure to the market until the purchase to which the trader has committed to at

$$t = 0$$

is completed. The focus here is on the case

$$X > 0$$

In the case of a portfolio trading problem X may be a vector, but the consideration here is only on a single asset.

2. The Problem Trade List Determination: $x(t)$ denotes the holdings at time t with

$$x(0) = X$$

and

$$x(T) = 0$$

The problem is to choose an optimal function $x(\cdot)$ so as to minimize a chosen cost functional. Later the limit

$$T \to \infty$$

will be taken in which the natural execution time emerges as a result of the analysis, but for now the consideration is on an exogenously imposed time horizon.

3. Origin of the *Static* Strategy: It is a rather surprising fact that in the absence of serial correlation in the asset price movements, the optimal price may be determined *statically* at the start of the trading. Unless the market parameters change, observations of price movements in the course of trading do not convey any information that would lead to a change in the strategy.
4. Evenly Spaced Discrete Time Intervals: The analysis starts with the construction of a discrete time model. Thus for a given trading interval

$$\tau > 0$$

$$t_k = k\tau$$

for

$$k = 0, \cdots, N$$

with

$$N = \frac{T}{\tau}$$

and let x_k be the holdings at time t_k with

$$x_0 = X$$

and

$$x_N = 0$$

The sales between times t_k and t_{k-1} are

$$n_k = x_k - x_{k-1}$$

corresponding to the velocity

$$v_k = \frac{n_k}{\tau} \ shares\ per\ unit\ time$$

Thus

$$x_k = X - \sum_{j=1}^{k} n_j = \sum_{j=k+1}^{N} n_j$$

$$k = 0, \cdots, N$$

5. Generating the Optimal Trade List: In the discrete time model there is no assumption that the shares are traded at a uniform rate *within* each interval. Rather the assumption is that the trader achieves the optimal execution possible subject to the constraint that n_k shares are to be traded in the next time interval τ. The functions introduced below are a model to describe the trader's best efforts.
6. Temporary/Permanent Market Impact Components: On a standard manner (Stoll (1985)) the impact is divided into a permanent and a temporary component. Thus S_k describes the price per share of the asset that is publically available in the market.
7. Discrete Arithmetic Permanent Impact Component: The price satisfies the arithmetic random walk

$$S_k = S_{k-1} + \sigma\sqrt{\tau_k}\xi_k - \tau_k g\left(\frac{n_k}{\tau_k}\right) = S_0 + \sigma \sum_{j=1}^{k} \sqrt{\tau_j}\xi_j - \sum_{j=1}^{k} \tau_j g(v_j)$$

   where ξ_j are independent random variables with zero mean and unit variance, σ is an *absolute* (not percentage) volatility, $g(v)$ is the *permanent impact function* representing the effect of the share price of the information conveyed by the trade. This effect is generally small, and below $g(v)$ is taken to be a linear function, in which case it will have no effect on determining the optimal strategy.
8. Nonlinear Temporary Impact Component: The price that one actually gets on the k^{th} trade is

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau_k}\right) + \frac{1}{\sqrt{\tau_k}} f\left(\frac{n_k}{\tau_k}\right)\tilde{\xi}_k$$

$$k = 1, \cdots, N$$

Here $h(v)$ is a nonlinear *temporary impact function* representing the price concession one must accept in order to trade

$$n_k = v_k \tau_k$$

shares in time τ_k. The random variables $\tilde{\xi}_k$ are independent of each other and of ξ_k with zero mean and unit variance. The new function $f(v)$ represents the uncertainty of the trade execution as a function of the block size.

9. Liquidity Volatility Term Time Dependence: The factor $\frac{1}{\sqrt{\tau_k}}$ in the last term of

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau_k}\right) + \frac{1}{\sqrt{\tau_k}} f\left(\frac{n_k}{\tau_k}\right) \tilde{\xi}_k$$

simply represents a scaling of the parameters if τ_k is fixed and finite. When τ_k varies, for example when the continuous time

$$\tau_k \to 0$$

is taken, this factor is necessary to preserve the effect of the trading-enhanced risk.

10. Liquidity Volatility Incremental Time Dependence: If the above term were not present, then breaking a block into several smaller blocks would diversify away the risk due to the uncertainty of each one, regardless of the form of the risk.
11. *Capture* of the Trade Program: The *capture* of the trade program is the total cash received

$$\sum_{k=1}^{N} n_k \tilde{S}_k = XS_0 + \sigma \sum_{k=1}^{N} \sqrt{\tau_k} x_k \xi_k - \sum_{k=1}^{N} \tau_k x_k g(v_k) + \sum_{k=1}^{N} \sqrt{\tau_k} v_k f(v_k) \tilde{\xi}_k$$
$$- \sum_{k=1}^{N} \tau_k v_k h(v_k)$$

12. <u>The Trade Program *Implementation Cost*</u>: Discounting is ignore since the trading horizon is short. The *implementation cost* is $XS_0 - \sum_{k=1}^{N} n_k \tilde{S}_k$ - a random variable due to uncertainties in price movements and realized prices.
13. <u>Components of the *Implementation Cost*</u>: Note that the implementation cost includes both the costs of finite liquidity and price uncertainty due to delayed execution. This is the *implementation shortfall* of Perold (1988) – see also Jones and Lipson (1999).
14. <u>Implementation Cost Mean and Variance</u>: Its expectation and variance at

$$t = 0$$

depend on the free parameters $x_1, \cdots, x_{N-1}$ of the trade strategy:

$$\mathbb{E}[x_1, \cdots, x_{N-1}] = \sum_{k=1}^{N} \tau_k x_k g(v_k) + \sum_{k=1}^{N} \tau_k v_k h(v_k)$$

$$\mathbb{V}[x_1, \cdots, x_{N-1}] = \sum_{k=1}^{N} \tau_k {\sigma_k}^2 {x_k}^2 + \sum_{k=1}^{N} \tau_k {v_k}^2 f^2(v_k)$$

15. <u>Mean Variance Optimal Static Strategies</u>: A rational trader will construct his or her own strategies to minimize some combination of $\mathbb{E}[x]$ and $\mathbb{V}[x]$. As t advances the values of $\mathbb{E}[x]$ and $\mathbb{V}[x]$ change, but if $\mathbb{E}[x]$ and $\mathbb{V}[x]$ are constructed using a classic mean-variance approach, the optimal strategy continues to be the one determined initially (Almgren and Chriss (2000), Huberman and Stanzl (2005)).
16. <u>Continuous Time Limit Trading Strategy</u>: Now, for analytical convenience, the continuous time limit

$$\tau \to 0$$

is taken. The trade strategy becomes a continuous path $x(t)$ and the block sizes n_k are assumed to be well-behaved so that

$$v_k \to v(\tau_k k)$$

with

$$v(t) = -\dot{x}(t)$$

17. Continuous Time Mean and Variance: The above expressions have finite limits

$$\mathbb{E}[x] = \int_0^T \left[x(t) g\big(v(t)\big) + v(t) h\big(v(t)\big)\right] dt$$

$$\mathbb{V}[x] = \int_0^T \left[\sigma^2 x^2(t) + v^2(t) f^2\big(v(t)\big)\right] dt$$

where the square brackets indicate that these are *functionals* of the entire continuous-time path $x(t)$.

18. Caveat behind Continuous Time Analytics: It needs to be emphasized that the continuous time limit is simply an analytical device for obtaining solutions when τ_k is reasonably small; in reality the discreteness of the trading intervals must be taken into account in order to correctly describe trading-enhanced risk.
19. Mean Variance Optimization Objective Function: Introducing the risk-aversion parameter λ, the combined quantity

$$\mathbb{U}[x] = \mathbb{E}[x] + \lambda \mathbb{V}[x]$$

is minimized. Whether or not mean variance optimization is appropriate for a particular case, λ may be considered to be a Lagrange/KKT multiplier for the constrained problem of minimizing $\mathbb{E}[x]$ for a given $\mathbb{V}[x]$ and used to construct an efficient frontier in the space of trading trajectories.

20. <u>VaR Based Optimization Objective Functions</u>: More general weightings of risk, including Value-at-Risk, present thorny conceptual problems for time-dependent strategies (Artzner, Delbaen, Eber, and Heath (1999), Basak and Shapiro (2001)).
21. <u>The Calculus of Variations Approach</u>: Minimizing $\mathbb{U}[x]$ is a standard problem in the calculus of variations:

$$\underset{x(t)}{min}\ \mathbb{U}[x(t)] = \underset{x(t)}{min} \int_{0}^{T} F\big(x(t), -\dot{x}(t)\big)dt$$

with

$$F(x, v) = xg(v) + vh(v) + \lambda\sigma^2 x^2 + \lambda v^2 f^2(v)$$

22. <u>Perturbation Stationarity: Euler-Lagrange Equation</u>: Stationarity to small perturbations requires that the optimal $x(t)$ solve the Euler-Lagrange equation

$$0 = \frac{\partial F\big(x(t), -\dot{x}(t)\big)}{\partial x(t)} + \frac{d}{dt}\left[\frac{\partial F\big(x(t), -\dot{x}(t)\big)}{\partial v(t)}\right]$$
$$= \frac{\partial F\big(x(t), -\dot{x}(t)\big)}{\partial x(t)} + \dot{x}(t)\frac{\partial^2 F\big(x(t), -\dot{x}(t)\big)}{\partial x(t)\partial v(t)}$$
$$- \ddot{x}(t)\frac{\partial^2 F\big(x(t), -\dot{x}(t)\big)}{\partial v^2(t)}$$

– a second order ordinary differential equation to be solved with respect to the given endpoints $x(0)$ and $x(T)$

23. Integration into the First Order Form: Since $F(x(t), -\dot{x}(t))$ does not depend explicitly on t multiplying throughout by $\dot{x}(t)$ and integrating results in the first-order equation

$$F(x(t), -\dot{x}(t)) + \dot{x}(t)\frac{\partial F(x(t), -\dot{x}(t))}{\partial v(t)} = constant$$

24. Application to Optimal Trajectory Determination: In the current case one obtains

$$P(-\dot{x}(t)) - P(v_0) = x[g(-\dot{x}(t)) + \dot{x}(t)g'(-\dot{x}(t))] + \lambda\sigma^2 x^2$$

with

$$P(v) = v^2\frac{\partial h(v)}{\partial v} + \lambda v^2\left[f^2(v) + 2vf(v)\frac{\partial f(v)}{\partial v}\right] = v^2\frac{\partial[h(v) + \lambda v f^2(v)]}{\partial v}$$

25. Properties of the Almgren "P" Function: The constant of integration

$$v_0 = -\dot{x}(t)|_{x=0}$$

is the velocity with which $x(t)$ hits

$$x = 0$$

For a sell program with

$$X > 0$$

$$v_0 \geq 0$$

and conversely for a buy program. Note that

$$P(0) = 0$$

additional assumption is that $P(v)$ is always an *increasing* function of v and hence invertible.

26. Explicit Solutions - Key Simplifying Assumptions: Almgren (2001) makes two simplifying assumptions to obtain explicit solutions.
    a. Permanent impact is linear in the trading rate.
    b. The imposed time horizon is infinite.
27. Linear Permanent Market Impact Function: A linear cost function

$$g(v) = \gamma v$$

gives a total cost γX independent of the path $x(t)$. The first term on the right side of

$$P\big(-\dot{x}(t)\big) - P(v_0) = x\big[g\big(-\dot{x}(t)\big) + \dot{x}(t)g'\big(-\dot{x}(t)\big)\big] + \lambda\sigma^2 x^2$$

vanishes, and then since $\dot{x}(t)$ appears only on the left side and x itself appears only on the right side, the general solution can be written in the quadrature form as

$$\int_{x(t)}^{X} \frac{dt}{P^{-1}[\lambda\sigma^2 x^2 + P(v_0)]} = t$$

28. Bid Ask Permanent Linear Absent: The constant v_0 is to be chosen so that

$$x = 0$$

corresponds to

$$T = 0$$

Note also that any constant in h disappears; the bid-ask spread does not affect the optimal strategy.

29. No Extraneously Specified Liquidation Time: Since $P(\cdot)$ is an increasing function, so is $P^{-1}(\cdot)$. It is thus clear that as v_0 decreases towards zero, the liquidation time T increases.
30. Invoking Longest Possible Liquidation Time: If no time horizon is exogenously imposed, the longest possible liquidation time can be obtained by setting

$$v_0 = 0$$

which leads to the quadrature problem

$$\int_{x(t)}^{X} \frac{dt}{P^{-1}[\lambda\sigma^2 x^2]} = t$$

31. Tractability of the above Solution: Often analytic solutions to the above problem can be found when

$$\int_{x(t)}^{X} \frac{dt}{P^{-1}[\lambda\sigma^2 x^2 + P(v_0)]} = t$$

with

$$v_0 \neq 0$$

would be too intractable. These solutions will still give nearly complete liquidations in finite time determined by market parameters.

Nonlinear Cost Functions

1. Power Law Temporary Impact Functions: Restricting the attention to the sell program, with

$$v \geq 0$$

the temporary impact functions are taken to be

$$h(v) = \eta v^k$$

$$f(v) = 0$$

with

$$k > 0$$

- for a buy program the signs will be changed in an obvious way. The linear case corresponds to

$$k = 1$$

2. Temporary Impact Almgren "P" Function: As noted above a possible constant in h corresponding to the bid-ask spread has been neglected. Then

$$P(v) = \eta k v^{k+1}$$

which, for the case of a general finite time horizon with

$$v_0 \geq 0$$

leads to the quadrature problem

$$\int_{x(t)}^{X} \left(\frac{\lambda\sigma^2}{\eta k} x^2 + {v_0}^{k+1}\right)^{-\frac{1}{k+1}} dx = t$$

3. Longest Optimal Trajectory Explicit Solution: Taking

$$v_0 \geq 0$$

explicit solutions for the longest optimal trajectories can be obtained:

$$\frac{x(t)}{X} = \begin{cases} \left(1 + \frac{1-k}{1+k}\frac{t}{T_*}\right)^{-\frac{1+k}{1-k}} & 0 < k < 1 \\ e^{-\frac{t}{T_*}} & k = 1 \\ \left(1 - \frac{k-1}{k+1}\frac{t}{T_*}\right)^{\frac{k+1}{k-1}} & k > 1 \end{cases}$$

4. Characteristic Time for Optimal Execution: Here the *characteristic time* is

$$T_* = \left(\frac{k\eta X^{k-1}}{\lambda\sigma^2}\right)^{\frac{1}{k+1}}$$

This is the analog of the *half-life* in the linear case. Only in the linear case

$$k = 1$$

is T_* independent of the initial portfolio size X. For

$$k \neq 1$$

the characteristic time depends on the initial size as

$$T_* \sim X^{\frac{k-1}{k+1}}$$

5. <u>Sub Linear Power Law Exponent</u>: For

$$k < 1$$

rapid trading is *under*-penalized relative to the linear case. As the portfolio size increases, volatility risk dominates the trading costs, and the optimal trading time *decreases* since the exponent is negative.

6. <u>Supra Linear Power Law Exponent</u>: For

$$k > 1$$

rapid trading is *over*-penalized relative to the linear case. As the portfolio size increases, the trading cost dominates the volatility risk, and the optimal trading time *increases*, since the exponent is positive. For example, if

$$k = 3$$

then

$$T_* \sim \sqrt{X}$$

7. <u>Characteristic Time vs Half Life</u>: As the portfolio size decreases to zero, reconciliation of the optimal trajectory would use a different starting value X and hence a different time T_*. The meaning of T_* is thus a little less fundamental than in the linear case. However, T_* scales in exactly the right way to make $x(t)$ still a static solution.

8. Intuition behind the Characteristic Time: For more intuition, note that the initial rate of selling is

$$-\dot{x}(0) = \frac{X}{T_*}$$

and T_* is the solution to the relation

$$\lambda\sigma^2 X^2 T = k\eta \left(\frac{X}{T}\right)^k X$$

9. Characteristic Time as a Cost Balance: The left side is the risk penalty associated with holding X shares for a time T, and the right side, up to a factor k, is $Xh\left(\frac{X}{T}\right)$, the impact cost associated with selling X shares over a time T (without the constant term representing the bid-ask spread, which does not impact the optimal solution).
10. The Longest Optimal Execution Time: For

$$k > 1$$

the trajectory reaches

$$x = 0$$

with

$$v = 0$$

at a finite time

$$T_{MAX} = \frac{k+1}{k-1} T_*$$

Thus these trajectories are the solution for finite imposed time T if

$$T > T_{MAX}$$

the trajectory reaches 0 at T_{MAX} and stays there till T

11. Self-Similar Scaling Trajectory Form: Almgren (2003) contains a graphical illustration of the optimal trajectories generated from

$$\frac{x(t)}{X} = \begin{cases} \left(1 + \frac{1-k}{1+k}\frac{t}{T_*}\right)^{-\frac{1+k}{1-k}} & 0 < k < 1 \\ e^{-\frac{t}{T_*}} & k = 1 \\ \left(1 - \frac{k-1}{k+1}\frac{t}{T_*}\right)^{\frac{k+1}{k-1}} & k > 1 \end{cases}$$

The form of the portfolio is independent of a particular choice of time scale T_* and initial portfolio size X; these solutions may be easily scaled to any case.

12. Time Realization of Trajectory Differences: A sense of the differences between the solutions may be gained by noting that for short times, all optimal trajectories are fairly close to each other; but the *tail* of the trajectories are extended for small value of k which strongly penalize trading at slow rates.
13. Example: Sell-Order Exponent Dependence: For example, as shown by Almgren (2003), at

$$t = T_*$$

the optimal trajectories reduce holdings to 30%, 37%, and 42% of the initial portfolio for

$$k = 2$$

$$k = 1$$

and

$$k = \frac{1}{2}$$

respectively.

14. Example Sell Time Exponent Dependence: Further as demonstrated by Almgren (2003), at

$$\frac{t}{T_*} = 3$$

the trajectory for

$$k = 2$$

has reached

$$x = 0$$

and remains there, the trajectory for

$$k = 1$$

retains 5% of its initial holdings, and the trajectory for

$$k = 2$$

retains 12.5% of the initial holdings. The relative differences become even more pronounced as time continues.

Objective Function

1. Determination of $\mathbb{E}[x]$ and $\mathbb{V}[x]$: $\mathbb{E}[x]$ and $\mathbb{V}[x]$ can be explicitly computed for these solutions from

$$\mathbb{E}[x] = \int_0^T [x(t)g(v(t)) + v(t)h(v(t))]dt$$

$$\mathbb{V}[x] = \int_0^T [\sigma^2 x^2(t) + v^2(t)f^2(v(t))]dt$$

and hence the frontier can be drawn. In doing this the contributions from $g(v)$ and the term ϵX in $\mathbb{E}[x]$ are neglected.

2. Closed Form for $\mathbb{E}_\lambda[x]$ and $\mathbb{V}_\lambda[x]$: Then for a general k

$$\mathbb{E}_\lambda[x] = \frac{k+1}{3k+1}\eta\left(\frac{X}{T_*}\right)^{k+1} T_* = \frac{k+1}{3k+1}\eta\left(\frac{\eta\sigma^{2k}X^{3k+1}}{k^k}\lambda^k\right)^{\frac{1}{k+1}} T_*$$

$$\mathbb{V}_\lambda[x] = \frac{k+1}{3k+1}\sigma^2 T_* X^2 = \frac{k+1}{3k+1}\left(\frac{k\eta\sigma^{2k}X^{3k+1}}{\lambda}\right)^{\frac{1}{k+1}}$$

3. The $(\mathbb{E}, \mathbb{V})$ Efficient Frontier Curve: As λ varies $(\mathbb{E}, \mathbb{V})$ moves along the hyperboloid-like curve

$$\mathbb{E}_\lambda[x](\mathbb{V}_\lambda[x])^k = \left(\frac{k+1}{3k+1}\right)^{k+1} \eta\sigma^{2k}X^{3k+1}$$

For any positive λ there is a unique solution.

4. <u>Asymptotics of $\lambda \to 0$ and $\lambda \to \infty$</u>: As

$$\lambda \to 0$$

one gets

$$T_* \to \infty$$

$$\mathbb{E}_\lambda[x] \to 0$$

and

$$\mathbb{V}_\lambda[x] \to \infty$$

i.e., optimizing expected cost without regard to variance leads to use of all available time. As

$$\lambda \to \infty$$

it can be seen that

$$T_* \to 0$$

$$\mathbb{E}_\lambda[x] \to \infty$$

and

$$\mathbb{V}_\lambda[x] \to 0$$

i.e., uncertainty is minimized regardless of the cost.

Almgren (2003) Example

1. Reference Trading Rate/Market Depth: Almgren (2003) provides a sample methodology for the estimation of the parameters. The first is to choose a representative level of trading rate v_{REF} If a specific time period τ is chosen v_{REF} is equivalent to a certain block size

$$n_{REF} = \tau v_{REF}$$

   traded in the time period; it may be interpreted as the market *depth* in the sense of Kyle (1985) or Bondarenko (2001).
2. Choice of Reference Trading Rate: The examples below consider stocks that trade one million shares a day, and v_{REF} is taken to be 10% of that rate, or

$$v_{REF} = 100{,}000\ share/day$$

   For time period

$$\tau = 1\ hour$$

   with $6.5\ periods\ per\ day$ this rate is equivalent to trading a block of approximately $15{,}300$ shares in each hour.
3. The Corresponding Temporary Price Impact: Next the price impact h_{REF} which would be incurred by the steady trading at the reference rate v_{REF} would be chosen. The share price is assumed to be $\$50/share$ and the assumption is that trading

$$v_{REF} = 100{,}000\ shares/day$$

   incurs a price impact of 1% or $\$0.50/share$

4. Choosing the Exponent - The Rationale: Finally a choice for the value of the exponent k is made that best fits the belief of how the price impact would depend upon the trading rate for rates smaller or larger than v_{REF}.
5. Reminder Trading Rate k Dependence: The choice

$$k = 1$$

corresponds to the linear dependence of price impact on rate;

$$k > 1$$

means that *large* trading rates or block sizes have a disproportionately *large* effect on price; while

$$k < 1$$

means that large trading rates or block sizes have a relatively *smaller* impact.
6. h_{REF}/v_{REF} Based Impact Model: This impact model is then written as

$$h(v) = h_{REF}\left(\frac{v}{v_{REF}}\right)^k$$

or

$$\eta = \frac{h_{REF}}{{v_{REF}}^k}$$

7. Daily Volatility and Initial Portfolio: It is also supposed that the stock has an annual volatility of 32% for the daily price change of

$$\sigma = \$1/share \cdot \sqrt{day}$$

A portfolio of initial size

$$X = 100{,}000$$

is considered, equal to $\frac{1}{10}^{th}$ of the daily volume.

8. Construction of the Efficient Frontier: There is now enough information to construct the efficient frontier

$$\mathbb{E}_{\lambda}[x] = \frac{k+1}{3k+1}\eta\left(\frac{X}{T_*}\right)^{k+1} T_* = \frac{k+1}{3k+1}\eta\left(\frac{\eta\sigma^{2k}X^{3k+1}}{k^k}\lambda^k\right)^{\frac{1}{k+1}} T_*$$

$$\mathbb{V}_{\lambda}[x] = \frac{k+1}{3k+1}\sigma^2 T_* X^2 = \frac{k+1}{3k+1}\left(\frac{k\eta\sigma^{2k}X^{3k+1}}{\lambda}\right)^{\frac{1}{k+1}}$$

from $\mathbb{E}_{\lambda}[x]$ and $\mathbb{V}_{\lambda}[x]$ for any chosen k describing the family of solutions as the risk aversion parameter λ ranges over all the possible values

$$0 < \lambda < \infty$$

To construct particular optimal solutions a specific value for λ needs to be set.

9. Variance/Cost Dependence on λ: The results are shown numerically in the table below. For any value of k the natural liquidation time T_* increases with the *risk tolerance* parameter $\frac{1}{\lambda}$; as both increase the expected cost decreases and the variance increases.

10. λ Impact: T_*, $\mathbb{E}_{\lambda}[x]$, $\sqrt{\mathbb{V}_{\lambda}[x]}$: The table below shows that the optimal time scale T_*, the expected cost $\mathbb{E}_{\lambda}[x]$, and the standard deviation of the cost $\sqrt{\mathbb{V}_{\lambda}[x]}$ as functions of the risk tolerance parameter $\frac{1}{\lambda}$ and the temporary impact exponent k. Market and portfolio parameters are as given in the treatment above (the initial portfolio value is

$5 *million*). As k is varied, the reference values h_{REF} and v_{REF} are held constant; thus the coefficient η varies as in

$$h(v) = h_{REF}\left(\frac{v}{v_{REF}}\right)^k$$

or

$$\eta = \frac{h_{REF}}{{v_{REF}}^k}$$

Time T_* is measured in days; $\frac{1}{\lambda}$, $\mathbb{E}_\lambda[x]$, and $\sqrt{\mathbb{V}_\lambda[x]}$ are in thousands of dollars.

11. Table of $T_*, \mathbb{E}_\lambda[x], \sqrt{\mathbb{V}_\lambda[x]}$:

	$\boldsymbol{\frac{1}{\lambda}}$	$\boldsymbol{k = \frac{1}{2}}$	$\boldsymbol{k = 1}$	$\boldsymbol{k = 2}$
T_*	1	0.02	0.07	0.22
	10	0.09	0.22	0.46
	100	0.40	0.71	1.00
	1000	1.84	2.24	2.15
	10000	8.55	7.07	4.64
$\mathbb{E}_\lambda[x]$	1	221	354	462
	10	103	112	99
	100	48	35	21
	1000	22	11	5
	10000	10	4	1
$\sqrt{\mathbb{V}_\lambda[x]}$	1	11	19	30
	10	23	33	45
	100	49	59	65
	1000	105	106	96
	10000	226	188	141

12. <u>Large λ Execution Time Dependence</u>: For large values of λ, optimal trajectories all execute rapidly, to reduce the volatility risk associated with the portfolio. When the trading rate is larger than v_{REF} costs *increase* with increasing k so larger k leads to slower trading.
13. <u>λ and k Combination Impact</u>: Conversely for small λ generally trading proceeds *more slowly* than v_{REF} in order to minimize the total expected cost. In this regime *smaller* k is more expensive and leads to relatively slower trading.
14. <u>Cross-Over λ Execution Time</u>: In the intermediate parameter regime the trajectories cross-over from one behavior to the other; larger k suggests slower trading at the beginning when the rate is large, then relatively more rapid in the tail.
15. <u>Characteristic Time Reference Rate Dependence</u>: Note that

$$T_* = \left(\frac{k\eta X^{k-1}}{\lambda\sigma^2}\right)^{\frac{1}{k+1}}$$

may be re-written as

$$T_* = \left(\frac{k h_{REF}}{\lambda\sigma^2 X}\right)^{\frac{1}{k+1}} \left(\frac{X}{v_{REF}}\right)^{\frac{k}{k+1}}$$

from which it is clear that

$$T_* \to \frac{X}{v_{REF}}$$

as

$$k \to \infty$$

regardless of the values of the other parameters.

16. <u>Trade Speed vs. Cost Balance</u>: In this limit, trading more rapidly than the reference rate is very strongly penalized, while trading more slowly is almost without cost, so the optimal strategy is to always trade exactly at the critical rate.
17. <u>λ Estimation from Reference Parameters</u>: Finally since λ is a difficult parameter to select in practice, it may be observed that it can be estimated if a time scale T_* is chosen from

$$T_* = \left(\frac{k\eta X^{k-1}}{\lambda\sigma^2}\right)^{\frac{1}{k+1}}$$

and

$$h(v) = h_{REF}\left(\frac{v}{v_{REF}}\right)^k$$

or

$$\eta = \frac{h_{REF}}{{v_{REF}}^k}$$

- it can be found that

$$\lambda = k\frac{h_{REF}\left(\frac{X/_{T_*}}{v_{REF}}\right)^k X}{\sigma^2 T_* X^2}$$

18. <u>Trading Cost vs. Variance Balance</u>: The numerator is the price concession per share for trading at a concession rate $\frac{X}{T_*}$ multiplied by the total number of shares X to get the expected cost; the denominator is the variance that would be incurred by holding X

shares for time T_*. This ratio is multiplied by k to correct for nonlinearities which are ignored by this simple description.

Trading-Enhanced Risk

1. Liquidity Volatility: Functional Form Considered: Now the following functional form is taken for a sell program with

$$v \geq 0$$

$$h(v) = \eta v$$

$$f(v) = \alpha + \beta v$$

The deterministic part of the temporary impact is the linear case

$$k = 1$$

of the previous section.
2. Trading Rate Independent Volatility Component: The constant term in $f(v)$, with coefficient α, represents a constant uncertainty in the realized sale price independent of the rate of selling and of the underlying process. The total risk associated with this term is minimized by splitting the sale into as many parts as possible; thus this term pushes towards the linear trajectory.
3. Trading Rate Dependent Volatility Component: The linear term, with coefficient β, represents the increase in variance caused by non-zero amounts of selling. This term can even more strongly push toward the linear trajectory.
4. Liquidity Volatility Almgren "P" Function: Then with

$$\dot{x} = -v \leq 0$$

$$P(-\dot{x}(t)) - P(v_0) = x\left[g(-\dot{x}(t)) + \dot{x}(t)g'(-\dot{x}(t))\right] + \lambda\sigma^2 x^2$$

becomes

$$P(v) = (\eta + \lambda\alpha^2)v^2 + 4\lambda\alpha\beta v^3 + 3\lambda\beta^2 v^4$$

5. <u>Behavior of the Almgren "P" Function</u>: The polynomial $P(v)$ has

$$P(0) = 0$$

and is increasing for

$$v \geq 0$$

so the graph of the trajectory is always convex, and the inverse of P^{-1} is well-defined. For a buy program with

$$x \geq 0$$

the sign of the odd term in $P(v)$ is reversed.

6. <u>No Hard Maximum Execution Time</u>: Since

$$P(v) \sim \mathcal{O}(v^2)$$

for v near zero, the integrand appearing in the quadrature formulation

$$\int_{x(t)}^{X} \frac{dt}{P^{-1}[\lambda\sigma^2 x^2 + P(v_0)]} = t$$

behaves as $\mathcal{O}(x^{-1})$ as

$$x \to 0$$

for

$$v_0 = 0$$

and there is no “hard” maximum time as was found above for

$$k > 1$$

Constant Enhanced Risk

1. Analytical Solution for $\beta = 0$ Case: Two special cases are considered for obtaining analytical solution. The first is

$$\beta = 0$$

With this assumption the price uncertainty on each trade is independent of the size of the trade.

2. Trading Trajectory and Execution Time: A solution can then be found for

$$v_0 = 0$$

$$x(t) = Xe^{-\frac{t}{T_*}}$$

$$T_* = \sqrt{\frac{\eta + \lambda\alpha^2}{\lambda\sigma^2}}$$

3. <u>Comparison with $f(v) = 0, k = 1$ Case</u>: This is a pure exponential solution, except that the time constant has been increased by adding the additional variance per transaction to the impact coefficient

$$\eta \longmapsto \eta + \lambda\alpha^2$$

4. <u>Expressions for $\mathbb{E}_\lambda[x]$ and $\mathbb{V}_\lambda[x]$</u>: The value functions are

$$\mathbb{E}_\lambda[x] = \frac{1}{2}\eta\frac{X^2}{T_*} = \frac{1}{2}X^2\sqrt{\frac{\lambda\eta^2\sigma^2}{\eta + \lambda\alpha^2}}$$

$$\mathbb{V}_\lambda[x] = \frac{1}{2}X^2\sigma^2 T_*\left(1 + \frac{\alpha^2}{\sigma^2 T_*}\right) = \frac{1}{2}X^2\frac{\sigma}{\sqrt{\lambda}}\sqrt{\frac{\eta + 2\lambda\alpha^2}{\eta + \lambda\alpha^2}}$$

5. <u>$\lambda \to 0$ $\mathbb{E}_\lambda[x]$ and $\mathbb{V}_\lambda[x]$ Behavior</u>: The optimal value functions change in a more complicated way than the trajectory. As

$$\lambda \to 0$$

   the behavior is the same as that found earlier;

$$\mathbb{E}_\lambda[x] \to 0$$

   and

$$\mathbb{V}_\lambda[x] \to \infty$$

   since there is less care about the enhanced risk.

6. <u>$\lambda \to 0$ $\mathbb{E}_\lambda[x]$ T_* $\mathbb{V}_\lambda[x]$ Behavior</u>: In contrast as

$$\lambda \to \infty$$

all quantities have finite limits;

$$T_* \to \frac{\sigma}{\alpha}$$

$$\mathbb{E}_\lambda[x] \to \frac{1}{2}\eta\frac{X^2}{T_*}$$

and

$$\mathbb{V}_\lambda[x] \to \alpha\sigma X^2$$

Since trading itself introduces risk, risk-aversion and cost reduction both encourage spreading the trade over several periods; the minimum variance solution takes finite time and has finite cost.

Linear Enhanced Risk

1. Analytical Solutions for the $\alpha = 0$ Case: The next special case is

$$\alpha = 0$$

and

$$P(v) = 0$$

becomes

$$P(v) = \eta v^2 + 3\lambda\beta^2 v^4$$

and thus

$$P^{-1}(\omega) = \sqrt{\frac{\sqrt{\eta^2 + 12\lambda\beta^2\omega} - \eta}{6\lambda\beta^2}}$$

2. <u>Trading Trajectory and Characteristic Fields</u>: This can be integrated to obtain

$$\frac{t}{T_*} = F\left(\frac{X}{X_*}\right) - F\left(\frac{x}{X_*}\right)$$

in which the characteristic time and the characteristic share level are

$$T_* = \sqrt{\frac{\eta}{\lambda\sigma^2}}$$

$$X_* = \frac{1}{\sqrt{3}}\frac{\eta}{\lambda\sigma\beta} = \frac{1}{\sqrt{3}}\frac{\sigma T_*{}^2}{\beta}$$

and the nonlinear function is

$$F(u) = 2z - \coth^{-1} z$$

where

$$z = \sqrt{\frac{1}{2}\left(1 + \sqrt{1 + 4u^2}\right)}$$

3. <u>Intuition behind the Characteristic Size</u>: The characteristic time is the same as in the earlier section for

$$k = 1$$

and does not depend on the new coefficient β. To understand the characteristic level X_* note that

$$\sqrt{3}\beta \frac{X_*}{T_*} \frac{1}{\sqrt{T_*}} = \sigma\sqrt{T_*}$$

4. <u>Trading Enhanced Market Volatility Balance</u>: In this expression the left side is the trading induced variance in the share price given by the model

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau_k}\right) + \frac{1}{\sqrt{\tau_k}} f\left(\frac{n_k}{\tau_k}\right) \tilde{\xi}_k$$

$$k = 1, \cdots, N$$

if an initial portfolio of size X_* were sold in a single period T_* The right side is the variance in the share price due to the volatility in the same time interval; at the characteristic share level these two quantities are of comparable size.

5. <u>X_* much bigger than x, X</u>: To compare with the previous results note that

$$F(u) \sim \log u + constant + \mathcal{O}(u^2)$$

$$u \to 0$$

If this limit is attained by taking a limit of the *parameters* so that

$$\frac{X_*}{X} \to \infty$$

so that

$$F(u) \sim \log u + constant + \mathcal{O}(u^2)$$

is valid uniformly over x, since

$$0 \leq x \leq X$$

- a pure exponential solution results

$$\frac{t}{T_*} = \log \frac{X}{x(t)} + \mathcal{O}\left[\left(\frac{\lambda\alpha\beta}{\eta}\right)^2\right]$$

$$\frac{\lambda\alpha\beta}{\eta} \to 0$$

which, in particular, recovers the previous result with

$$k = 1$$

in the limit

$$\beta \to 0$$

6. Behavior Towards $x \to 0$; Trajectory *Tail*: And for any fixed value of the parameters

$$F(u) \sim \log u + constant + \mathcal{O}(u^2)$$

$$u \to 0$$

describes the tail of the solution as

$$x \to 0$$

the time constant of the decay is not affected by the addition of β.

7. <u>X_* much smaller than x, X</u>: For

$$x \gg X_*$$

i.e., the initial behavior when

$$X \gg X_*$$

using the expansion

$$F(u) \sim 2\sqrt{u} - \mathcal{O}\left(\frac{1}{\sqrt{u}}\right)$$

$$u \to \infty$$

gives

$$x(t) \sim X_* \left(C - \frac{1}{2}\frac{t}{T_*}\right)^2$$

$$x \gg X_*$$

with

$$C = \frac{1}{2}F(1)$$

This is the same solution constructed in the earlier Section with

$$k = 3$$

with

$$\eta = \lambda \beta^2$$

8. Almgren (2003) Asymptotic Solution Illustration:

$$\frac{t}{T_*} = F\left(\frac{X}{X_*}\right) - F\left(\frac{x}{X_*}\right)$$

together with

$$\frac{t}{T_*} = \log \frac{X}{x(t)} + \mathcal{O}\left[\left(\frac{\lambda \alpha \beta}{\eta}\right)^2\right]$$

$$\frac{\lambda \alpha \beta}{\eta} \rightarrow 0$$

and

$$x(t) \sim X_* \left(C - \frac{1}{2}\frac{t}{T_*}\right)^2$$

$$x \gg X_*$$

are illustrated in elaborate figures in Almgren (2003) Figure 5.

9. Strategy Construction Approach: Starting Trajectory: Thus the optimal strategy for construction would be as follows. Assuming

$$x > X_*$$

the initial trades are done using the trajectories of the temporary impact power law with

$$k = 3$$

with

$$\eta = \lambda\beta^2$$

That is, the volatility due to trading completely dominates the intrinsic volatility σ

10. Strategy Construction Approach: Tail Trajectory: As $x(t)$ reaches the level X_* switch is done to the optimal solution in the linear case

$$k = 1$$

with the other parameters taking their market values. In the tail trading-enhanced risk is a negligible quantity compared to the volatility.

Almgren (2003) Nonlinear Example Sample

1. Working out the $\alpha = 0$ Case: In this case the focus is on the previous section in which

$$\alpha = 0$$

and

$$\beta \neq 0$$

so that trading enhanced risk increases linearly with block size with no constant term.

2. <u>The Corresponding Discrete Price Equation</u>: To estimate the coefficients, one starts with the discrete time model. With

$$h(v) = \eta v$$

and

$$f(v) = \beta v$$

the price model

$$\tilde{S}_k = S_{k-1} - h\left(\frac{n_k}{\tau_k}\right) + \frac{1}{\sqrt{\tau_k}} f\left(\frac{n_k}{\tau_k}\right) \tilde{\xi}_k$$

$$k = 1, \cdots, N$$

becomes

$$\tilde{S}_k = S_{k-1} - \eta \frac{n_k}{\tau_k} + \beta n_k {\tau_k}^{-\frac{3}{2}} \tilde{\xi}_k$$

3. <u>Liquidity Risk as Price Volatility Fraction</u>: Assuming that for a particular choice of the trading interval τ the standard deviation of price concession associated with trading-enhanced risk is a fraction ϱ of the deterministic impact – a plausible assumption since both quantities are linearly proportional to the block size.
4. <u>The Corresponding Characteristic Price</u>: That is

$$\beta n_k {\tau_k}^{-\frac{3}{2}} = \varrho \eta \frac{n_k}{\tau_k}$$

or

$$\beta = \varrho\eta\sqrt{\tau}$$

which gives

$$X_* = \frac{1}{\sqrt{3}\varrho}\frac{1}{\lambda\sigma\sqrt{\tau}}$$

5. <u>X_* Dependence on Risk Aversion</u>: At this portfolio size, the volatility risk of holding the portfolio roughly balances the risk of selling along the optimal trajectory. Al though both of these quantities are risks, X_* involves λ through its influence on trading time T_*.
6. <u>Choice of λ, τ, ϱ</u>: The market parameters are taken to be the same as in the earlier section, with

$$\frac{1}{\lambda} = \$10{,}000$$

corresponding to the case where the liquidation is one a day. The trading is divided into one hour time intervals, so

$$\tau = \frac{2}{13}\ days$$

and

$$\varrho = \frac{1}{2}$$

7. <u>Estimate of β and X_*</u>: One obtains

$$\beta = 10^{-6}\$ \cdot day^{\frac{3}{2}} \cdot share^{-2}$$

and

$$X_* = 30{,}000\ shares$$

corresponding to a portfolio size of $\$1.5m$. A liquidation problem with initial value greater than will begin in the large x regime where trading-enhanced risk is dominant and end in the small x regime where it is negligible.

Conclusions: Summary and Extensions

1. Summary: Power Law Temporary Component: The treatment seen above obtains explicit analytical solutions for certain cases of the impact model. First, it neglects the effects of trading-enhanced risk, and takes the impact function to be a simple power law. The solutions in this case are straightforward nonlinear extensions of Almgren and Chriss (2000); the exponential solutions obtained there are a particular dividing case of these power law solutions.
2. Summary: Constant Trading-Enhanced Risk: With trading-enhanced risk, two particular cases with linear impact functions were considered. If the price uncertainty per transaction is independent of the transaction size, then the optimal trajectories are given by the previous results, simply augmenting the impact coefficient by the additional variance. A risk-averse trader lengthens his trade program, diversifying some variance away by spreading the execution over more different transactions at the expense of slightly higher volatility risk.
3. Summary: Linear Trading-Enhanced Risk: If price uncertainty per transaction is linearly proportional to the transaction size, then a characteristic portfolio size emerges, above which reduction of the added variance is the dominant effect. In this regime trade trajectories are equivalent to the previous power law solutions with exponent equal to 3. For portfolios smaller than this size the new effect may be neglected compared to the deterministic impact costs and volatility.

4. Extension #1: Optimal Numerical Trajectories: Throughout this treatment, the focus has been on obtaining explicit solutions for the sake of analytical insight. Numerical solutions would be quite straightforward, and allow lifting the restrictions described above, and consideration of a more general class of models.
5. Extension #2: Linear Impact Portfolios: Portfolios of assets are an interesting extension. Already in the linear case (Almgren and Chriss (2000)), to obtain explicit solutions it is necessary to make simplifying assumptions about cross-impacts, for example, that trading in each asset affects only the price of that asset.
6. Extension #3: Nonlinear Impact Portfolios: Even with that assumption, the nonlinear formulation opens a wide class of possible models; for example, should the exponent be the same for each asset? Determination and characterization of optimal trajectories in this case is a topic for future work.

References

- Almgren, R. F., and N. Chriss (1999): Value under Liquidation *Risk* **12 (12)** 61-63.
- Almgren, R. F., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.
- Almgren, R. F. (2003): Optimal Executions with Nonlinear Impact Functions and Trading-Enhanced Risk *Applied Mathematical Finance* **10 (1)** 1-18.
- Artzner, P., F. Delbaen, J. M. Eber, and D. Heath (1999): Coherent Measures of Risk *Mathematical Finance* **9** 203-228.
- Barra (1997): *Market Impact Model Handbook*.
- Basak, S., and A. Shapiro (2001): Value-at-risk Based Risk management: Optimal Policies and Asset Prices *Review of Financial Studies* **14** 371-405.
- Bessembinder, H., and H. M. Kaufmann (1997): A Comparison of Trade Execution Costs for NYSE and NASDAQ-Listed Stocks *Journal of Financial and Quantitative Analysis* **32** 287-310.
- Bondarenko, O. (2001): Competing Market Makers, Liquidity Provisions, and Bid-Ask Spreads *Journal of Financial Markets* **4 (3)** 269-308.

- Chakravarthy, S. (2001): Stealth Trading: Which Traders' Trades moves Prices? *Journal of Financial Economics* **61** 289-307.
- Chan, L. K. C., and J. Lakonishok (1993): Institutional Trades and Intra-day Stock Price Behavior *Journal of Financial Economy* **33** 173-199.
- Chan, L. K. C., and J. Lakonishok (1995): The Behavior of Stock Prices around Institutional Trades *Journal of Finance* **50** 1147-1174.
- Chordia, T., A. Subrahmanyam, and V. R. Anshuman (2001): Trading Activity and Expected Stock Returns *Journal of Financial Economics* **59** 3-32.
- Grinold, R. C., and R. N. Kahn (1999): *Active Portfolio Management 2nd Edition* **McGraw-Hill**.
- Hasbrouck, J., and R. A. Schwartz (1988): Liquidity and Execution Costs in Equity Markets *Journal of Portfolio Management* **14** 10-16.
- Hasbrouck, J., and D. J. Seppi (2001): Common Factors in Prices, Order Flows, and Liquidity *Journal of Financial Economics* **59** 383-411.
- Holthausen, R. W., R. W. Leftwich, and D. Mayers (1987): The Effects of Large Block Transactions on Security Prices: A Cross-Sectional Analysis *Journal of Financial Economy* **19** 237-267.
- Holthausen, R. W., R. W. Leftwich, and D. Mayers (1990): Large Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Price Effects *Journal of Financial Economy* **26** 71-95.
- Huang, R. D., and H. R. Stoll (1997): The Components of the Bid-ask Spread: A General Approach *Review of Financial Studies* **10 (4)** 995-1034.
- Huberman, G., and W. Stanzl (2005): Optimal Liquidity Trading *Review of Finance* **9 (5)** 165-200.
- Jones, C. M., and M. L. Lipson (1999): Execution Costs of Institutional Liquidity Orders *Journal of Financial Intermediation* **8** 123-140.
- Kahn, R. N. (1993): How the Execution of Trades is best Operationalized, in: *Execution Techniques, True Trading Costs, and the Microstructure of Markets*, K. F. Sherrerd (Editor) **AIMR**.

- Keim, D. B., and A. Madhavan (1995): Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Traders *Journal of Financial Economy* **37** 371-398.
- Keim, D. B., and A. Madhavan (1997): Transaction Costs and Investment Style: An Inter-exchange Analysis of Institutional Equity Trades *Journal of Financial Economy* **46** 265-292.
- Konishi, H., and N. Makimoto (2001): Optimal Slice of a Block Trade **3 (4)** 33-51.
- Koski, J. L., and R. Michaely (2000): Price, Liquidity, and the Information Content of Trades *Review of Financial Studies* **13** 659-696.
- Kraus, A., and H. R. Stoll (1972): Price Impacts of Block Trading at the New York Stock Exchange *Journal of Finance* **27** 569-588.
- Kyle, A. S. (1985): Continuous Auctions and Insider Trading *Econometrica* **53** 1315-1336.
- Loeb, T. F. (1983): Trading Costs: The Critical Link between Investment Information and Results *Financial Analysts Journal* **39** 39-44.
- Perold, A. F. (1988): The Implementation Short-fall: Paper versus Reality *Journal of Portfolio Management* **14** 4-9.
- Perold, A. F., and R. S. Salomon Jr. (1991): The Right Amount of Assets under Management *Financial Analysts Journal* **47** 31-39.
- Rickard, J. T., and N. G. Torre (1999): Information Systems for Optimal Transaction Implementation *Journal of Management Information Systems* **16** 47-62.
- Stoll, H. R. (1989): Inferring the Components of the Bid-Ask Spread: Theory and Empirical tests *Journal of Finance* **44** 115-134.
- Wagner, W. H., and M. Banks (1992): Increasing Portfolio Effectiveness via Transaction Cost Management *Journal of Portfolio Management* **19** 6-11.

Market Impact Function/Parameters Estimation

Introduction, Overview, and Background

1. Power Law Temporary Impact Function: The impact of large trades on prices is very important and widely discussed, but rarely measured. Using a large data set from a major bank and a simple but theoretical model, Almgren, Thum, Hauptmann, and Li (2005) propose that the impact is a $\frac{3}{5}$ law of the block size, with specific dependence on trade duration, daily volume, volatility, and shares outstanding.
2. Incorporation into Scheduling/Cost Estimation Algorithms: The results can be directly incorporated into an optimal trade scheduling algorithms and into pre- and post-trade estimation systems.
3. Performance Impact of Transaction Costs: Transaction costs are widely recognized as an important determinant of investment performance (see, for example, Freyre-Sanders, Guobuzaite, and Byrne (2004)). Not only do they affect the realized results of an active investment strategy, they also control how rapidly assets can be converted into cash should the need arise.
4. Direct Fixed Transaction Cost Component: Such costs generally fall into two categories. First are the direct costs such as commissions and fees that are explicitly stated and easily measured. These are important and should be minimized, but are not the focus here.
5. Indirect Controllable Transaction Cost Component: Indirect costs are not explicitly stated. For large trades the most important component of these is the impact of the traders' own actions on the market. These costs are notoriously difficult to measure but they are most amenable to careful trade management and execution.
6. Calibration of Market Impact Costs: Almgren, Thum, Hauptmann, and Li (2005) present a quantitative analysis of the market impact costs based on a large of

Citigroup US brokerage executions. A simple theoretical model that brings in the very important role of execution is used.

7. Out-of-Sample Cross Validation: The model and its calibration are constructed to satisfy two criteria. First, the predicted costs are quantitatively accurate, as determined by direct fit and out-of-sample back testing, as well as extensive consultations with the traders and the other market participants.
8. Deployability with External Execution Schedulers: The results may be directly used as an input into optimal portfolio scheduling systems, although the scheduling algorithm itself may be non-trivial.
9. Use in Citigroup's BECS System: The results of this study have been incorporated into Citigroup's Best Execution Consulting Services (BECS) software for use internally at all desks as well as the clients of the equity division. While this work has focused on US markets, it has been extended to global equities. BECS is the delivery platform for Citigroup's next generation of trading analytic tools, both pre- and post-execution.
10. Extension to the Standard Trading Model: This pre-trade analyzer is an extension to the market standard existing model that has been delivered through the Stock Facts Pro software for the past 25 years (Sorensen, Price, Miller, Cox, and Birnbaum (1998)).
11. Solid Empiricals and Real-Data Verifications: This model is based on better developed empirical foundations; it is based on real trading data taking time into consideration while verifying the results through post-trade analysis. The table below summarizes some of the advantages/disadvantages of this approach.
12. Distinguishing Features of the Model:
    a. Advantages
        i. Calibrated from Real Data
        ii. Includes Time Component
        iii. Incorporates Intra-day Profiles
        iv. Uses non-linear Impact Functions
        v. Confidence Levels for Coefficients
    b. Disadvantages

i. Based only on Citigroup Data
ii. Little Data for Small-Cap Stocks
iii. Little Data for very large Trades

13. Academic Industrial Market Data Quantification: Much work in both the academic and industrial communities has been devoted to understanding and quantifying market impact costs. Many academic studies have only worked with publically available data in Trade and Quote (TAQ) tick-record from the New York Stock Exchange (NYSE).
14. Buy-Sell Market Imbalance Analysis: Breen, Hodrick, and Korajczyk (2002) regress the net markets movements over five-minute and half-hour time periods against a net buy-sell impact during the same period, using a linear impact model. A similar model is developed in Kissell and Glantz (2003).
15. Impact Cost Function Dependence Analysis: Rydberg and Shephard (2003) develop a rich econometric foundation for describing price motions; Dufour and Engle (2000) investigate the key role of waiting period between successive trades. Using techniques from statistical physics, Lillo, Farmer, and Mantegna (2003) look for a power law scaling in the impact cost function, and find significant dependence on total market capitalization as well as daily volume, and Bouchaud, Gefen, Potters, and Wyart (2004) discover non-trivial serial correlation ion volume and price data.
16. Limitations of Public Data Sets: The publically available data sets lack the reliable classification of individual trades as buyer- or seller- initiated. Even more significantly, each transaction exists in isolation; there is no information on the sequence of trades that form part of the larger transaction.
17. Transaction Trade Sequence Incorporation Studies: Some academic studies have used limited data sets made available by asset managers that do have this information, where the date, but not the time duration of the trade is known (Holthausen, Leftwich, and Mayers (1990), Chan and Lakonishok (1995), and Keim and Madhavan (1996)).
18. Permanent and Temporary Impact Costs: The transaction cost model embedded in this analysis is based on the model presented by Almgren and Chriss (2000) with non-linear extensions from Almgren (2003). The essential features of this model, as described below, is that it explicitly divides the market impact costs into a permanent

component associated with information and a temporary component arising from the liquidity demands made by the execution in a short time.

Data Description and Filtering Rules

1. Data Generation Period and Universe: The data on which the analysis was based on contains, before filtering, almost 700,000 US stock orders executed by the Citigroup equity trading desks for a 19-month period from December 2001 to June 2003.
2. Orders, Transactions, and Resulting Executions: Each order is broken down into one or more transactions, each of which may generate one or more executions. The information presented below is available for each order.
3. Symbol, Size, and Order Type: The stock symbol, the requested order size, and the sign (buy/sell) of the entire order – the client information is removed.
4. Order Submission Time and Method: The times and the methods by which the transaction was submitted by the Citigroup trader to the market. The time t_0 of the first transaction is taken to be the start of the order. Some of the transactions are sent as market orders, some as limit order, and some are submitted to the Citigroup's automated VWAP server. Except for the starting time and except to exclude the VWAP orders, no use is made of this transaction information.
5. Execution Times, Sizes, and Prices: The times, the prices, and the sizes corresponding to the execution of each transaction is used. Some transactions are cancelled or inly partly executed; only completed sizes and prices are used. The execution times re denoted by $t_1, \cdots, t_n$, sizes by $x_1, \cdots, x_n$, and prices by $S_1, \cdots, S_n$.
6. Order Completion Finish Time Frame: All orders are completed within one day, though not necessarily filled.
7. Additional Types of Information Available: In addition, various additional pieces of information – such as instructions given by the client to the trader for the order – e.g., 'market on close', 'market on open', 'over the day', VWAP or blank – are available.
8. Sample Subset Filtering Criteria: the total sample contains 682,582 orders, but only a subset is used for the data analytics.

9. S & P 500 Constituent Stocks Only: To exclude small and thinly traded stocks, only orders on the Standard and Poor's Index are considered, which represent about half of the total number of orders, but a large majority of the dollar value. Even within this universe there is enough diversity to explore the dependence on the market capitalization, as there are both NYSE and OTC stocks.
10. Exclusion of Highly Volatile Stocks: Also excluded are approximately 400 orders for which the stocks exhibit more than 12.5% daily volatility – 200% annual volatility.
11. Orders that Match the Analysis Objectives: Furthermore, only those orders that are reasonably representative of the actual scheduling strategies that are the ultimate goal are considered.
12. Excluding Market-on-Open/Close: The orders for which the client requested market-on-open or market-on-close executions are excluded. These orders are likely to be executed with strongly non-linear profiles that do not satisfy the modeling assumption – there are only a few hundred of these.
13. Excluding Client Requested VWAP Orders: The orders for which the client requested VWAP execution are excluded. These orders have consistently long execution times and represent very small rates of trading relative to the market volume. These are about 16% of the total number of orders.
14. Excluding Later-in-the-Day Transactions: Also excluded are orders for which any executions are recorded after 16:10 EST, approximately 10% of the total. In many cases these use Citigroup's block desk for some or all of the transactions, and the fills are reported sometime after the order is completed. Therefore the time information is not reliable.
15. Caveat - Impact of Orders Exclusion: This exclusion, together with the use of fill size in place of originally requested size, could be a source of significant bias. For example, if clients and traders consistently used limit orders, orders may be filled only if the prices moved in a favorable direction. Analysis of the data set suggests that this effect is not significant – for example the same coefficients are obtained with or without partially filled orders – and informal discussions with the traders confirm the belief that partial fills are not the result of a limit order strategy.

16. Other Minimum Cut off Criteria: Most significantly, small order are excluded since the goal is to estimate transaction costs in the range where they are significant. Specifically only the following orders are included:
    a. The order has at least two completed transactions.
    b. Orders are at least 1,000 shares.
    c. Orders are at least 0.25% of average daily volume in that stock.
17. Range of Execution Sizes/Times: The results of the model are reasonably stable to changes in the above criteria. After this filtering there are 29,509 orders in the data set; the largest number of executions for any order is 548, and the median is around 5. The median time is around 5 minutes.
18. Order/Volume Ratio Range: The table below shows some descriptive statistics for the sample. Most of the orders constitute only a few percent of the typical market volume, and the model is designed to work within this range of values. Orders greater than a few percent of daily volume have substantial sources of uncertainty that are not modeled here, and the model does not represent them.
19. Summary Statistics of the Sample Orders:

	Mean	**Minimum**	**Q1**	**Median**	**Q3**	**Maximum**
Total Cost %	0.04	-3.74	-0.11	0.03	0.19	3.55
Permanent Cost % I	0.01	-3.95	-0.17	0.01	0.19	2.66
Temporary Cost % J	0.03	-3.57	-0.11	0.02	0.17	2.33
Shares/ADV % $\lvert X \rvert$	1.51	0.25	0.38	0.62	1.36	88.62
Time Days	0.39	0.00	0.10	0.32	0.65	1.01
Daily Volatility %	2.68	0.70	1.70	2.20	3.00	12.50
Mean Spread %	0.14	0.03	0.03	0.11	0.16	2.37

Data Model - Variables

1. Market Impact Input Dependence Estimation: The goal of the study is to determine the market impact in terms of a small number of input variables. Below is a list of

precisely which market impacts are measured, and what primary and auxiliary variables will be used to model them.

2. Pre- and Post- Market Prices: Let $S(t)$ be the price of the asset being traded. For each order the following price points of interest are defined: S_0 is the market price before this order begins executing; S_{POST} is the market price after this order is completed; and $\bar{S}$ is the average realized price on the order.
3. The Transaction Weighted Average Price: The realized price

$$\bar{S} = \frac{\sum_{j=1}^{N} x_j S_j}{\sum_{j=1}^{N} x_j}$$

is calculated from the transaction data set. The market price S_0 and S_{POST} are the bid-ask mid points from TAQ.

4. First Transaction Pre- Trade Price: The pre-trade price S_0 is the price before the impact on the trade begins to be felt (this is an approximation, since some information may leak before any record enters the system). S_0 is computed from the latest quote just preceding the first transaction.
5. Post-trade Price Capture - Caveat: The post-trade price S_{POST} should capture the permanent effects of the trade program. That is, it should be long enough after the last execution that any effects of temporary liquidity have dissipated.
6. Accounting for the Permanent Impact: In reportedly performing the fits, Almgren, Thum, Hauptmann, and Li (2005) have found that 30 minutes after the last execution is enough to achieve this. For shorter time intervals, the regressed values depend on the time lag, and about this level the variation stops. That is, they define

$$t_{POST} = t_n + 30\ minutes$$

7. t_{POST} Delay Date Roll Over: The price S_{POST} is taken from the first quote following t_{POST}. If t_{POST} is after the market close, it carries over to the next morning. This risks distorting the results by including excessive overnight volatility, but Almgren, Thum,

Hauptmann, and Li (2005) have found this to give more consistent results than truncating at the market close.

8. Permanent Realized Impact Variables Definition: Based on these prices the following dimensionless impact variables are defined. The dimensionless permanent impact is

$$I = \frac{S_{POST} - S_0}{S_0}$$

and the dimensionless realized impact is

$$I = \frac{\bar{S} - S_0}{S_0}$$

9. Conversion into Observed Market Impacts: The “effective dimensionless impact J is the quantity of most interest, since it specifies the actual cash spent or received on the trade. In the model below the temporary impact will be defined to be J minus a suitable fraction of I and this temporary impact will be the quantity described by the theory.
10. Signs of the Impact Variables: On any individual order, the signs of I, J can be positive or negative. In fact since volatility is a very large contributor to either values, they are almost likely to have either sign. They are defined so that positive cost is experienced if I, J have the same sign as the total order X; for a buy order with

$$X > 0$$

positive cost means that the price $S(t)$ moves upwards. The average values of I, J taken across many orders is expected to have the same sign as X.
11. Intra-day Volume Weighted Time: The level of market activity is known to vary substantially and consistently over different periods of the trading day; this intra-day variation affects both the volume profile and the variance of prices. To capture this

effect, all computations are performed in volume time τ which represents the fraction of the average day's volume that has executed up to the time t.

12. Intra-day Volume Weighted Trajectory: Thus a constant rate trajectory in the τ variable corresponds to a VWAP execution in real time. The relationship between t and τ is independent of the daily trading volume; it is scaled so that

$$\tau = 0$$

at market open and

$$\tau = 1$$

at market close.

13. Intra-day Volume Weighted Times: Each of the clock times $\tau_0, \cdots, \tau_n$ in the data set is mapped to the corresponding volume time $t_0, \cdots, t_n$.Since the stocks in the sample are heavily traded a non-parametric estimator that directly measures the differences in τ is used; the shares traded during the period correspond to the execution of each order.

14. Time Volume vs. Price Volatility: Almgren, Thum, Hauptmann, and Li (2005) display an illustration of the empirical profiles. The fluctuations in each time period in these illustrations correspond to the approximate size of statistical error in the volume calculation for a 15-minute trade; these errors are typically less than 5%, and are smaller for longer periods.

15. The "Dimensional" Parametric Explanatory Variables: The impacts I and J are to be described in terms of the following quantities.
    a. Total executed size in shares

$$X = \sum_{j=1}^{N} x_j$$

    b. Volume Duration of Active Trading:

$$T = \tau_n - \tau_0$$

c. Volume Duration of the Impact:

$$T_{POST} = \tau_{POST} - \tau_0$$

16. Caveats around the Explanatory Variables: As noted above, X is positive for a buy order, and negative for a sell order. Explored defining T using a size weighted average of execution times, but the results are not substantially different. The intermediate execution times $\tau_1, \cdots, \tau_{n-1}$ are not used, and the execution sizes are not used either except in calculating the order size and the mean realized prices.
17. Fixed Trade - Optimal Time Nodes: In the eventual application for trajectory optimization, the size X will be assumed given, and the execution schedule here represented by T will be optimized.
18. Execution Time as Optimizing Parameter: In general the solution will be a complicated time dependent trajectory parametrized by a time scale T. For the purposes of data modeling the trajectory optimization is ignored and the schedules are taken to be determined only by a single number T.
19. Market Core Empirical Parametric Inputs: Although the goal is to explain the dependence of the impact costs I, J on order size X and trade time T, other market variables will influence the solution. The most important of these are: V – which is the average daily volume in shares, and σ – the daily volatility.
20. Daily Volume/Volatility "Wander" Scale: V is a 10 day moving average. For volatility, an intra-day estimator that makes use of every transaction in the day is used. It is important to track changes in these variables not only between different stocks but also across time for the same stock.
21. Order Size/Daily Volume Normalization: These values serve primarily to "normalize" the active variables across the stocks with widely varying properties. It seems natural that order size X should be measured as a fraction of the average daily volume V: $\frac{X}{V}$ is a more natural variable than V itself.

22. Intrinsic Notion of Volume Time: In the model presented below, the order size as a fraction of the average volume traded during the time of execution will also be seen to be important. VT is estimated directly by taking the average volume that executed between the times t_0 and t_n over the previous 10 days. In fact, since in the model the trade duration T appears only in the combination VT this avoids the need to measure T directly.
23. "Wander" Scale of Market Impact: The volatility is used to scale the impacts – a certain level of participation in the daily volume should cause a certain level of participation in the 'normal' motion of the stock. Empirical investigation by Almgren, Thum, Hauptmann, and Li (2005) shows that volatility is the most important scale factor for cost impact.

Trajectory Cost Model

1. Constant Volume Time Trading Rate: The model used is based on the framework developed by Almgren and Chriss (2000), and Almgren (2003), with simplifications made to facilitate data fitting. The main simplification is that the rate of trading is constant (in volume time). In addition cross impact is neglected, since the data has no information about the effect of trading one stock on the price of the other.
2. The Permanent Impact Market Component: The price impact is decomposed into two components. First is a permanent component that reflects the information transmitted to the market by the buy/sell imbalance. This component is believed to be roughly independent of trade scheduling; 'stealth' trading is not admitted by this construction. In the data fit this component will be independent of the execution time T.
3. The Temporary Market Impact Component: A temporary component reflects the price concession needed to attract counterparties within a specified short time interval. This component is highly sensitive to trade scheduling; here it will strongly depend on T.
4. Other Elaborate Market Impact Frameworks: More detailed conceptual frameworks have been developed (Bouchaud, Geffen, Potters, and Wyart (2004)), but this easily

understood model has become standard in industry and academic literature (Madhavan (2000)).

5. Decomposition of the Realized Market Impact: The realized price impact is a combination of the above two effects. In terms of the realized and the permanent impact defined above and observed from the data, the model may be summarized as

$$Realized = Permanent + Temporary + Noise$$

with suitable coefficients and scaling depending upon T. Thus the temporary impact is obtained as a difference between the permanent impact and the realized impact; it is not directly observed, although there is a direct model for it.

6. Uniform Rate of Order Liquidation: The starting point is the initial order demand of X shares. This is assumed to be completed by a uniform rate of trading over a volume interval T. That is, the trade rate in volume units is

$$v = \frac{X}{T}$$

and is held constant until the program is completed.

7. Sign of the Trade Rate: Constant rate in these units is equivalent to VWAP execution during the time of execution. Note that v has the same sign as X; thus

$$v > 0$$

for a buy order and

$$v < 0$$

for a sell order. Market impact will move the price in the same direction as v.

Permanent Impact

1. Volatility/Permanent Impact Price Change: The model postulates that the asset price $S(\tau)$ follows an arithmetic Brownian motion with a drift term that depends on the trade rate term v. That is

$$\Delta S = S_0 g(v)\Delta\tau + S_0\sigma\Delta B(\tau)$$

where $B(\tau)$ is a standard Brownian motion (or a Bachelier process); and $g(v)$ is the permanent impact function; the only assumptions made are that $g(v)$ is increasing and has

$$g(0) = 0$$

2. Integrated Form of Price Change: As noted above, τ is volume time, representing the fraction of an average day's volume that has executed so far. This expression can be integrated in time taking v to equal $\frac{X}{T}$ for

$$0 \le \tau \le T$$

to obtain the permanent impact

$$I = Tg\left(\frac{X}{T}\right) + \sigma\sqrt{T_{POST}}\xi$$

where

$$\xi \sim \mathcal{N}(0, 1)$$

is a standard Gaussian variable.

3. Linearity of the Permanent Impact Function: Note that if $g(v)$ is a linear function, then the accumulated drift at time τ is equal to $\frac{X\tau}{T}$, the number of shares executed to time τ, and the permanent impact I is proportional to the total order size X independently of the time scale T.

Temporary Impact

1. Temporary Impact Price Change Realization: The actual price received from the trade is

$$\tilde{S}(\tau) = S(\tau) + S_0 h\left(\frac{X}{T}\right)$$

where $h(v)$ is the temporary impact function. For convenience, it has been scaled by the market price at the start of trading, since the time intervals involved are all less than one day.

2. Discretization of the Price Impact: This expression is a continuous time approximation to a discrete process. A more accurate description would be to imagine that the time intervals would be broken down into intervals such as, say, one hour or 30 minute intervals. Within each interval the average price realized on the trade during that interval would be less favorable that the average price that an unbiased observer would measure during that time interval.
3. Unbiased Price Plus Liquidity Concession: The unbiased price is affected by the previous trades that have been executed before this interval (as well as the volatility) but not on their timing. The additional concession during this time interval is strongly dependent on the number of shares executed in this interval.
4. Closed Form Temporary Impact Expression: At a constant liquidation rate, calculating the time average of the execution price gives the temporary impact expression

$$J - \frac{I}{2} = h\left(\frac{X}{T}\right) + \sigma\left[\sqrt{\frac{T}{12}\left(4 - 3\frac{T}{T_{POST}}\right)}\chi - \frac{T_{POST} - T}{2\sqrt{T_{POST}}}\xi\right]$$

where

$$\chi \sim \mathcal{N}(0, 1)$$

is independent of ξ. The term $\frac{I}{2}$ reflects the effect on the later execution prices of permanent impact caused by the earlier parts of the program.

5. Estimate of the Heteroscedastic Corrections: The rather complicated error expression reflects the fluctuations on the middle part of the Brownian motion on $[0, T]$ relative to the end point at T_{POST}. It is only used for the heteroscedastic corrections for the regression fits below.
6. Fluctuations and Error Residuals Estimation: The equations and provide explicit expressions for the permanent and the temporary impact components I, J in terms of the values of the functions g and h at known trade rates, together with the estimates of the magnitude of the error coming from the volatility.
7. Regression Based Impact Form Estimation: The data fitting procedure above is in principle straightforward, the impacts I and J are computed from the transaction data, and those values are regressed against order sizes and times as indicated to directly extract the functions $g(v)$ and $h(v)$.

Choice of the Functional Form

1. Permanent/Temporary Impact Function Structure: The next question that needs to be addressed is what should the structure of the permanent impact function $g(v)$ and the temporary impact function $h(v)$ be. Even with a large sample it is not possible to extract these functions purely from data, so a hypothesis must be made about their structure.

2. Power Law Impact Functional Forms: The postulate is that these functions are power laws, that is, that:

$$g(v) = \pm\gamma|v|^{\alpha}$$

and

$$h(v) = \pm\eta|v|^{\beta}$$

where the numerical values of the dimensionless coefficients γ and η and the exponents α and β are to be determined by linear and non-linear regressions on the data. The sign is to be chosen so that $g(v)$ and $h(v)$ have the same sign as v.

3. Range of Power Law Representation: This class of power law is extremely broad. It includes concave functions (exponent < 1), convex functions (exponent > 1), and linear functions (exponent = 1). It is the functional form that is implicitly assumed by fitting straight lines on a log-log plot as is very common in physics, and has been used in this context, for example, by Lillo, Farmer, and Mantegna (2003).
4. Order Type/Exchange Parameter Independence: The same coefficients are taken for buy orders

$$v > 0$$

and sell orders

$$v < 0$$

It would be a trivial modification to introduce different coefficients $\gamma_{\pm}$ and $\eta_{\pm}$ for the two sides, but the exploratory analysis by Almgren, Thum, Hauptmann, and Li (2005) has not indicated a string need for this. Similarly it would be possible to use different coefficients for stocks traded on different exchanges, but this does not appear to be necessary either.

5. Quasi Arbitrage Permanent Impact Elimination: There is reason to be specific in the choice of the exponents. For the permanent impact function there is a strong reason to prefer the linear model with

$$\alpha = 1$$

   This is the only value for which the model is free from quasi-arbitrage (Huberman and Stanzl (2004)).
6. Linearity of the Permanent Impact: Furthermore, the linear function is the only one for which the permanent price impact is independent of the trading time. Of course this substantial conceptual simplification must be supported by the data.
7. Concave Nature of the Temporary Exponents: For temporary impacts, there is ample evidence indicating that the function should be concave, that is

$$0 < \beta < 1$$

   This evidence dates back to Loeb (1983) and is strongly demonstrated by the fits in Lillo, Farmer, and Mantegna (2003). In particular theoretical arguments (Barra (1997)) suggest that the particular value of

$$\beta = \frac{1}{2}$$

   is especially plausible, resulting in a square root impact function.
8. Verification of Power Exponent Values: The approach is then as follows. Unprejudiced fits to the power law functions shall be made to the entire data set to determine the best estimates for the exponents α and β. The validity of the values

$$\alpha = 1$$

   and

$$\beta = \frac{1}{2}$$

will then be tested to validate the linear and the square root candidate functional forms.

9. Determination of the Impact Coefficients: Once the exponents have been selected, simple linear regression is adequate to determine the coefficients. In this regression heteroscedastic weightings are used, with the error magnitudes from

$$I = Tg\left(\frac{X}{T}\right) + \sigma\sqrt{T_{POST}}\xi$$

and

$$J - \frac{I}{2} = h\left(\frac{X}{T}\right) + \sigma\left[\sqrt{\frac{T}{12}\left(4 - 3\frac{T}{T_{POST}}\right)}\chi - \frac{T_{POST} - T}{2\sqrt{T_{POST}}}\xi\right]$$

The result of this regression is not only the values for the coefficients, but also a collection of the error residuals ξ and χ which must be tested for normality as the theory supposes.

Cross-Sectional Description

1. Motivations for the Properties Normalization: The above analysis has assumed an 'ideal' asset, all of whose properties remain constant in time. For any real asset, the parameters that determine the market impact will vary with time. For example one would that the execution of a given number of shares would incur higher impact costs on a day with unusually low volume or unusually high volatility.

2. Basis for the Normalizer Choice: That is, the impact of the cost functions should be expressed in terms of the dimensionless quantity $\frac{X}{VT}$ rather than X itself, where V is the average number of shares per day defined above.
3. Normalization of the Price Moves: Furthermore, the motion of the price should not be given as a raw percentage figure, but it should be expressed as a fraction of 'normal' daily motion of the price, as expressed by the volatility σ.
4. Normalized Expressions for I, J: With these assumptions, the equations

$$I = Tg\left(\frac{X}{T}\right) + \sigma\sqrt{T_{POST}}\xi$$

and

$$J - \frac{I}{2} = h\left(\frac{X}{T}\right) + \sigma\left[\sqrt{\frac{T}{12}\left(4 - 3\frac{T}{T_{POST}}\right)}\chi - \frac{T_{POST} - T}{2\sqrt{T_{POST}}}\xi\right]$$

can be modified to

$$I = \sigma Tg\left(\frac{X}{T}\right) + [\![noise]\!]$$

and

$$J - \frac{I}{2} = \sigma h\left(\frac{X}{T}\right) + [\![noise]\!]$$

respectively, where $[\![noise]\!]$ is the error expression depending on the volatility.
5. Dimensionless Permanent Temporary Function Inputs: Now g and h are dimensionless functions of a dimensionless variable. They are assumed to be constant in time for a single stock across days when σ and V vary. The next step is to investigate these functions for their dependence on cross-stock variables.

Model Determination

1. Specification of Extraneous Model Regressors: To bring the full size of the data into play, one must address the more complex and the less precise question of how the impact functions vary across the stocks, that is, how much they depend variables such as market capitalization, shares outstanding, bid-ask spread, or other quantities. Temporary and permanent impact must be considered separately.
2. Permanent Impact Function Liquidity Regressor: A 'liquidity factor' $\mathcal{L}$ is inserted into the permanent cost function $g(v)$, where $\mathcal{L}$ depends on the market parameters characterizing each stock (in addition to daily volume and liquidity). There are several candidates for inputs into $\mathcal{L}$.
3. Liquidity Regressor Candidate - Inverse Turnover: The form of $\mathcal{L}$ is constrained to be

$$\mathcal{L} = \left[\frac{\Theta}{V}\right]^{\delta}$$

   where Θ is the total number of shares outstanding, and δ is the exponent to be determined. The dimensionless quantity $\frac{\Theta}{V}$ is the inverse of the 'turnover' – the fraction of the company's value traded each day. This is a natural explanatory variable, and has been used in empirical studies such as Breen, Hodrick, and Korajczyk (2002).
4. Liquidity Regressor Candidate – Bid-Ask: Almgren, Thum, Hauptmann, and Li (2005) did not find any consistent dependence on the bid-ask spread across the sample, so it is not included in $\mathcal{L}$.
5. Liquidity Regressor Candidate - Market Capitalization: This differs from the shares outstanding by the price per share, so including this factor is equivalent to including a 'price effect'. The study by Almgren, Thum, Hauptmann, and Li (2005) found that there is a persistent price effect, as also found by Lillo, Farmer, and Mantegna (2003),

but that the dependence is weak is enough that it may be neglected in favor of the conceptually simpler quantity $\frac{\Theta}{V}$.

6. Temporary Impact Function - Regressor Candidate: In further extensive preliminary exploration, it was found that the temporary cost function $h(v)$ does not require any stock-specific modification; liquidity costs as a fraction of volatility only depends upon the fraction of the shares traded as a fraction of the average daily volume.
7. Revised I, J Functional Forms: After assuming the functional form defined above, the model is validated and the exponent δ is determined by performing a non-linear regression of the form

$$\frac{I}{\sigma} = \gamma T sgn(X) \left|\frac{X}{VT}\right|^{\alpha} \left[\frac{\Theta}{V}\right]^{\delta} + [\![noise]\!]$$

and

$$\frac{1}{\sigma}\left[J - \frac{I}{2}\right] = \eta sgn(X) \left|\frac{X}{VT}\right|^{\beta} + [\![noise]\!]$$

where $[\![noise]\!]$ is the again the heteroscedastic error term from

$$I = Tg\left(\frac{X}{T}\right) + \sigma\sqrt{T_{POST}}\xi$$

and sgn is the sign function.
8. Estimates and Residuals of Exponents: A modified Gauss-Newton optimization algorithm was used to determine the values of α, β, and δ that minimized the normalized residuals. The results are:

$$\alpha = 0.891 \pm 0.10$$

$$\delta = 0.267 \pm 0.22$$

$$\beta = 0.600 \pm 0.038$$

9. Errors Represented as One Sigma Amounts: Here, as throughout this chapter, the error bars are expressed with $\pm$ are one standard deviation, assuming Gaussian error model. Thus the 'true' value can be expected to be within this range with 67% probability, and within a range twice as large with 95% probability.
10. Choice of Linear Permanent Impact: From these values the following conclusions can be drawn. First the value

$$\alpha = 1$$

for linear impact cannot be reliably rejected. In view of enormous practical simplification of the linear permanent impact

$$\alpha = 1$$

is chosen.
11. Permanent Impact Liquidity Exponent Estimation: The liquidity factor is very approximately

$$\delta = \frac{1}{4}$$

12. Temporary Impact Power Law Exponent: For temporary impact, the analysis confirms the concavity of the function with β strictly less than one. This confirms the fact that the bigger the trades made by the fund managers on the market, the less additional cost they experience per share traded. At 95% confidence level, the square root model

$$\beta = \frac{1}{2}$$

is rejected. The temporary cost exponent is therefore fixed at

$$\beta = \frac{3}{5}$$

In comparison with the square root model, this gives slightly smaller costs for smaller trades, and slightly larger costs for large trades.

13. <u>Permanent Dependence on Shares Outstanding</u>: Note that because

$$\delta > 0$$

for fixed values of number X of shares in the order, and the average daily volume V, the cost increases with Θ, the total number of shares outstanding. In effect a large number of outstanding shares means that a smaller fraction of the company is traded each day, so a given fraction of that flow has a greater impact.

14. <u>Linear Permanent Concave Temporary Impact</u>: Therefore the results confirm empirically the theoretical arguments of Huberman and Stanzl (2004) for permanent impact that is linear in the block size, and the concavity of the temporary impact has been widely described in the literature for both theoretical and empirical reasons.

Determination of the Coefficients

1. <u>Estimation of the Impact Coefficients</u>: After fixing the exponent values, the values of γ and η are determined by linear regression of the models.

$$\frac{I}{\sigma} = \gamma T sgn(X) \left|\frac{\mathrm{X}}{VT}\right|^{\alpha} \left[\frac{\Theta}{V}\right]^{\delta} + [\![noise]\!]$$

and

$$\frac{1}{\sigma}\left[J-\frac{I}{2}\right]=\eta sgn(X)\left|\frac{\mathrm{X}}{VT}\right|^{\beta}+[\![noise]\!]$$

using the heteroscedastic error estimates given in

$$I=Tg\left(\frac{X}{T}\right)+\sigma\sqrt{T_{POST}}\xi$$

and

$$J-\frac{I}{2}=h\left(\frac{X}{T}\right)+\sigma\left[\sqrt{\frac{T}{12}\left(4-3\frac{T}{T_{POST}}\right)}\chi-\frac{T_{POST}-T}{2\sqrt{T_{POST}}}\xi\right]$$

It is found that

$$\gamma=0.314\pm0.041$$

with

$$t=7.7$$

and

$$\eta=0.142\pm0.0062$$

with

$$t=23$$

2. Interpretation of the t statistic: The t statistic is calculated assuming that the Gaussian model expressed in

$$I = Tg\left(\frac{X}{T}\right) + \sigma\sqrt{T_{POST}}\xi$$

and

$$J - \frac{I}{2} = h\left(\frac{X}{T}\right) + \sigma\left[\sqrt{\frac{T}{12}\left(4 - 3\frac{T}{T_{POST}}\right)}\chi - \frac{T_{POST} - T}{2\sqrt{T_{POST}}}\xi\right]$$

is valid; the error estimates are the values divided by the t statistic. Although the actual residuals are fat-tailed as discussed below, these estimates indicate that the coefficient values are highly significant.
3. Permanent Impact Signal Contribution: The $\mathbb{R}^2$ values are typically less than 1% indicating that only a small part of the dependent variables I and J is explained by the model in terms of the independent variables. This is precisely what is expected given the small size of the random impact term relative to the random motion of the price due to the volatility arising from the trade execution.
4. Importance of the Permanent Cost: The permanent persistent cost, though small, is of major importance since it is on the average the cost incurred by the fund managers while trading. Furthermore since most orders are part of large portfolio trades, the volatility costs experienced on the portfolio level is considerably lower than exhibited in the stock-level analysis, increasing the significance of the fraction of the impact cost estimated. As previously mentioned the non-linear optimization of the volatility versus impact cost trade-off would reveal additional profitable strategies.
5. Universal Coefficients of Market Impact: The dimensionless numbers γ and η are the universal coefficients of market impact. According to the model, they apply to every order and every asset in the entire data set.
6. Interpretation Caveat I vs. J: To summarize, they are to be inserted into the equations

$$I = \gamma\sigma \frac{X}{V} \left[\frac{\Theta}{V}\right]^{\frac{1}{4}} + [\![noise]\!]$$

and

$$J = \frac{I}{2} + \eta\sigma sgn(X) \left|\frac{X}{VT}\right|^{\frac{3}{5}} + [\![noise]\!]$$

giving the expectation of the impact costs; in any particular order the realized values will vary greatly due to the volatility. To reiterate, I does not signify the total cost, but is simply the net price motion from pre-trade to post-trade. The actual cost experienced by the trade is signified by J.

7. Sub Group Impact Parameters Determination: Almgren, Thum, Hauptmann, and Li (2005) have chosen the simple form above to have a single model that applies reasonably well across the entire data set which consists entirely of large cap stocks in the US market. More detailed models could be constructed to capture more limited data or assets, or to account for variations across global markets. In practice, it is expected that the coefficients, perhaps even the exponents, or maybe even the functional forms, will be continually updated to reflect the most recent data.
8. Example of Impact Cost: The table below shows the impact cost functions and the numerical examples for two large cap stocks when the customer buys 10% of the average daily volume. The permanent cost is independent of the time of execution, the temporary cost depends on the time of execution, but across different asset it is the same fraction of the daily volatility. K is written as

$$K = J - \frac{I}{2}$$

9. Example of Impact Costs Table:

			IBM	**DRI**

Average Daily Volume	Million	V	6.561			1.929		
Shares Outstanding	Million	Θ	1728			168		
Inverse Turnover		$\frac{\Theta}{V}$	263			87		
Daily Volatility	%	σ	1.57			2.26		
Normalized Trade Rate		$\frac{X}{V}$	0.1			0.1		
Normalized Permanent Impact		$\frac{I}{\sigma}$	0.126			0.096		
Permanent Price Impact	bp	I	20			22		
Trade Duration	Days	T	0.1	0.2	0.5	0.1	0.2	0.5
Normalized Temporary Impact		$\frac{K}{\sigma}$	0.142	0.094	0.054	0.142	0.094	0.054
Temporary Impact Cost	bp	K	22	15	8	32	21	12
Realized Cost	bp	J	32	25	18	43	32	23

10. Example of Impact Cost – Analysis: In the above example, because DRI turns over $\frac{1}{87}$ of its float each day, whereas IBM turns over only $\frac{1}{263}$, trading 10% of the day's volume causes a permanent price move of only 0.1 times volatility for DRI, but 0.13 times for IBM; half of this is experienced as cost. Because the permanent cost function is linear, the permanent cost numbers are independent of the times of execution.

Residual Analysis

1. Results of Impact and Market: The results of the above analysis are not simply the values of the coefficients presented. In addition, the error formulation provides specific predictions for the nature of the residuals ξ and χ for the permanent and the temporary impact respectively from

$$I = Tg\left(\frac{X}{T}\right) + \sigma\sqrt{T_{POST}}\xi$$

and

$$J - \frac{I}{2} = h\left(\frac{X}{T}\right) + \sigma\left[\sqrt{\frac{T}{12}\left(4 - 3\frac{T}{T_{POST}}\right)}\chi - \frac{T_{POST} - T}{2\sqrt{T_{POST}}}\xi\right]$$

2. Validating the Independence of the "Wanderers": Under the assumption that the asset price is a Brownian motion with the drift caused by the impact, these two variables should be independent standard Gaussians. This assumption has already been used in heteroscedastic regression, now it needs to be verified.
3. Residual Mean and Covariance: Almgren, Thum, Hauptmann, and Li (2005) demonstrate histograms and $Q - Q$ plots of ξ and χ. The means are quiet close to zero, the variances a r reasonably close to 1, and the correlation is reasonably small.
4. Fat-Tailed Nature of the Distribution: But the distribution is extremely fat-tailed, as is normal for returns distributions on short-time intervals (see Rydberg (2000)), and hence does not indicate that the model is poorly specified. Nonetheless, the structure of the residuals confirms that the model is close to the best that can be obtained within the Brownian framework.

References

- Almgren, R. F., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.
- Almgren, R. F. (2003): Optimal Executions with Nonlinear Impact Functions and Trading-Enhanced Risk *Applied Mathematical Finance* **10 (1)** 1-18.
- Almgren, R. F., C. Thum, E. Hauptmann, and H. Li (2005): Equity Market Impact *Risk* **18 (7)** 57-62.
- Barra (1997): *The Market Impact Handbook.*

- Bouchaud, J. P., Y. Gefen, M. Potters, and M. Wyart (2004): Fluctuations and Responses in Financial Markets: The "Subtle" Nature of Random Price Changes *Quantitative Finance* **4 (2)** 176-190.
- Breen, W., L. Hodrick, and R. Korajczyk (2002): Predicting Equity Liquidity *Management Science* **48 (4)** 470-483.
- Chan, L. K. C., and J. Lakonishok (1995): The Behavior of Stock Prices around Institutional Trades *Journal of Finance* **50** 1147-1174.
- Dufour, A. and R. Engle (2000): Time and the Price Impact of a Trade *Journal of Finance* **55 (6)** 2467-2498.
- Freyre-Sanders, A., R. Guobuzaite, and K. Byrne (2004): A Review of Trading Cost Models *Journal of Investing* **13** 93-115.
- Holthausen, R. W., R. W. Leftwich, and D. Mayers (1990): Large Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Price Effects *Journal of Financial Economy* **26** 71-95.
- Huberman, G., and W. Stanzl (2004): Price Manipulation and Quasi-arbitrage *Econometrica* **72 (4)** 1247-1275.
- Keim, D. B., and A. Madhavan (1995): Anatomy of the Trading Process: The Upstairs Market for Large Block Transactions; Analysis and Measurement of Price Effects *Review of Financial Studies* **9** 1-36.
- Kissell, R., and M. Glantz (2003): *Optimal Trading Strategies* **Amacom**.
- Lillo, F., J. Farmer, and R. Mantegna (2003): Master Curve for Price-Impact Function *Nature* **421** 129-130.
- Loeb, T. (1983): Trading Cost: The Critical Link between Investment Information and Results *Financial Analysts Journal* **39 (3)** 39-44.
- Rydberg, T. (2000): Realistic Statistical Modeling of Financial Data *International Statistical Review* **68 (3)** 233-258.
- Rydberg, T., and N. Shephard (2003): Dynamics of Trade-by-trade Price Movements; Decomposition and Models *Journal of Financial Economics* **1 (1)** 2-25.
- Sorensen, E., L. Price, K. Miller, D. Cox, and S. Birnbaum (1998): *The Solomon Smith Barney Global Equity Impact Cost Model* Technical Report **Solomon Smith Barney**.

Optimal Execution of Program Trades

Introduction

1. The Definition of the Program Trade: The program trade involves the sale or the purchase of a basket of stocks that is too large to be traded immediately in the market. When such trades are brokered, they take on of two forms.
2. The Agency Program Trade Type: In the *agency trade* type the broker executes the trade on behalf of the client on a commission basis, and all the risk of the trade is borne by the client.
3. The Principal Program Trade Type: In a *principal trade*, also called a principal basket, principal bid, or risk bid, the broker directly purchases the entire basket for a fixed price, usually expressed as a discount to the fair market value. By design, principal trades transfer all of the risk from the client to the broker in exchange for a single price, which therefore proxies for the risk of the market portfolio.
4. Program Trades Share of NYSE: Program trading represents an increasing percentage of the overall stock market volume. In 2002 program trades averaged over 30% of the New York Stock Exchange trading volume, up from approximately 20% in 19999 and 2000.
5. Share of the Principal Trades: Overall 50% of all trading volume took place in the NYSE, and of this 30-40% was done on a principal basis (NYSE (2002)).
6. Attributes of a Principal Trade: Almgren and Chriss (2003) begin by making two key observations about the program trading business. First the principal trade consists of two attributes – a basket of stocks, and a price.
7. The Basket Components and Price: The basket is determined by the client, but the price – usually expressed as a per share discount to the current market value – is agreed upon by the client and the broker, and the potential profitability of the trade depends upon the precise price that can be secured for trading the basket. Throughout

this chapter, the trade will be treated as the logical unit which consists of both the basket and its price.

8. Corporate Finance View of the Principal Trade: The second observation is that since program trading represents an investment of the firm's capital, the correct way to view performance is on an annualized basis.
9. First Result - The "Optimal" Execution: The treatment here constructs a mathematical framework for pricing and trading of principal baskets, yielding two main results. First it shows that for a broad class of measures of annualized risk-adjusted return there is a unique optimal way to trade the basket.
10. Second Result - The Information Ratio: Second, the measure called the *information ratio* of a trade is introduced, which is the ratio of the annualized expected profit to the annualized standard deviation of the profit.
11. Information Ratio as a Relative Value Metric: Given a proposed trade one can calculate the information ratio from the known information about the basket. Because the information ratio is annualized by the expected time to completion of the liquidation, it provides a way to compute the profitability of trades of different sizes and levels of liquidity. This yields a powerful tool for analyzing principal businesses.
12. Accommodating Various Market Impact Models: One does not need to know all of the constituents of a basket to compute its information ratio – just the volatility, the liquidity, and a proposed price for the basket. In particular, one must know the effect of trading a basket on the prices of its constituents, that is, the market impact. The methods in this treatment fit in with a wide variety of existing models.
13. The "Hurdle Rate" Minimum Price: The information ratio is primarily a pricing tool. By specifying a hurdle rate – a minimum information ratio that every principal basket must exceed – one can determine for a given principal basket what the minimum price will be that exceeds the hurdle rate.
14. Viability of the "Going Price": Alternatively it may be used as an evaluation tool. If one knows the "going price" of a trade, one can compute the information ratio of the trade based on that price to determine whether it is worthwhile to submit a winning bid for the business.

Efficient Frontier Pricing of Program Trades

1. *Ex-Ante* Risk Adjusted Return: The program trade is essentially the use of capital by a trading desk. In this treatment, the use of this capital is evaluated using an *ex-ante* risk adjusted return ratio – the ratio of the predicted profit to the standard deviation of the predicted profit. This turns out to the familiar information ratio, analogous to the familiar Sharpe ratio of the Information Theory.
2. The Investment Horizon Cost/Variance: Almgren and Chriss (2003) argue that for a principal desk engaged in an ongoing business, the correct approach is to *annualize* the cost and the variance, placing them in the context of other investment opportunities.
3. Discount Based Optimal Execution Point: Remarkably for each value of the discount received for trading the basket, there is then a *single* optimal point, independent of the risk preferences. This point corresponds to the single best value for the overall information ratio.
4. "Put Out To Bid" Call: The value of this ratio is therefore a potential tool to be used in evaluating whether to accept a certain piece of business at a certain price. In many situations the program trading business is "put out to bid". That is, the portfolio manager contacts multiple desks about a particular portfolio.
5. Information Ratio as a Decision Tool: Each business responds with a certain bid – the discount to the fair market value required to do the business. Program trading desks often know the level of bid required to win the business, and therefore the information ratio can be used as a hurdle or an evaluation tool to decide whether or not to bid at a level to win the business.

The Efficient Frontier Including Discount

1. Incorporating the Execution Discount Premium: This section connects the problem of liquidating a basket to the price of a basket. The aim is to eventually compute the

information ration of liquidating a basket incorporating the value of the discount to the fair value received in the transaction.

2. Units of the Execution Discount: Thus the assumption is that the trader will receive a discount of D dollars per share for a basket in the principal trade, and explicitly calculate the cash received in trading out the basket and its variance. For example, if the program trading desk were to be able to dump the entire portfolio onto the market without any market impact, it would earn a profit of DX dollars.
3. Execution Profit Under market Impact: In general, because of the market impact, the total expected profit of a trade would be

$$E = DX - C_{PERM} - C_{TEMP}$$

   and the variance is the same as stated in other publications (Almgren and Chriss (2000), Almgren (2003)).
4. Execution Profit Expectation and Variance: The expected profit and its variance can be explicitly calculated as functions of the execution time T:

$$\mathbb{E}_P[T] = \left(D - \frac{1}{2}\gamma X\right)X - \frac{k+1}{3k+1}\eta\left(\frac{X}{T}\right)^k X$$

$$\mathbb{V}_P[T] = \frac{k+1}{3k+1}\sigma^2 T X^2$$

5. Market Impact Reduction of Profit: Shortly $\mathbb{E}_P[T]$ and $\mathbb{V}_P[T]$ will be used to construct the information ratio of a trade, but for now some of its properties are examined. Clearly it is seen that the expected profit of a trade is its total discount DX reduced by a temporary impact and a permanent impact amount.
6. Difference between the Permanent and the Temporary Impact Costs: The effect of the permanent impact is to reduce the expected profit per share, as reflected in the size of the discount D, by an amount equal to the portfolio size, while temporary impact is proportional to a *per share* reduction in the expected profit of $\left(\frac{X}{T}\right)^k$.

7. Time Horizon Dependence of Profit: It is worth noting the dependence of $\mathbb{E}_P[T]$ and $\mathbb{V}_P[T]$ on T. Short liquidation times

$$T \to 0$$

correspond to

$$\mathbb{E}_P[T] \to -\infty$$

and

$$\mathbb{V}_P[T] \to 0$$

All profit is dissipated in impact costs, but no variance is incurred.

8. Intuitive Interpretation of Long Times: Long times

$$T \to \infty$$

correspond to

$$\mathbb{E}_P[T] = \left(D - \frac{1}{2}\gamma X\right)X$$

and

$$\mathbb{V}_P[T] \to \infty$$

Temporary impact costs are avoided completely by essentially holding the portfolio forever, but at the expense of any certainty of profit.

9. Elimination of Permanent Impact Costs: In principle, if the portfolio were held forever without trading, then the permanent impact costs will also be avoided:

$$T = \infty$$

is not the same as

$$T \to \infty$$

10. <u>The Zero Net Profit Trade</u>: Assuming that the discount is at least enough to compensate for the permanent impact, there is an intermediate point at which

$$\mathbb{E}_P[T] = 0$$

- the zero profit trade; impact costs are exactly compensated by the discount on average, but risk is taken to achieve this.

11. <u>Applying Customized Mean Variance Objective</u>: The previous work by Almgren and Chriss (2000) focused on various ways to draw the expectation cost expectation/variance combination frontier, in order to maximize either a mean-variance criterion $\mathbb{E}_P[T] + \lambda_u \mathbb{V}_P[T]$ or a value-at-risk measure $\mathbb{E}_P[T] + \lambda_v \sqrt{\mathbb{V}_P[T]}$ for a single trade in isolation. The next step is to consider this trade as part of an ongoing business.

Performance Measures

1. <u>Basket Specific Optimal Liquidation Time</u>: In this section the information ratio of a single trade is determined assuming a given discount of D dollars a share. First observe that the above analysis did not take into account the fact that different baskets will have different optimal liquidation times.
2. <u>Comparison across Different Principal Bids</u>: If the principal bids are to be considered in the context of ongoing business in relation to multiple investment opportunities,

then the expected profit per trade must be viewed in units that are comparable across different optimal times.

3. Per Basket Annualized Expected Return: This is done by directly annualizing the expected return by the expected amount of time it takes to liquidate substantially all of the basket, as determined by the characteristic time of the trade T. If a positive profit can be made, i.e., if

$$\mathbb{E}_P[T] > 0$$

then the trader prefers a shorter liquidation time to a longer liquidation time, other things being equal.

Annualization

1. Characteristic Time as the Investment Horizon: In order to annualize the expected profit and its variance, it is assumed that the entire invested capital becomes available for re-investment after one characteristic time T.
2. Intermittent Release of Invested Capital: In fact, liquidation is a continuous process; some capital is available immediately, and recovery of the full capital formally requires an infinite amount of time. Nonetheless T is a reasonable average value, and it is the simplest way to compare different trajectories.
3. Annualized Expected Return and Variance: Assuming that T is measured in years, annualizing is simply a matter of dividing by T. The expectation and the variance per year of trading is

$$\frac{\mathbb{E}_P[T]}{T} = \frac{\left(D - \frac{1}{2}\gamma X\right)X}{T} - \frac{k+1}{3k+1}\eta\left(\frac{X}{T}\right)^{k+1}$$

$$\frac{\mathbb{V}_P[T]}{T} = \frac{k+1}{3k+1}\sigma^2 X^2$$

4. Annualization of the Discount and the Permanent Costs: The annualized expectation is composed of two terms. The first term is the average rate at which the discount payment D is accepted, reduced by the cost of the permanent impact; since this is a fixed amount per portfolio, it is increased by rapidly trading.
5. Annualization of the Temporary Costs: The second term is the impact cost incurred by trading at a constant rate $\frac{X}{T}$ adjusted by a numerical coefficient to account for the non-linear shape of the trajectory.
6. Annualized Variance Independent of the Liquidation Time: Note that the annualized variance is independent of the liquidation time; this can be interpreted as saying that in the course of repeated execution, one is always invested in the market by the same amount on average.
7. Impact on the Efficient Frontier: This has an important implication. That is that if the efficient frontier is re-cast in terms of annualized expectation and annualized variance, it collapses to a single point.
8. Almgren Chriss (2003) Efficient Frontier Illustration: As illustrated in Almgren and Chriss (2003) the feasible region collapses into a half-infinite vertical line and the frontier itself has collapsed into a single point. This is a direct consequence of the annualized variance being independent of the trading time.
9. The Mandatory Capital Market Line Point: The striking consequence of this is that any measure of risk-adjusted profitability as constructed from the tangent line on the curve, regardless of the functional form or the parameter values, will pass through the highest point on this line.
10. The corresponding Optimal Trading Time: This means that, in particular, there is a unique best way to trade for any reasonable risk-adjusted return measure *regardless* of any particular risk-reward preferences, and is found simply by finding the value of T that maximizes $\frac{\mathbb{E}_P[T]}{T}$. This gives

$$T_{OPT} = \left[\frac{(k+1)^2}{3k+1}\right]^{\frac{1}{k}} \frac{\eta^{\frac{1}{k}} X}{\left(D - \frac{1}{2}\gamma X\right)^{\frac{1}{k}}}$$

11. Quasi Universal Optimal Trading Time: To emphasize the point made once more, T_{OPT} is the parameter representing *the* optimal trading strategy across a broad spectrum of possible risk-adjusted return measures. It is independent of the risk/reward preferences, but depends on the discount D. The aim next is to define and evaluate a particular risk-adjusted return measure.

Definition of the Information Ratio

1. Mathematical Definition of Information Ratio: Almgren and Chriss (2003) define and compute the *information ratio* of a trade, incorporating the effect of the discount D received in the transaction. For a given characteristic time T the information ratio with respect to T represents the annualized risk-adjusted expected profit that may be achieved by trading along a trajectory with parameter T

$$I(T) = \frac{\frac{\mathbb{E}_P[T]}{T}}{\sqrt{\frac{\mathbb{V}_P[T]}{T}}}$$

2. Risk Adjusted Basket Return: Note that $I(T)$ is the risk-adjusted return for a basket implicitly assuming a discount D received for the trade and a trading time parameter T.
3. Estimating the Maximal Information Ratio: The question is, for which T is $I(T)$ maximal? The answer to this is T_{OPT}, which can be substituted into the expression for $I(T)$ to get

$$I_{MAX} = \frac{(3k+1)^{\frac{k+2}{2k}}}{(k+1)^{\frac{3k+4}{2k}}} \frac{\left(D - \frac{1}{2}\gamma X\right)^{\frac{k+1}{k}}}{\eta^{\frac{1}{k}} X \sigma}$$

4. <u>Alternative Interpretation of $\mathbb{E}_P[T]$ and $\mathbb{V}_P[T]$</u>: Since the numerator and the denominator in the definition of I are both proportional to the portfolio size, it would be equivalent to considering $\mathbb{E}_P[T]$ and $\sqrt{\mathbb{V}_P[T]}$ above as *percentage* return and risk.
5. <u>Units of the Information Ratio</u>: Thus this quantity allows comparison of baskets and other investment opportunities of arbitrary size. It has units of $year^{-\frac{1}{2}}$ and thus should be compared only with other annualized measures.

Applications of the Information Ratio

1. <u>Discount Level Implied Information Ratio</u>: The two main applications of information ratio can now be states as answers to two questions. First, for a given level of discount than can be demanded for the trade, what is the information ratio of the basket?
2. <u>Information Ratio Implied Discount Hurdle</u>: Second, for a given information ratio *hurdle*, what minimum discount must be demanded in order to clear it?
3. <u>Discount Implied Maximum Information Ratio</u>: The answer to the first question is simply I_{MAX}, the information ratio of the basket assuming a discount of D. Put a different way

$$I_{MAX} = \frac{(3k+1)^{\frac{k+2}{2k}}}{(k+1)^{\frac{3k+4}{2k}}} \frac{\left(D - \frac{1}{2}\gamma X\right)^{\frac{k+1}{k}}}{\eta^{\frac{1}{k}} X \sigma}$$

shows that given a discount D that a desk can demand for a trade, it will yield a peak information ratio I_{MAX} which can then be used to determine if the trade clears a particular hurdle.

4. Information Ratio Implied Minimum Discount: The second question may be answered by simply inverting

$$I_{MAX} = \frac{(3k+1)^{\frac{k+2}{2k}}}{(k+1)^{\frac{3k+4}{2k}}} \frac{\left(D - \frac{1}{2}\gamma X\right)^{\frac{k+1}{k}}}{\eta^{\frac{1}{k}} X \sigma}$$

to yield

$$D_{MIN} = \frac{1}{2}\gamma X \left[\frac{(k+1)^{\frac{3k+4}{2k}}}{(3k+1)^{\frac{k+2}{2k}}} X \eta^{\frac{1}{k}} \sigma I_{HURDLE}\right]^{\frac{k}{k+1}}$$

The expression for D_{MIN} above gives the maximum that a desk can be bid for a given basket while still clearing the minimum information ratio threshold of I_{HURDLE}

5. Power Law Process Impact Table:

$\boldsymbol{k}$	$\boldsymbol{T}$	$\boldsymbol{I_{MAX}}$
$\frac{1}{2}$	$0.810\frac{\eta^2 X}{\widetilde{D}^2}$	$1.063\frac{\widetilde{D}^3}{\eta^2 X \sigma}$
1	$\frac{\eta X}{\widetilde{D}}$	$0.707\frac{\widetilde{D}^2}{\eta X \sigma}$
2	$1.134\frac{\eta^{\frac{1}{2}} X}{\widetilde{D}^{\frac{1}{2}}}$	$0.449\frac{\widetilde{D}^{\frac{3}{2}}}{\eta^{\frac{1}{2}} X \sigma}$

6. Power Law Process Impact – Legend: The table above shows the optimal trading time and maximum information ratio for three different values of the market exponent k. $\widetilde{D}$ is computed from

$$\widetilde{D} = D - \frac{1}{2}\gamma X$$

the discount reduced by the anticipated permanent impact costs.

7. Execution Time/IR Exponent Dependence: The table above presents the specific forms of these expressions for a few particular and important values of k. Although the analytical expressions above are complex, for a specific choice of k they reduce to simple numerical coefficients.
8. Execution Time/IR Discount Dependence: What is particularly noteworthy is the relationship between the price of the principal basket – as embodied in the discount to fair – and both the information ratio and the optimal time for liquidation. It depends on the market impact functions assumed, and can be quiet sensitive to small movements.
9. Example - BARRA Market Impact Exponent: For example, for the BARRA model

$$k = \frac{1}{2}$$

the maximum information ratio increases as the *cube* of the discount, after allowing for permanent impact, and the optimal time decreases as the square of the discount.

10. Risk-Adjusted Profit Discount Dependence: One interpretation of these results is that small changes in the price of the principal bid, expressed as cents per share discount to fair value, can have significant impact on both the risk-adjusted profitability of the trade and the time it takes to liquidate the trade. For instance, a basket that commands $2.5c$ per share discount to fair is twice as profitable on a risk-adjusted basis versus one that commands a $2c$ per share discount.

References

- Almgren, R. F., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.

- Almgren, R. F. (2003): Optimal Executions with Nonlinear Impact Functions and Trading-Enhanced Risk *Applied Mathematical Finance* **10 (1)** 1-18.
- Almgren, R. F., and N. Chriss (2003): Bidding Principals *Risk* **16 (6)** 97-102.
- NYSE (2002): Press Release.

Bayesian Trading with a Daily Trend

Overview, Motivation, and Synopsis

1. Parametric Estimation using Updated Information: Standard models of algorithmic trading neglect the presence of a daily trend. Almgren and Lorenz (2006) construct a model in which the trader uses information from the observations of price evolution during the day to continuously update his estimate of other traders' target sizes and directions.
2. Constraint Based Optimal Trajectory Generation: The trader uses this information to determine an optimal trade schedule to minimize total expected costs of trading, subject to sign constraints, e.g., never buy as part of a sell program.
3. Dynamic Strategies using Projected Cost: It is argued that these strategies are determined using very simple dynamic reasoning – at each moment they assume that the current condition will last until the end of trading – they are in fact globally optimal strategies as would be determined by dynamic programming.

Introduction and the Associated Literature

1. Market Information Based Learning Updates: The work of Almgren and Lorenz (2006) presents a model for price dynamics and optimal trading that explicitly includes the daily trend and the trader's attempt to learn the targets of other market participants.
2. Drawbacks of Current Approaches: This is in contrast to most current models of optimal trading strategies that view time as an undifferentiated continuum, and other traders as a collection of random noise sources. This approach has two primary motivations.

3. Incorporating the Explicit Daily Trend: The first set of motivations is the academic articles by Brunnermeier and Pedersen (2005) and Carlin, Lobo, and Viswanathan (2007). In these articles, institutional trading has an explicit daily cycle, based on the assumption that at the beginning of each day, each informed market participant, or institutional investor, is given an exogenously specified trade target.
4. Targets of the Informed Traders: These participants know the targets of the other informed traders, and they must decide whether to cooperate with their peers so as to not lose value to the uninformed traders, or whether to compete and take value from their peers.
5. Dynamic Estimation of Peer Targets: The novel feature of this market is that the participants do not know each other's targets, but must guess them by observing the prices throughout the day. It is taken for granted that informed participants will use all available information to compete with each other.
6. Dynamic Determination of Execution Trajectory: The second set of motivations is the popularity of execution algorithms that adapt to the changes in the prices of the asset being traded, either by accelerating execution when the prices move in the traders' favor, or conversely.
7. Momentum and/or Mean Reversion: Although these optimal trade models may be handled by introducing various forms of risk aversion (Kissell and Malamut (2006), Almgren and Chriss (2007)), the most common justification for them is a belief in mean reversion or momentum of the asset price.
8. Underlying Institutional Investor Drift Factor: The model introduced in this chapter may be interpreted as one plausible way to model price momentum. There is an underlying drift factor, caused by the net positions being executed by the other institutional investors.
9. Daily Institutional Trader Price Momentum: This factor is approximately constant throughout the day because the other traders execute across the entire day. Thus price increases in the early part of the day suggest that this factor is positive, which suggests that the prices will continue to increase throughout the day.
10. Short Term Price Change Correlation: This is different from the short term momentum model in which the price change across one short period of time is

correlated with the price change across the preceding period; most empirical evidence shows that such correlation is weak if it exists at all.

11. Incorporating the Daily Price Momentum: The strategies presented in this chapter exploit this momentum to minimize the expected value of the trading costs, somewhat in the spirit of Bertsimas and Lo (1998), except that because the focus is on long term momentum, higher gains can be obtained.
12. Origin of the Daily Trend: The daily trend is an essential feature of this model. Large institutional participants make investment decisions overnight and implement them through the following day.
13. Trend Period vs. Implementation Horizon: Within each day morning is different from the afternoon, since an intelligent trader will spend the early hours collecting information about the targets of the other traders, and will use this information to trade in the rest of the day.
14. Random Nature of Trade Decisions: By contrast, in the market that is implicitly assumed by most models, trade decisions are made at random times, and trade programs have random durations, with no regard to the daily trends. Thus if one observes a buy pressure from the market as a whole, one has no reason to believe that this pressure will last more than a short term. From the point of view of optimal trading, price motions are purely random.
15. Constraint on the Trade Direction: In addition, the very important feature of constraints on the trade direction is incorporated; the trader must never sell as part of a buy program even if this yields lower expected costs – or even expected profit – because of the anticipated negative drift in the price.
16. Reasons for the Constraint Imposition: This is for two reasons. First the point of view of a broker/dealer executing an agency trade for a client is taken. Second, the bid/offer spread and the other fixed costs are neglected, which greatly reduce the profitability of such reversing strategies. These adaptive strategies simply sift the buying or selling from one period to another.
17. Binding Nature of the Constraint: This constraint is often binding, and globally affects the structure of optimal strategies. In many cases it leads to the determination

of an optimal end time for trading, and sometimes directs the strategy to stop completely for a finite period in the middle of the execution.

18. Bayesian Learning of Institutional Direction: In the section below a model is presented of Brownian motion with a drift whose distribution is updated continuously using Bayesian inference.
19. Best Estimate Based Optimal Trajectory: Subsequently, optimal strategies are presented which, surprisingly, can be computed by determining a "static" optimal Trajectory at each moment, assuming that the best parameter estimates as of that moment will persist through the end of the trading period.

Price Model Using Bayesian Update

1. Arithmetic Brownian Motion Price Dynamics: Trading in a single asset is considered, whose price $S(t)$ obeys an arithmetic random walk

$$S(t) = S_0 + \alpha t + \sigma B(t)$$

for

$$t \geq 0$$

where $B(t)$ is a standard Brownian motion, σ is the absolute volatility, and α is the drift. In the presence of intra-day seasonality, t is interpreted as a volume time relative to a historical profile.

2. Origin of Volatility - Uninformed Traders: The interpretation of volatility is that it comes from the activity of the "uninformed" traders, whose average behavior can be predicted reasonably well. Mathematically, the value of σ is assumed to be known precisely – for a Brownian process σ can be estimated arbitrarily precisely from an arbitrarily short observation of the process.

3. Origin of Drift - Institutional Traders: The drift is interpreted as coming from the activities of the other institutional traders, who have made trade decisions before the market opens, and who expect to execute these trades throughout the day. If these decisions are in the aggregate weighted to buys, then this will cause a positive price pressure and an upwards drift – conversely for overall selling.
4. Trade Direction Based Drift Estimate: No knowledge of the net direction of the trade estimates is presumed, but inferred by observing the prices. It is implicitly assumed that the traders are using VWAP-like strategies rather than the pure arrival price, so that their trading is not "front-loaded". This assumption is questionable; if their strategies are front-loaded, then the drift coefficient would vary throughout the day.
5. Drift Belief - Mean and Confidence: Thus the drift α is assumed consant throughout the day, but its value is unknown. At the beginning of the day the prior belief

$$\alpha \sim \mathcal{N}(\bar{\alpha}, \nu^2)$$

   will be updated using price observations throughout the day.
6. "Frequentist" Volatility vs. "Bayesian" Drift: There are two sources of randomness in the problem – the continuous Brownian motion representing the uninformed traders, and the single drift coefficient representing the constant trading of the large traders.

Bayesian Inference

1. Drift Estimate from Realized Price: Intuitively as the trader observes the prices from the beginning of the day onwards, he/she starts to get a feel for the day's overall flow. Mathematically the stock price trajectory $S(\tau)$ is known for

$$0 \leq \tau \leq t$$

   In fact all the information about the drift comes from the final value $S(t)$

2. Bayesian Formulation of the Price Evolution: Conditional on the value of α the distribution of $S(t)$ is

$$S(t) - S_0 \sim \mathcal{N}(\alpha t, \nu^2 t)$$

The unconditional distribution can be found after some calculation as

$$S(t) - S_0 \sim \mathcal{N}(\bar{\alpha} t, [\sigma^2 + \nu^2 t] t)$$

3. The Posterior Conditional Drift Distribution: The Bayes' rule is then used:

$$Prob\ (\alpha | S(t)) = \frac{Prob\ (S(t)|\alpha) \cdot Prob\ (\alpha)}{Prob\ (S(t))}$$

to obtain the posterior conditional distribution

$$\alpha \sim \mathcal{N}\left(\frac{\bar{\alpha}\sigma^2 + \nu^2[S(t) - S_0]}{\sigma^2 + \nu^2 t}, \frac{\sigma^2}{\sigma^2 + \nu^2 t}\nu^2\right)$$

conditional on $S(t)$.

4. Best Estimate of Mean/Variance: This represents the best estimate of the true drift α as well as the uncertainty in this estimate based on the combination of the prior belief with the price information observed to time t.
5. Fully Certain Estimate of Drift: This formulation accommodates a wide variety of belief structures. If the belief in the initial information is perfect then one sets

$$\nu = 0$$

and the updated belief is always

$$\alpha = \bar{\alpha}$$

with no incremental updating.

6. Fully Uncertain Estimate of Drift: If one believes that there is no reliable prior information then

$$\nu^2 \to \infty$$

and the estimate is

$$\alpha \sim \mathcal{N}\left(\frac{S(t) - S_0}{t}, \frac{\sigma^2}{t}\right)$$

coming entirely from intra-day observations.

7. The $t = 0$ and $t \to \infty$ Asymptotes: For

$$t = 0$$

one has

$$S(0) = S_0$$

and the belief is just the prior. As

$$t \to \infty$$

the estimate becomes

$$\alpha \sim \mathcal{N}\left(\frac{S(t) - S_0}{t}, 0\right)$$

so much information has been accumulated that the prior belief becomes irrelevant.

Trading and Price Impact

1. The Order Size and Horizon: The trader has an order of X shares which begins at time

$$t = 0$$

and must be completed by the time

$$t = T < \infty$$

For concreteness it is supposed that

$$X > 0$$

which is interpreted as a buy order.

2. Trade Rate and Trading Trajectory: A *trading trajectory* is a function $x(t)$ with

$$x(0) = X$$

and

$$x(T) = 0$$

representing the number of shares to buy at time t. The corresponding *trading rate* is

$$v(t) = -\frac{dx(t)}{dt}$$

It shall be required that

$$v(t) \geq 0$$

for all t so that a program never sells as part of the buy order. Together with the endpoint constraints this requires

$$0 \leq x(t) \leq X$$

but it may also be binding in the interior of the region.

3. Linear Temporary Market Impact Function: A linear temporary market impact function is used for simplicity, although the empirical work of Almgren, Thum, Hauptmann, and Li (2005) suggests a concave function. The actual execution price is

$$\tilde{S}(t) = S(t) + \eta v(t)$$

where

$$\eta > 0$$

is the coefficient of temporary market impact.

4. Execution Trajectory Implementation Shortfall: $\mathcal{C}$ is the total cost of executing the buy program relative to the initial value

$$\mathcal{C} = \int_0^T \tilde{S}(t) v(t) dt - X S_0 = \sigma \int_0^T x(t) dB(t) + \eta \int_0^T v^2(t) dt + \alpha \int_0^T x(t) dt$$

5. Deterministic and Random Cost Components: Here α is the true drift, and this determines cost, whether or not its true value is known. $\mathcal{C}$ is a random variable, both because $S(t)$ is random, and because the optimal trading trajectory $v(t)$ may be adapted to S.

Optimal Trading Strategies

1. Classic Mean-Variance Risk Aversion: This section addresses the question of what trading strategies are optimal given the above model for price evolution and market impact. In the classic arrival price framework of Almgren and Chriss (2000) trajectories are determined by a trade-off between market impact and aversion to risk caused by volatility.
2. Balance between Slow/Fast Trading: The trader wants to complete the trade quickly to reduce exposure to price volatility; He or she wants to trade slowly to reduce the cost of market impact. The optimal trajectory is determined as a balance between these two effects, parametrized by a coefficient of risk aversion.
3. Optimal Cost Adaptive Trading Strategies: Risk-averse trading strategies can behave strangely in time even in the classic mean variance framework (Almgren and Lorenz (2007)) depending on the precise formulation of the mean-variance trade-off.
4. Complication Introduced by the Drift Variance: In this case the problem is complicated by the need to account for the variance in the estimate of α. Almgren and Lorenz (2006) claim to have obtained partial solutions for the risk-averse problem, but the resulting complexity obscures the underlying structure.
5. Neglecting the Mean Variance Risk Aversion: To focus on the drift, which is the most important new aspect of this problem, risk aversion is neglected here; only the expectation of the trading cost is sought to be minimized.
6. Cost associated with the Drift: That is, it is assumed that the pressure to complete the trade rapidly comes primarily by a desire to capture the price motion expressed by the drift α, and it is this effect that must be balanced against the desire to reduce the impact costs by trading slowly.
7. Positive Baseline Drift Assumption: To support this description it is generally supposed that the original buy decision was made because of the traders belief that

$$\bar{\alpha} > 0$$

Thus it is expected

$$\alpha > 0$$

in

$$\mathcal{C} = \int_0^T \tilde{S}(t)v(t)dt - XS_0 = \sigma \int_0^T x(t)dB(t) + \eta \int_0^T v^2(t)dt + \alpha \int_0^T x(t)dt$$

and the term $\alpha \int_0^T x(t)dt$ is a positive cost. It may be that the true value has

$$\alpha < 0$$

or that the intermediate price movements result in the formation of a negative estimate.

8. <u>Hard Trade Completion Time $t = T$</u>: Because the point of view is that of a broker/dealer executing an agency trade, it shall always be required that the trade be completed by

$$t = T$$

unless the instructions are altered.

9. <u>Conditional Expectation of Unrealized Cost</u>: For any deterministic trajectory $x(t)$ specified at

$$t = 0$$

$\mathcal{C}$ is a Gaussian variable. Conditional on the true value of α it has the expected value

$$\mathbb{E}[\mathcal{C}] = \eta \int_0^T v^2(t)dt + \alpha \int_0^T x(t)dt$$

10. The Bayesian Estimate for α: From

$$\alpha \sim \mathcal{N}\left(\frac{\bar{\alpha}\sigma^2 + \nu^2[S(t) - S_0]}{\sigma^2 + \nu^2 t}, \frac{\sigma^2}{\sigma^2 + \nu^2 t}\nu^2\right)$$

conditional on $S(t)$ the best estimate at time t for the value of α is

$$\alpha_*(t, S) = \frac{\bar{\alpha}\sigma^2 + \nu^2[S(t) - S_0]}{\sigma^2 + \nu^2 t}$$

where

$$S \equiv S(t)$$

11. Bayesian Estimate of Unrealized Cost: Because the expectation

$$\mathbb{E}[\mathcal{C}] = \eta \int_0^T v^2(t)dt + \alpha \int_0^T x(t)dt$$

conditional on α is linear α one may substitute the expected value of α_* to see that, conditional on the information available at time t the expected cost of the remaining program is

$$\mathbb{E}[t, x(t), S, \{x(t)\}] = \eta \int_t^T v^2(\tau)d\tau + \alpha_*(t, S) \int_t^T x(\tau)d\tau$$

12. Nomenclature - Description of the Terms: On the left t is the current time, $x(t)$ is the number of shares currently remaining to buy, S is the current price, and $\{x(t)\}$ denotes the liquidation strategy that will used on the remaining time

$$t \leq \tau \leq T$$

13. Strategy Objective - Minimizing Transaction Cost: The trading goal is to choose the remaining strategy to minimize this expected cost, i.e., determine $x(\tau)$ for

$$t \leq \tau \leq T$$

so that

$$\underset{\{x(\tau)\}}{min} \mathbb{E}[t, x(t), S, \{x(\tau)\}]$$

14. Invariance of the Drift Estimate: In computing this solution, it is assumed that the drift estimate $\alpha_*(t, S)$ does not change during the interval

$$t \leq \tau \leq T$$

In fact it will change as new price is obtained.

15. Dynamically Recomputed Unrealized Trajectory Cost: The actual strategy will only use the instantaneous trade rate of this trajectory, continuously responding to price information. This is equivalent to following the strategy only for a very small time interval Δt then re-computing. Thus the strategy is highly dynamic.
16. Equivalence with Full Dynamic Optimization: It shall be argued that the trajectory thus determined is the true optimum strategy that would be computed by a full dynamic optimization. Loosely speaking this will be because the expected value of future updates is zero, and thus they do not change the strategy of a risk-neutral trader.

Trajectory by Calculus of Variations

1. Trajectory Perturbation at the End-points: A small perturbation of the path $x(\tau) \mapsto x(\tau) + \Delta x(\tau)$ is considered for $t \leq \tau \leq T$ Since $x(\tau)$ is fixed at $\tau = t$ and $\tau = T$ this perturbation must have $\Delta x(t) = \Delta x(T) = 0$
2. The Corresponding Bayesian Cost Impact: The associated trade rate perturbation is $\Delta v(\tau) = -\Delta x'(\tau)$ and the perturbation in cost – assuming that $x(\tau)$ and $\Delta x(\tau)$ are twice differentiable – is $\Delta\mathbb{E}[t, x(t), S, \{x(t)\}] = \eta \int_t^T 2v(\tau)[\Delta v(\tau)]d\tau + \alpha_*(t, S) \int_t^T [\Delta x(\tau)]d\tau = \int_t^T \{-2\eta x''(\tau) + \alpha_*(t, S)\}[\Delta x(\tau)]d\tau$
3. Cost Optimized Trajectory - Necessary Condition: Here $\alpha_* \equiv \alpha_*(t, S)$ is the best available drift estimate using information available at time t which we assume is constant for $t \leq \tau \leq T$ If $x(\tau)$ is an optimal solution then there must not exist any admissible $\Delta x(\tau)$ that gives $\Delta\mathbb{E}[t, x(t), S, \{x(\tau)\}] > 0$
4. Unconstrained Trajectories - The Holdings ODE: For now the sign constraints on $x'(\tau)$ is neglected. The $\Delta x(\tau)$ may have either positive or negative values independently for each τ and optimizing $x(\tau)$ must satisfy the ordinary differential equation (ODE) $x''(\tau) = \frac{\alpha_*}{2\eta}\ t \leq \tau \leq T$
5. Unconstrained Trajectory - The Holding Solution: The solution to this equation that satisfies the boundary conditions is $x(\tau) = \frac{T-\tau}{T-t}x(t) - \frac{\alpha_*}{4\eta}(\tau - t)(T - \tau)\ t \leq \tau \leq T$ and the corresponding instantaneous trade rate is $v(t, x) = -x'(\tau)|_{\tau=T} = \frac{x(t)}{T-t} + \frac{\alpha_*}{4\eta}(T - t)$ as a function of time and shares remaining.
6. Unconstrained Trajectory Holdings Constraint Violation: This solution may violate the constraints; if α_* is large then the quadratic term in $x(\tau) = \frac{T-\tau}{T-t}x(t) - \frac{\alpha_*}{4\eta}(\tau - t)(T - \tau)\ t \leq \tau \leq T$ may cause $x(\tau)$ to dip below zero, which would cause $v(t)$ in $v(t, x) = -x'(\tau)|_{\tau=T} = \frac{x(t)}{T-t} + \frac{\alpha_*}{4\eta}(T - t)$ to become negative.

7. Unconstrained Trajectory - The Holdings Component: The unconstrained solution is the sum of two parts. The first piece is proportional to $x(t)$ and represents the linear (VWAP) liquidation of the current position; it is the optimal strategy to reduce the expected impact costs with no risk aversion.
8. Unconstrained Trajectory - Holdings Drift Component: The second piece is independent of $x(t)$ and would exist even if the trader has no initial position. Just as in the solutions of Bertsimas and Lo (1998), this second piece is effectively a proprietary trading strategy superimposed on liquidation.
9. Unconstrained Trajectory - Relative Component Contribution: The magnitude of this strategy, and hence the possible gains, are determined by the ratio between the expected drift and the liquidity coefficient. Imposition of this constraint will couple these pieces together.
10. Constrained Trajectories - Consequences of Violation: If the constraint becomes binding then it is no longer clear that the integration by parts procedure used to derive $\Delta\mathbb{E}[t, x(t), S, \{x(t)\}] = \eta \int_t^T 2v(\tau)[\Delta v(\tau)]d\tau + \alpha_*(t, S) \int_t^T [\Delta x(\tau)]d\tau = \int_t^T \{-2\eta x''(\tau) + \alpha_*(t, S)\}[\Delta x(\tau)]d\tau$ is valid.
11. Constrained Trajectories - Non Smooth Edges: For example if a trajectory that crosses the axis $x = 0$ is simply clipped to satisfy $x \geq 0$ then the derivative will be discontinuous. A more refined use of the calculus of variations gives the additional condition that $v(\tau)$ must be continuous though not differentiable.
12. Constrained Trajectories - Smoothing the Edges: Thus when solutions meet the constraint, they must do so smoothly. Solutions are obtained by combining the ODE's $x''(\tau) = \frac{\alpha_*}{2\eta}\ t \leq \tau \leq T$ in the regions of smoothness, with "smooth pasting" conditions at the boundary points.
13. Constrained Trajectories - The "Critical" Drift: The results may be summarized as follows. There is a critical drift value α_C such that if $|\alpha_*| \leq \alpha_C$ then the constraint in binding. The solution is the one given in $x(\tau) = \frac{T-\tau}{T-t}x(t) - \frac{\alpha_*}{4\eta}(\tau - t)(T - \tau)\ t \leq \tau \leq T$ and $v(t, x) = -x'(\tau)|_{\tau=T} = \frac{x(t)}{T-t} + \frac{\alpha_*}{4\eta}(T - t)$

14. Constrained Trajectories – Super-critical Drift Horizon: If $\alpha_* > \alpha_C$ the solution is still $x(\tau) = \frac{T_* - \tau}{T_* - t} x(t) - \frac{\alpha_*}{4\eta}(\tau - t)(T_* - \tau)$ $t \le \tau \le T_*$ and $v(t, x) = -x'(\tau)|_{\tau = t} = \frac{x(t)}{T_* - t} + \frac{\alpha_*}{4\eta}(T_* - t)$ but with a shortened end-time $T_* < T$ determined by $T_* - t = \sqrt{\frac{4\eta x(t)}{\alpha_*}}$
15. Constrained Trajectories Super-critical Drift Value: The values for T_* above is determined so that $x'(T_*) = x(T_*) = 0$ The threshold value α_C is the value of α_* for which $T_* = T$ $\alpha_C(x(t), T - t) = \frac{4\eta x(t)}{(T - t)^2}$
16. Constrained Trajectories Sub Critical Drift: If $\alpha_* < -\alpha_C$ then the solution is one of $x(\tau) = \frac{T - \tau}{T - t_*} x(t) - \frac{\alpha_*}{4\eta}(\tau - t_*)(T - \tau)$ $t_* \le \tau \le T$ and $v(t, x) = -x'(\tau)|_{\tau = t_*} = \frac{x(t_*)}{T - t_*} + \frac{\alpha_*}{4\eta}(T - t_*)$ except that trading does not begin until a starting time t_* determined by $T - t_* = \sqrt{\frac{4\eta x(t)}{-\alpha_*}}$ This value is determined so that $x'(t_*) = 0$ and $x(t_*) = x(t)$ The threshold value α_C is the value $-\alpha_*$ for which $t_* = t$
17. Constrained Trajectories - Illustration: Almgren and Lorenz (2006) illustrate the constrained solutions $x(\tau)$ starting at time t with shares $x(t)$ and drift estimate α. For $\alpha > 0$ the trajectories go below the linear profile to reduce the expected purchase cost. As shown in the shaded region in the illustration for $|\alpha| \le \alpha_C$ the constraint is not binding. At $\alpha = \alpha_C$ the solutions become tangent to the line $x = 0$ at $\tau = T$ and for larger values they hit $x = 0$ with zero slope at $\tau = T_* < T$ For $\alpha < -\alpha_C$ trading does not begin until $\tau = t_* > t$

References

- Almgren, R. F., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.
- Almgren, R. F., C. Thum, E. Hauptmann, and H. Li (2005): Equity Market Impact *Risk* **18 (7)** 57-62.

- Almgren, R. F., and J. Lorenz (2006): Bayesian Adaptive Trading with a Daily Cycle *Journal of Trading* **1 (4)** 38-46.
- Almgren, R. F., and J. Lorenz (2007): Adaptive Arrival Price, in: *Algorithmic Trading III (B. R. Bruce, editor)* **Institutional Investor** 59-66.
- Bertsimas, D., and A. W. Lo (1998): Optimal Control of Execution Costs *Journal of Financial Markets* **1** 1-50.
- Brunnermeier, M., K., and L. H. Pedersen (2005): Predatory Trading *Journal of Finance* **60 (4)** 1825-1863.
- Carlin, B. I., M. S. Lobo, and S. Viswanathan (2007): Episodic Liquidity Crises: Cooperative and Predatory Trading **62 (5)** 2253-2274.
- Kissell, R., and R. Malamut (2006): Algorithmic Decision-Making Framework *Journal of Trading* **1 (1)** 12-21.

Cost Adaptive Arrival Price Trading

Synopsis and Key Results

1. Cost Uncertainty Balance in Trading: Electronic trading of Equities and other securities makes heavy use of *arrival price* algorithms that determine optimal trade execution strategies by balancing the market impact cost of rapid execution against the volatility risk of slow execution.
2. Optimal Static Mean Variance Strategies: In the standard formulation mean-variance optimal strategies are static; they do not modify the execution speed in response to price motions observed during trading.
3. Dynamically Adjusting Risk Tolerant Profiles: Almgren and Lorenz (2007) show that with a more realistic formulation of the mean-variance tradeoff, and even with no momentum or mean-reversion in the price process, substantial improvements are possible for adaptive strategies that spend trading gains to reduce risk by accelerating execution when the price moves in the traders' favor. The improvement is larger for large initial portfolios.

Introduction, Background, and Motivation

1. Breaking-Down a Given Order: Algorithmic trading represents a large and growing fraction of the total order flow, especially in the equity markets. When the size of a requested buy or sell order is larger than what the market can immediately supply or absorb, then the order must be worked across some period of time, exposing the trader to price volatility.

2. Tailoring Execution to Risk Preferences: The algorithm attempts to achieve an average execution price whose profitability is suited to the client's preferences. Almgren and Lorenz (2007) propose a way to dramatically improve this distribution.
3. Benchmark Implementation Shortfall: Arrival price algorithms, which are currently the most widely used framework, take as their benchmark the *pre-trade* or *decision* price. The difference between the execution price and the benchmark is the *implementation shortfall* which is an uncertain quantity since the order execution takes a finite amount of time.
4. Expected Value of Shortfall: In the most straightforward implementation of this model, the expected value of the implementation shortfall is entirely due to the market impact incurred by trading at a non-zero rate (neglecting any anticipated price drift); this expected cost is minimized by trading as slowly as possible, for example, using a VWAP strategy across the maximum allowed time horizon.
5. Variance of Implementation Shortfall: Since market impact is assumed to be deterministic, the variance of the implementation shortfall is entirely due to price volatility; this variance is minimized by trading rapidly.
6. Efficient Frontier of Optimal Trading: This risk-reward trade-off is very common in finance, and a variety of criteria can be used to determine risk-averse optimal solutions. Arrival price algorithms compute the set of *efficient* strategies that compute the risk for a specified level maximum of expected cost or the converse; the set of such strategies is summarized in the *efficient frontier of optimal trading* introduced by Almgren and Chriss (1999, 2000).
7. Independence from the Portfolio Size: This simple mean-variance approach has the advantage that the risk-reward tradeoff is independent of the initial wealth – a useful property in an institutional setting.
8. Static vs. Dynamic Execution Schemes: A central question is whether the trade schedule should be *static* or *dynamic*; should the list of shares to be executed at each time interval be computed and fixed before the trading begins, or should the list be updated in "real time" using information revealed during the execution?
9. Static vs. Dynamic Schemes Equivalence: The observations of Almgren and Chriss (2000) is that, under very realistic assumptions about the price process (arithmetic

random walk with no serial correlation), static strategies are *equivalent* to dynamic strategies. No value is added by considering "scaling" strategies in which the execution speed changes in response to price motions.

10. <u>Static Strategy - Initial Time Determination</u>: To be more specific, two different specifications of the trade scheduling problem are considered. For a static strategy the entire trade schedule is required to be fixed in advance – Huberman and Stanzl (2005) suggest that a reasonable example of this is insider trading – where trades must be announced in advance.
11. <u>Pre-determination of the Trajectory Cost Distribution</u>: For any candidate schedule the mean and the variance are calculated at the initial time, and the optimal schedule is determined for a specific risk aversion level.
12. <u>Dynamic Determination of the Trading Schedule</u>: For a dynamic strategy – as is usually understood in dynamic programming – arbitrary modification of the strategy is allowed at any time. To re-calculate the trade list all available information is used at that time, and the strategies are valued using a mean-variance trade-off of the remaining cost, using a constant parameter of risk aversion.
13. <u>Reduction of Dynamic to Static</u>: In the model of Almgren and Chriss (2000), the first and the second strategies have the same solution. Liquidity and volatility are assumed known in advance, so the only information revealed is the asset price motion.
14. <u>Price Distribution Independence from the Realization</u>: Price distribution revealed in the first part of the execution does not change the probability distribution of future price changes. Because the mean-variance trade-of is independent of the initial wealth, trading gains or losses incurred in the first part of the program are "sunk costs" and therefore do not influence the strategy for the remainder.
15. <u>Trade Rate Determination Rule</u>: Almgren and Lorenz (2007) present an alternate formulation. In this, they pre-compute the *rule* determining the trade rate as a function of price, using a mean-variance tradeoff measures at the initial time. Once trading begins the rule may not be modified, even if the trader's preferences re-evaluated at an intermediate time would lead him or her to choose a different strategy, as in the second strategy above.

16. Differences among the Computed Trajectories: The optimal solution to the third strategy is generally not the same as the solutions to the first and the second strategies.
17. Example: Comparison of the Strategies: As an illuminating contrast, in the well-known problem of option hedging, the optimal hedge position, once the trade list, depend on the price, and hence are not known until the price is observed, although the *rule* giving this hedge position is computed in advance using dynamic programming. Thus the first strategy is dramatically sub-optimal, and the second gives the same results as the third.
18. When is the Third Strategy Optimal? For algorithmic trading, the improvement of the third strategy over the first and the second come from introducing the negative correlation between the trading gains or losses in the first part of the execution and the market impact costs incurred in the second part.
19. Extraneously Imposed Serial Correlation Rule: Trading gains and losses due to price movements are serially uncorrelated, but can be correlated with the market impact costs via a simple rule; if the price moves in the traders' favor in the early part of the trading, then those gains are spent on the market impact costs by accelerating the remainder of the program.
20. Adaptive Strategy - Contra Price Move: If the price moves against, then the future costs are reduced by trading more slowly, despite the increased exposure to risk of future fluctuations. The result is an overall decrease in the variance measured at the initial time, which can be traded for a decrease in the expected cost.
21. *Ex ante* vs *ex post* Optimization: In practice there are no artificial constraints in the adaptivity of the trading strategies. The key observation contained in this chapter is that the *ex ante* mean-variance optimization expressed by the third formulation corresponds better to the way the trading results are measured in practice, via *ex post* sample mean and variance over a collection of similar programs.

Adaptive Strategies – A Simple Illustration

1. Universe of Available Sample Bets: Suppose that two bets are available. Bet A pays 0 or 6 with equal probability; its expected value is 3 and its variance is 9. Bet B pays 1 with certainty; its expected value is 1 and its variance is 0.
2. Per-Strategy Objective Utility Value: Consider the case of a risk-averse investor whose coefficient of risk aversion is $\frac{1}{9}$; he assigns an *ex ante* value of $E - \frac{1}{9}V$ to a random payout with an expected value E and variance V. For this investor a single pay of A has a value 2 and a single pay of B has a value 1 so he prefers A.
3. Two Plays with Outcome Independence: Now suppose that the investor plays this game twice, with independence between the outcomes. Three ways in which he chooses his bets are considered.
4. First Strategy - Optimal Outcome: In a static strategy, the sequence AA, AB, BA, or BB must be fixed before the game begins. By independence choice AA has twice the value of A and is preferred. Its value is 4.
5. Second Strategy - Constant Wealth Effect: In a dynamic strategy, the second bet is chosen after the result of the first play is learnt. By that time the first result will be a constant wealth effect, so A will always be chosen on the second play.
6. Second Strategy - Optimal Outcome: Knowing that that will be the future choice A is chosen on the first bet as well to maximize the total value measured at the initial time. Thus the strategy and the payoff are the same as in the static case.
7. Third Strategy - Sequential Play Rule: In the new formulation the investor specifies *three* choices; his bet on the first play, his bet on the second play if he wins the first one, and his bet on the second play if he loses the first.
8. Third Strategy - Optimal Outcome: The optimal rule is to bet A on the first play, then if he wins to choose B, if he loses to play A again, giving payouts of 0, 6, 6, and 7 with equal probability. Its value is 4.06, better than the first two strategies.
9. Optimally using Slow/Fast Trading: In this model bet A corresponds to slow trading, with high expected value (low cost) and high variance, and B is fast trading. If the random outcome (trading gain) in the first period is positive, then the trader spends some of this gain on reducing the variance in the second period.

10. Extension to Multi-Play: Now suppose the investor plays this game many times in sequence, and wishes to optimize the sample mean and variance, combined the coefficient of risk aversion.
11. Ex Post vs Ex Ante Single Play: If the results are reported over individual plays, then the *ex post* sample mean and variance will be close to the *ex ante* expectation and variance of a single play, and the optimal strategy would be to bet A each time, as in the first and the second strategies above.
12. Aggregation Over Play Pairs: However, suppose the results are aggregated over *pairs* of plays. That is, the gains of play 1 and play 2 are added together, play 3 and play 4 are added, *etc*.
13. Pairs Connected via Rule: Then the third strategy above, which is adaptive, will give the best results; within each pair choose the second bet based on the result of the first one. If the results are grouped into larger sets, then a more complicated strategy will be even more optimal.

Trading in Practice

1. Reporting Driven Aggregation Granularity: As in the simple example, the question of which formulation is more realistic depends on how the trading results are reported. At Banc of America securities, and probably at other firms, clients of the agency trading desk are provided with a post-trade report daily, weekly, or monthly depending on their trading activity.
2. Aggregation along the Reporting Dimensions: Typically these reports show sample average and standard deviation of execution price relative to the implementation shortfall benchmark across all the trades executed for that client during the reporting period. The results are further broken down into subsets across a dozen dimensions such as strategy type, buy or sell, primary exchange, trade sector, industry sector, market capitalization, *etc*.
3. Supra-Order Reporting Challenges: Because of these kinds of subsets, it is difficult to identify a larger unit than an individual order. Therefore it can be argued that the

broker-dealer's goal is to design algorithms that optimize sample mean and variance at the per-order level so that the post-trade report will be as favorable as possible.

4. <u>Order-Level Aggregation corresponds to Third Strategy</u>: As in the simple example this criterion translates to the third strategy above which is not optimized by the typical arrival price algorithms.
5. <u>Reporting Consistency with Client Goals</u>: Of course, the broker also has a responsibility to design the post-trade report so that it will be maximally useful to the client; that is, it corresponds as closely as possible to the client's investment goals.
6. <u>Execution Metrics under Finer Resolution</u>: One interpretation of these results is that the reports should show details with a finer resolution. For instance, it can show the mean and the variance of shortfall for each one thousand dollars of client money spent. The best choice of the reporting intervals is an open question.

Other Adaptive Strategies

1. <u>"Aggressive-in-the-money" AIM</u>: The new optimal strategies of Almgren and Lorenz (2007) are "aggressive-in-the-money" in the sense of Kissell and Malamut (2006); execution accelerates when the price moves in the traders' favor, and slows when the price moves adversely.
2. <u>"Passive-in-the-money" PIM</u>: A "passive-in-the-money" (PIM) strategy would react oppositely. Adaptive strategies of this form are called "scaling" strategies, and they can arise for a number of different reasons beyond those considered here.
3. <u>Traders' Preference and Prospect Theory</u>: A decrease in risk tolerance following a gain, and an increase following a loss, is consistent with traders' risk preferences (Shefrin and Statman (1985)) and id well-known in "prospect theory" (Kahneman and Tversky (1979)).
4. <u>Mathematical Foundations of Scaling Strategies</u>: Perhaps for these reasons scaling strategies often seem intuitively reasonable, though such qualitative preferences properly have no place in quantitative institutional trading. The formulation here is straightforward mean-variance optimization.

5. PIM as Optimal Momentum Strategy: One important reason for using an AIM or a PIM strategy would be the expectation of serial correlation in the price process. If the price is believed to have momentum, i.e., positive serial correlation, then a PIM strategy is optimal; if the price moves favorably, one should slow down to capture even more favorable prices in the future.
6. AIM fir Mean Reversion Optimality: Conversely if the price is believed to be mean-reverting, then favorable prices should be captured quickly before they mean-revert (Kissell and Malamut (2006)). The strategies presented in this chapter arise from pure random walk with no serial price correlation, using pure classic mean-variance optimization.
7. Caveats behind AIM/PIM Deployment: These models do provide an important caveat for their formulation. The AIM strategy suggest to "cut the gains and let the losses run".
8. Adverse Impact of "Moneyness" Guesses: If the price process does have any significant momentum, even on a small fraction of the real orders, then this strategy can cause much more serious losses than the gains that it provides. Thus implementing them in practice should be done only after doing extensive empirical tests.
9. The "Market Power" Parameter: The next section presents the market and the trading model, and shows the general importance of the "market power" parameter.
10. Single Update vs. Continuous Time: Two simple "proofs of concept" are considered; first a single update time, then a continuous response function that depends linearly on the asset price. The final section describes some approaches towards a full continuous time model.

The Market Model

1. Asset Price Arithmetic Random Walk: Trading in a single asset whose price is $S(t)$ is considered. $S(t)$ obeys the arithmetic random walk

$$S(t) = S_0 + \sigma B(t)$$

where $B(t)$ is a standard Brownian motion and σ is an absolute volatility. This process has neither momentum nor mean reversion; future price changes are completely independent of past changes.

2. Intra-day Profile Adapted Randomness: The Brownian motion $B(t)$ is the only source of randomness in the formulation. In the presence of intra-day seasonality t is interpreted as a time relative to a historical profile, and volume is assumed to constant under this transformation.
3. The Trade Order Execution Settings: The trader has an order of X shares, which begins at time

$$t = 0$$

and must be completed by time

$$t = T < \infty$$

X is taken to be

$$X > 0$$

and this in interpreted as a buy order. The benchmark value of this position at the start of the strategy is XS_0
4. The Estimation Output Trading Strategy: A *trading trajectory* is a function $x(t)$ with

$$x(0) = X$$

and

$$x(T) = 0$$

representing the number of shares remaining to buy at time t. For s static trajectory $x(t)$ is determined at

$$t = 0$$

but in general $x(t)$ may be any non-anticipating random functional of $B(t)$.

5. Observability of the Equilibrium Price: Permanent market impact is also important but has no effect on the optimal trade trajectory is it is linear – Almgren and Chriss (2000) carry out a detailed discussion of this model. The model parameters are assumed to be known with certainty and thus the underlying price $S(t)$ is observable based on the execution price $\tilde{S}(t)$ and the trade rate $v(t)$.
6. The Arrival Price Shortfall: The *implementation shortfall* $\mathcal{C}$ is the total cost of executing the buy program relative to its initial value.

$$\mathcal{C} = \int_0^T \tilde{S}(t)v(t)dt - XS_0 = \sigma \int_0^T x(t)dB(t) + \eta \int_0^T v^2(t)dt$$

7. Randomness of the Cost Components: The first term above represents the trading gains or losses. Since the trader is buying a positive price motion gives a positive cost. The second term represents the market impact cost. For an adaptive strategy both terms are random since $x(t)$ and $v(t)$ are both random.
8. Adaptive Strategies not necessarily Gaussian: Mean-variance optimization solves the problem

$$\min_{x(t)} \{\mathbb{E}[\mathcal{C}] + \lambda\mathbb{V}[\mathcal{C}]\}$$

for each

$$\lambda \geq 0$$

where $\mathbb{E}[\mathcal{C}]$ and $\mathbb{V}[\mathcal{C}]$ are the expected values of the mean and the variance of $\mathcal{C}$. As λ varies the resulting set of points $\{\mathbb{V}_{\lambda}[\mathcal{C}], \mathbb{E}_{\lambda}[\mathcal{C}]\}$ trace out an efficient frontier. For adaptive strategies $\mathcal{C}$ is not Gaussian, but mean-variance optimization is still used.

Static Trajectories

1. Non-random Optimal Execution Trajectory: If $x(t)$ is fixed independently of $B(t)$ then $\mathcal{C}$ is a Gaussian random variable with mean and variance

$$\mathbb{E}_{\lambda}[\mathcal{C}] = \eta \int_{0}^{T} v^{2}(t)dt$$

and

$$\mathbb{V}_{\lambda}[\mathcal{C}] = \sigma^{2} \int_{0}^{T} x^{2}(t)dt$$

2. Optimal Execution Trajectory Closed Form: The solution to

$$\min_{x(t)} \{\mathbb{E}[\mathcal{C}] + \lambda\mathbb{V}[\mathcal{C}]\}$$

is then obtained as

$$x(t) = Xh(t, T, \kappa)$$

where the static trajectory function is

$$h(t, T, \kappa) = \frac{\sinh[\kappa(T - t)]}{\sinh[\kappa T]}$$

for

$$0 \leq t \leq T$$

and the static *urgency* parameter is

$$\kappa = \sqrt{\frac{\lambda \sigma^2}{\eta}}$$

3. Portfolio Size Dependence of Urgency: The units of κ are inverse time, and $\frac{1}{\kappa}$ is the desired time scale for liquidation – the "half-life" as described in Almgren and Chriss (2000). The static trajectory is effectively an exponential with adjustments made to reach

$$x = 0$$

at

$$t = T$$

For a fixed λ the optimal time scale is independent of the portfolio size X since both the expected costs and the variance scale as X^2.

4. Static vs Dynamic Trajectory Equivalence: Equivalence of the static and the dynamic trajectories is demonstrated by observing that

$$h(t, T, \kappa) = h(s, T, \kappa) h(t - s, T - s, \kappa)$$

for

$$0 \le s \le t \le T$$

That is, the trajectory recomputed at time s, using the same urgency parameter, is the same as the tail of the original trajectory.

5. Low Urgency Limit VWAP Trading: By taking

$$\kappa \to 0$$

the linear profile

$$x(t) = X\frac{T - t}{T}$$

is recovered, which is equivalent to a VWAP profile under volume time transformation. The profile has expected cost

$$E_{LIN} = \frac{\eta X^2}{T}$$

and variance

$$V_{LIN} = \frac{\sigma^2 X^2 T}{3}$$

Non-dimensionalization

1. Dimensional Constants Determining the Solution: The optimal trajectory and the cost depend on 5 dimensional constants; the initial shares X, the time horizon T, the volatility σ, the impact coefficient η, and the risk aversion λ. To simplify the structure of the solution, it is convenient to define scaled variables.

2. <u>Non-dimensionalization of Horizon/Time</u>: Time is measured relative to T and shares relative to X. That is, the non-dimensional time is defined as

$$\hat{t} = \frac{t}{T}$$

and the non-dimensional holdings as

$$\hat{x}(\hat{t}) = \frac{x(\hat{t}T)}{X}$$

so that

$$0 \le \hat{t} \le 1$$

and

$$\hat{x}(0) = 1$$

The non-dimensional velocity is

$$\hat{v}(\hat{t}) = \frac{v(\hat{t}T)}{X/_T} = -\frac{d\hat{x}}{d\hat{t}}$$

3. <u>Non-dimensionalization of the Cost</u>: The cost is scaled by a dollar cost of a typical move due to volatility. That is, defining

$$\hat{\mathcal{C}} = \frac{\mathcal{C}}{\sigma X\sqrt{T}}$$

one then has

$$\hat{C} = \int_0^1 \hat{x}(\hat{t})d\hat{B}(\hat{t}) + \mu \int_0^1 \hat{v}^2(\hat{t})d\hat{t}$$

where

$$\hat{B}(\hat{t}) = \frac{B(\hat{t}T)}{\sqrt{T}}$$

and the "market power" parameter is

$$\mu = \frac{\eta X / T}{\sigma\sqrt{T}}$$

4. Definition of the "Market Power": Here the numerator is the price concession for trading at a constant rate, and the denominator is the typical size of the price motion due to volatility over the same period. The ratio μ is the non-dimensional preference free measure of the portfolio size, in terms of its ability to move the market.
5. "Market Power" Estimation – ATHL Model: To estimate realistic sizes for this parameter one recalls that Almgren, Thum, Hauptmann, and Li (2005) introduced the non-linear model

$$\frac{K}{\sigma} = \eta \left(\frac{X}{VT}\right)^{\alpha}$$

where K is the temporary impact (the only kind relevant here), σ is the daily volatility, X is the trade size, V is the average daily volume (ADV), and T is the fraction of the day over which the trade is executed.
6. ATHL Model "Market Power" Estimates: The coefficient was estimated as

$$\eta = 0.142$$

as was the exponent

$$\alpha = \frac{3}{5}$$

Therefore a 100% ADV executed across one full day gives

\

$$\mu = 0.142$$

7. "Market Power" Estimate Typical μ: Although the estimate above is only an approximate parallel to the linear model used here, it does suggest that for realistic trade sizes μ will be substantially smaller than one.
8. Non-dimensionalization of the Urgency: The problem

$$\min_{x(t)} \{\mathbb{E}[\mathcal{C}] + \lambda\mathbb{V}[\mathcal{C}]\}$$

has the scaled form

$$\min_{\hat{x}(\hat{t})} \{\mathbb{E}[\hat{\mathcal{C}}] + \mu\bar{\kappa}^2\mathbb{V}[\hat{\mathcal{C}}]\}$$

and the static urgency is

$$\bar{\kappa} = \kappa T$$

with κ from

$$\kappa = \sqrt{\frac{\lambda\sigma^2}{\eta}}$$

or

$$\bar{\kappa}^2 = \frac{\lambda\sigma^2 T^2}{\eta}$$

9. Non-dimensionalization of the Risk Aversion: The scaled risk aversion parameter $\mu\bar{\kappa}^2$ depends on X via the factor μ though the scaled time $\bar{\kappa}$ is independent of X.
10. Non dimensionalization of the Trajectory: $\bar{\kappa}$ will be used as the parameter to trace the frontier in place of λ. The result will be a trajectory $\hat{x}(\hat{t}; \bar{\kappa}, \mu)$ with the scaled cost values $\mathbb{E}\left[\hat{C}(\bar{\kappa}, \mu)\right]$ and $\mathbb{V}\left[\hat{C}(\bar{\kappa}, \mu)\right]$.
11. Non dimensionalization of Cost Distribution: For each of

$$\mu \geq 0$$

there will be an efficient frontier obtained by tracing $\mathbb{E}\left[\hat{C}(\bar{\kappa}, \mu)\right]$ and $\mathbb{V}\left[\hat{C}(\bar{\kappa}, \mu)\right]$ as functions of $\bar{\kappa}$ over

$$0 \leq \bar{\kappa} < \infty$$

The profile has expected cost

$$\hat{E}_{LIN} = \mu$$

and variance

$$\hat{V}_{LIN} = \frac{1}{3}$$

Small Portfolio Limit

1. Limit of Small "Market Power": Next the limit

$$\mu \to 0$$

is considered, keeping $\bar{\kappa}$ constant. Since X appears in μ but not in $\bar{\kappa}$ and all other dimensional variables do appear in $\bar{\kappa}$ this is equivalent to taking

$$X \to 0$$

with $T, \sigma, \eta,$ and λ fixed. Almgren and Lorenz (2006) show that for small portfolios status strategies are optimal.

2. Variance of the Non-dimensional Cost: When μ is small, assuming that $x(t)$ and $v(t)$ have reasonable limits, the second term in

$$\hat{C} = \int_0^1 \hat{x}(\hat{t}) d\hat{B}(\hat{t}) + \mu \int_0^1 \hat{v}^2(\hat{t}) d\hat{t}$$

is small compared to the first, and the variance of the non-dimensional cost is approximately

$$\mathbb{V}[\hat{C}] \sim \mathbb{V}\left[\int_0^1 \hat{x}(\hat{t}) d\hat{B}(\hat{t})\right] = \int_0^1 \mathbb{E}[\hat{x}^2(\hat{t})] d\hat{t}$$

$$\mu \to 0$$

3. Market Impact Contribution to Volatility: That is, the uncertainty in realized price comes primarily from the price volatility. Even if the strategy is adapted to the price process so that $\hat{x}(\hat{t})$ is random the market impact cost itself is a small number and the uncertainty in that number can be neglected next to the price volatility.
4. Expectation of Non-dimensional Cost: The first term in

$$\hat{C} = \int_0^1 \hat{x}(\hat{t}) d\hat{B}(\hat{t}) + \mu \int_0^1 \hat{v}^2(\hat{t}) d\hat{t}$$

has strictly zero expected value for any non-anticipating strategy – it is an Ito integral – and hence the expectation comes entirely from the second term.

5. <u>Corresponding Non-dimensional Objective Utility</u>: Thus

$$\mathbb{E}[\hat{C}] = \mu \mathbb{E}\left[\int_0^1 \hat{v}^2(\hat{t}) d\hat{t}\right]$$

and the complete risk-aversion cost function is approximately

$$\mathbb{E}[\hat{C}] + \mu \bar{\kappa}^2 \mathbb{V}[\hat{C}] \sim \mu \int_0^1 \mathbb{E}[\hat{v}^2(\hat{t}) + \bar{\kappa}^2 \hat{x}^2(\hat{t})] d\hat{t}$$

$$\mu \to 0$$

6. <u>Quadratic Nature of Objective Utility</u>: Consider a candidate adaptive strategy $\hat{x}(\hat{t})$. Since the quadratic is convex the static strategy

$$\bar{x}(\hat{t}) = \mathbb{E}[\hat{x}(\hat{t})]$$

will give a lower value of the objective function (thus $x(t)$ and $v(t)$ have limits, thereby justifying the original assumption).

7. <u>Consequence of the "Market Power"</u>: When μ is not small adaptive strategies can create negative correlation between the two terms in

$$\hat{C} = \int_0^1 \hat{x}(\hat{t}) d\hat{B}(\hat{t}) + \mu \int_0^1 \hat{v}^2(\hat{t}) d\hat{t}$$

thereby reducing the overall variance below its value for purely static strategies.

Portfolio Comparison

1. Optimal Strategy Portfolio Size Dependence: In the simplest form, the goal is to determine the optimal strategy $\hat{x}(\hat{t})$ for any specific set of parameters. But to understand results it is useful to compare strategies and costs for portfolios of different sizes.
2. Quadratic Scaling of Static Trajectories: Consider two portfolios X_1 and X_2 with

$$X_2 = 2X_1$$

and all other parameters the same including risk aversion; thus

$$\mu_2 = 2\mu_1$$

and $\bar{\kappa}$ is the same. Portfolio X_2 will in general cost four times to trade as much as portfolio X_1. For example, static trajectories for the two portfolios will have identical shapes, and the cost will satisfy

$$\mathbb{E}\left[\hat{\mathcal{C}}_2\right] = 4\mathbb{E}\left[\hat{\mathcal{C}}_1\right]$$

and

$$\mathbb{V}\left[\hat{\mathcal{C}}_2\right] = 4\mathbb{V}\left[\hat{\mathcal{C}}_1\right]$$

3. Adaptive Strategies: Sub-Quadratic Scaling: For adaptive strategies, the large portfolio is still more expensive to trade than the small portfolio, but it can take advantage of the negative correlation. Thus one will have

$$\mathbb{E}[\hat{C}_2] + \lambda \mathbb{V}[\hat{C}_2] \leq 4(\mathbb{E}[\hat{C}_1] + \lambda \mathbb{V}[\hat{C}_1])$$

for each λ although it is generally not true that separately

$$\mathbb{E}[\hat{C}_2] < 4\mathbb{E}[\hat{C}_1]$$

AND

$$\mathbb{V}[\hat{C}_2] < 4\mathbb{V}[\hat{C}_1]$$

4. Cost Ratio - Adaptive to Static: The ratio of an adaptive cost to a static cost will be less for a large portfolio than for a small portfolio, though all costs are higher for the large portfolio. Therefore these solutions will be of most interest to the large investors.
5. Representative Illustration of the Relative Costs: To highlight the differences in relative costs, Almgren and Lorenz (2007) draw efficient frontiers which show the expectation of the cost and its variance *relative* to their values for the linear trajectory.
6. Static Strategies correspond to $\mu = 0$: Thus the static efficient frontiers for all the values of

$$\mu > 0$$

super-impose, since the cost of all static trajectories scale precisely as X^2. This common static frontier appears as the limit of adaptive frontiers as

$$\mu \to 0$$

As μ increases the adaptive frontiers move down and to the left, away from the static frontier.

Single Update

1. <u>Non-dimensional Decision Time Instant</u>: Here the assumption is that the urgency update occurs at a single update time $\hat{T}_*$ where

$$0 < \hat{T}_* < 1$$

2. <u>Starting with the Initial Urgency</u>: On the first trading period

$$0 < \hat{t} < \hat{T}_*$$

an initial urgency $\bar{\kappa}_0$ is used, that is, the trajectory is

$$\hat{x}(\hat{t}) = h(\hat{t}, 1, \bar{\kappa}_0)$$

with h from

$$h(\hat{t}, \hat{T}, \bar{\kappa}) = \frac{\sinh\left[\bar{\kappa}\left(\hat{T} - \hat{t}\right)\right]}{\sinh\left[\bar{\kappa}\hat{T}\right]}$$

for

$$0 < \hat{t} < \hat{T}_*$$

3. <u>The Set of Decision Urgencies</u>: Let

$$\hat{x}_*\left(\bar{\kappa}_0, \hat{T}_*\right) = h\left(\hat{T}_*, 1, \bar{\kappa}_0\right)$$

be the shares remaining at the decision time. At time $\hat{T}_*$ one switches to one of the n new urgencies $\bar{\kappa}_1, \cdots, \bar{\kappa}_n$; with urgency $\bar{\kappa}_i$ one sets

$$\hat{x}(\hat{t}) = \hat{x}_*(\bar{\kappa}_0, \hat{T}_*) h(\hat{t} - \hat{T}_*, 1 - \hat{T}_*, \bar{\kappa}_i)$$

for

$$\hat{T}_* < \hat{t} < 1$$

4. Urgency Based on Realized Cost: The new urgency is chosen based on the non-dimensional realized cost up to time $\hat{T}_*$:

$$\hat{C}_0 = \int_0^{\hat{T}_*} \hat{x}(\hat{t}) d\hat{B}(\hat{t}) + \mu \int_0^{\hat{T}_*} \hat{v}^2(\hat{t}) d\hat{t} + \int_{\hat{T}_*}^1 [\hat{B}(\hat{t}) + \mu \hat{v}(\hat{t})] \hat{v}(\hat{t}) d\hat{t} + \hat{x}_* \hat{B}(\hat{T}_*)$$

5. Trajectory Cost Decomposition - Terms Explain: To measure $\hat{C}_0$ at time $\hat{T}_*$, as can be seen in the second expression above, the first term is the total dollar cost paid to acquire the shares so far, minus the value of those shares at the pre-trade price.
6. Trading Cost Decomposition - Position Remaining: The second term is the estimation of the additional cost that will need to be paid on the remaining shares relative to the pre-trade price, due to price movements observed so far.
7. Observability of the Realized Price Brownian: As noted before $\hat{B}(\hat{t})$ is observable if the execution price, the trade rate, and the coefficient of the market impact are all known.
8. Outcome Partitioning at the Decision Instant: The real-line is partitioned into n intervals $I_1, \cdots, I_n$ and $\bar{\kappa}_j$ is used if

$$\hat{C}_0 \in I_j$$

For large n this approaches a continuous dependence

$$\bar{\kappa} = f\left(\hat{C}_0\right)$$

9. Price vs Cost Bound $\bar{\kappa}$: The intuition seen in the Introduction section indicates that using the accumulated cost should be more effective than using the instantaneous price at time $\hat{T}_*$.
10. Pre-fixing Decision Time Urgencies: Before trading begins, the decision time $\hat{T}_*$, the interval break-points, and the $n + 1$ urgencies $\bar{\kappa}_1, \cdots, \bar{\kappa}_n$ are all fixed. However it is not known which trajectory shall actually be executed until $\hat{C}_0$ is observed at time $\hat{T}_*$.
11. Cost of the Decision Trajectory: $\hat{C}_j$ is the cost incurred in the second part of the trajectory if urgency $\bar{\kappa}_j$ is used:

$$\hat{C}_j = \int_{\hat{T}_*}^{1} \hat{x}(\hat{t}) d\hat{B}(\hat{t}) + \mu \int_{\hat{T}_*}^{1} \hat{v}^2(\hat{t}) d\hat{t} + \int_{\hat{T}_*}^{1} \left[\hat{B}(\hat{t}) + \mu \hat{v}(\hat{t})\right] \hat{v}(\hat{t}) d\hat{t} - \hat{x}_* \hat{B}\left(\hat{T}_*\right)$$

12. Total Cost across the Trajectory: The total cost is then

$$\hat{C} = \hat{C}_j + \hat{C}_{\mathcal{J}(\hat{C}_0)}$$

where

$$\mathcal{J}\left(\hat{C}_0\right) = i$$

if

$$\hat{C}_0 \in I_i$$

Although the total cost is not Gaussian the optimal frontier is still computed using mean-variance optimization.

Single Update Mean and Variance

1. Initial Time Decision Mean/Variance: As described before the mean and the variance are calculated at the initial time. Each variable $\hat{\mathcal{C}}_i$ is Gaussian with mean

$$E_i = \mu F_i$$

and variance V_i where F_i and V_i are integrals of the form

$$F_i = \int_{\hat{T}_*}^{1} \hat{v}^2(\hat{t}) d\hat{t}$$

and

$$V_i = \int_{\hat{T}_*}^{1} \hat{x}^2(\hat{t}) d\hat{t}$$

which do not depend on μ.

2. Initial and Decision Marginal Distributions: Next the mean and the variance of each possible decision trajectory cost needs to be evaluated.
3. Decision Cost Trajectory Mean and Variance: The following integrals ar4e readily determined:

$$F_0 = \bar{\kappa}_0 \frac{\sinh(2\bar{\kappa}_0) - \sinh\left[2\bar{\kappa}_0\left(1 - \hat{T}_*\right)\right] + 2\bar{\kappa}_0 \hat{T}_*}{4\sinh^2 \bar{\kappa}_0}$$

$$V_0 = \frac{\sinh(2\bar{\kappa}_0) - \sinh\left[2\bar{\kappa}_0\left(1 - \hat{T}_*\right)\right] - 2\bar{\kappa}_0 \hat{T}_*}{4\sinh^2 \bar{\kappa}_0}$$

and

$$F_i = \frac{\sinh^2\left[\bar{\kappa}_0\left(1-\hat{T}_*\right)\right]}{\sinh^2\left[\bar{\kappa}_i\left(1-\hat{T}_*\right)\right]}\bar{\kappa}_i\frac{\sinh\left[2\bar{\kappa}_i\left(1-\hat{T}_*\right)\right]+2\bar{\kappa}_i\left(1-\hat{T}_*\right)}{4\sinh^2\bar{\kappa}_0}$$

and

$$V_i = \frac{\sinh^2\left[\bar{\kappa}_0\left(1-\hat{T}_*\right)\right]}{\sinh^2\left[\bar{\kappa}_i\left(1-\hat{T}_*\right)\right]}\frac{\sinh\left[2\bar{\kappa}_i\left(1-\hat{T}_*\right)\right]-2\bar{\kappa}_i\left(1-\hat{T}_*\right)}{4\sinh^2\bar{\kappa}_0}$$

for

$$i = 1, \cdots, n$$

4. Trajectory Cost Distribution Density Expression: Each $\hat{C}_i$ is a Gaussian with a mean F_i and a variance V_i so its density is

$$f_i\left(\hat{C}_i\right) = \frac{1}{\sqrt{2\pi V_i}}e^{-\frac{\left(F_i-\hat{C}_i\right)^2}{2V_i}}$$

$$i = 1, \cdots, n$$

5. Partitioning the Cost Decision Space: The intervals are defined as

$$I_i = \left\{b_{j-1} < \hat{C}_0 < b_j\right\}$$

with

$$b_j = E_0 + a_j\sqrt{V_0}$$

where $a_0, \cdots, a_n$ are fixed constants with

$$a_0 = -\infty$$

and

$$a_n = +\infty$$

6. Cost Convolution over Decision Segments:

$$f(c)\Delta c = Prob\{\hat{C} \in [c, c + \Delta c]\}$$
$$= \sum_{i=1}^{n} Prob\{\hat{C}_0 \in I_i \; AND \; \hat{C}_i \in [c - \hat{C}_0, c - \hat{C}_0 + \Delta c]\}$$

so

$$f(\hat{C}) = \sum_{i=1}^{n} \int_{b_{i-1}}^{b_i} f(\hat{C}_0) f(\hat{C} - \hat{C}_0) d\hat{C}_0$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi V_0 V_i}} \int_{b_{i-1}}^{b_i} e^{-\left[\frac{(\hat{C}_0 - E_0)^2}{2V_0} + \frac{(\hat{C} - \hat{C}_0 - E_i)^2}{2V_i}\right]} d\hat{C}_0$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi V_0 V_i}} e^{-\frac{1}{2}\left[\frac{E_0{}^2}{V_0} + \frac{(\hat{C} - E_i)^2}{V_i} - \frac{\{E_0 V_i + (\hat{C} - E_i) V_0\}^2}{V_0 V_i (V_0 + V_i)}\right]} \int_{b_{i-1}}^{b_i} e^{-\frac{1}{2}\frac{V_0 + V_i}{V_0 V_i}\left[\hat{C}_0 - \frac{\{E_0 V_i + (\hat{C} - E_i) V_0\}^2}{V_0 + V_i}\right]} d\hat{C}_0$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi (V_0 + V_i)}} e^{-\frac{(\hat{C} - \hat{C}_0 - E_i)^2}{2V_i}}$$
$$\times \left[\Phi\left(\frac{\{\hat{C} - E_i - b_{i-1}\} V_0 + \{E_0 - b_{i-1}\} V_i}{\sqrt{V_0 V_i (V_0 + V_i)}}\right) - \Phi\left(\frac{\{\hat{C} - E_i - b_i\} V_0 + \{E_0 - b_i\} V_i}{\sqrt{V_0 V_i (V_0 + V_i)}}\right)\right]$$

7. Incremental Cost Distribution and Density: To calculate the mean and the variance of the composite cost $\hat{\mathcal{C}}$ the following non-dimensional fixed costs are defined.

$$p_j = \Phi(a_j) - \Phi(a_{j-1})$$

and

$$q_j = \phi(a_{j-1}) - \phi(a_j)$$

for

$$j = 1, \cdots, n$$

ϕ is the standard normal density, and Φ is its cumulative. Thus

$$Prob\{\hat{\mathcal{C}}_0 \in I_i\} = p_j$$

and

$$\mathbb{E}[\hat{\mathcal{C}}_0 | \hat{\mathcal{C}}_0 \in I_i] = E_0 + \frac{q_j}{p_j}\sqrt{V_0}$$

8. Total Cost Mean and Variance: By linearity of expectation one readily gets

$$E = \mu(F_0 + \bar{F})$$

with

$$\bar{F} = \sum p_i F_i$$

The variance is more complicated because of the dependence between the two terms in

$$\hat{C} = \hat{C}_0 + \hat{C}_{\mathcal{J}(\hat{C}_0)}$$

9. <u>The Full Trajectory Cost Variance</u>: One uses the conditional variance expression

$$\mathbb{V}[X] = \mathbb{E}\big[\mathbb{V}[X|Y]\big] + \mathbb{V}\big[\mathbb{E}[X|Y]\big]$$

to write, using

$$\bar{V} = \sum p_i V_i$$

$$\begin{aligned}
\mathbb{V}[\hat{C}] &= \mathbb{E}\left[\mathbb{V}\left[\hat{C}_0 + \hat{C}_{\mathcal{J}(\hat{C}_0)}|\hat{C}_0\right]\right] + \mathbb{V}\left[\mathbb{E}\left[\hat{C}_0 + \hat{C}_{\mathcal{J}(\hat{C}_0)}|\hat{C}_0\right]\right] \\
&= \mathbb{E}\left[\mathbb{V}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right]\right] + \mathbb{V}\left[\hat{C}_0 + \mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right]\right] \\
&= \bar{V} + \mathbb{V}\big[\hat{C}_0\big] + 2\ Covar\left[\hat{C}_0, \hat{C}_{\mathcal{J}(\hat{C}_0)}\right] + \mathbb{V}\left[\mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right]\right]
\end{aligned}$$

10. <u>Full Trajectory Cost Variance Components</u>: By definition

$$\mathbb{V}\big[\hat{C}_0\big] = V_0$$

and

$$\mathbb{V}\left[\mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right]\right] = \mu^2 \sum p_i (F_i - \bar{F})^2$$

Further

$$
\begin{aligned}
Covar\left[\hat{C}_0, \hat{C}_{\mathcal{J}(\hat{C}_0)}\right] &= \mathbb{E}\left[\hat{C}_0 \mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right]\right] - \mathbb{E}\left[\hat{C}_0\right]\mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right] \\
&= \sum Prob\{\hat{C}_0 \in I_i\}\mathbb{E}\left[\hat{C}_0 \mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right] | \hat{C}_0 \in I_i\right] - \mathbb{E}\left[\hat{C}_0\right]\mathbb{E}\left[\hat{C}_{\mathcal{J}(\hat{C}_0)}\right] \\
&= \mu\sqrt{V_0}\sum q_i F_i
\end{aligned}
$$

11. Bringing all the Parts Together: Putting all this together one gets

$$
V = \mathbb{V}\left[\hat{C}\right] = V_0 + \bar{V} + \mu\sqrt{V_0}\sum q_i F_i + \mu^2 \sum p_i (F_i - \bar{F})^2
$$

12. The Non-dimensional Objective Function: The overall objective function is

$$
U = \frac{E + \mu\bar{\kappa}^2 V}{\mu}
$$

or

$$
\begin{aligned}
&U\left(\bar{\kappa}_1, \cdots, \bar{\kappa}_n, \hat{T}_*; \bar{\kappa}, \mu\right) \\
&\quad = F_0 + \bar{F} + \bar{\kappa}^2(V_0 + \bar{V}) + \mu\bar{\kappa}^2\sqrt{V_0}\sum q_i F_i + \mu^2\bar{\kappa}^2 \sum p_i (F_i - \bar{F})^2
\end{aligned}
$$

13. Negative Two Period Cross Correlation: The $\mathcal{O}(\mu)$ term is approximately $2\sum\left(\hat{C}_0 - E_0\right)p_i E_i$ and can be made negative by making E_i negatively related to $\hat{C}_0$ corresponding to anti-correlation between second period impact costs and the first period trading losses.
14. The Optimizer Input/Search Space: For a given market power μ and static urgency $\bar{\kappa}$, U is minimized numerically over $\bar{\kappa}_0, \cdots, \bar{\kappa}_n$ and the decision time $\hat{T}_*$
15. Efficient Frontier Curve over $\bar{\kappa}$: As $\bar{\kappa}$ varies the resulting set of points (V, E) traces the efficient frontier. This results in a one-parameter family of efficient frontiers, depending on μ. The static trajectories appear at the limit of

$$\mu = 0$$

Almgren and Lorenz (2007) Results

1. Decision Urgency Based Efficient Frontier: Almgren and Lorenz (2000) illustrate the complete set of efficient frontiers for the single update problem. Each curve is computed by varying a static urgency parameter $\bar{\kappa}$ from 0 to ∞ for a fixed value of μ.
2. Discretization of the Decision Urgency: The solution for each pair of $(\bar{\kappa}, \mu)$ is computed using a fixed set of 32 equal-probability breakpoints. As described earlier $\hat{E}$ and $\hat{V}$ are plotted relative to their values for the linear trajectories to clearly see the improvement due to the adaptivity.
3. Improved Execution Strategy Cost Distribution: The frontiers are used to obtain adaptive strategies that are better than the cost distribution for any static strategies.
4. Static Urgency Trajectory and Cost: First Almgren and Lorenz (2007) compute a static trajectory using

$$\bar{\kappa} = 8$$

and generate the resulting cost distribution that is Gaussian. For a portfolio with

$$\mu = 0.1$$

this distribution has an expectation

$$\hat{E} \approx 4 \times \hat{E}_{LIN} \approx 4 \times \mu = 0.4$$

and variance

$$\hat{V} \approx 0.2 \times \hat{V}_{LIN} = \frac{0.2}{3} = 0.067$$

5. <u>Market Power Based Efficient Frontier</u>: Likewise they generate adaptive efficient frontiers for different values of the market power μ. They identify the set of values accessible to a static strategy as well as the static frontier – which is also the limit

$$\mu \to 0$$

with a static strategy $\bar{\kappa}$. The improved values accessible to the adaptive strategies are also identified; the improvement is greater for larger portfolios. The actual cost distributions corresponding to different $\bar{\kappa}$ are also estimated.

6. <u>Improvement available over the Static Trajectory</u>: The region in the $(\hat{V}, \hat{E})$ space accessible to an adaptive strategy with

$$\mu = 0.1$$

that are strictly preferable to a static strategy since they have lower expected cost and/or variance can be readily observed.

7. <u>Cost Profile Adaptive Urgency Range</u>: On the efficient frontier for

$$\mu = 0.1$$

these solutions are obtained by computing adaptive solutions with parameters approximately in the range

$$4.9 \leq \bar{\kappa} \leq 7.1$$

There is no need to use the same value for $\bar{\kappa}$ for the adaptive strategy as for the static strategy to which it is compared.

8. <u>Adaptive Trajectory Urgency - Cost/Variance</u>: Detailed cost distributions associated with these adaptive strategies can also be generated. For

$$\bar{\kappa} = 4.9$$

the adaptive distribution has a lower expected cost than the static distribution with the same variance. For

$$\bar{\kappa} = 7.1$$

the adaptive distribution has a lower variance than the static distribution with the same mean.

9. Strictly Optimal Adaptive Strategy Urgency: These distributions are the extreme points of a one-parameter family of distributions, each of which is strictly preferable to the given static strategy, regardless of the traders' risk preferences. For example the adaptive solution for

$$\bar{\kappa} = 6$$

has both lower expected cost and lower variance than the static distribution.

10. Strongly Positive Optimal Distribution Skew: These cost distributions are strongly skewed toward positive distribution costs suggesting that the mean-variance optimization may not give the best possible solutions.

11. Realized Static vs. Adaptive Trajectories: Almgren and Lorenz (2007) compare the adaptive trading trajectories for

$$\mu = 0.1$$

and

$$\bar{\kappa} = 6$$

against the static optimal trajectory with urgency

$$\bar{\kappa} = 8$$

The adaptive strategy clearly delivers both lower expectation of cost and lower variance.

12. Urgency Dependence on Trading Cost: They also demonstrate the dependence of the decision urgency on the initial trading cost $\hat{C}_0$ - in their plots they normalize $\hat{C}_0$ by its initial *ex ante* expectation and trading cost.
13. Adaptation under Favorable Price Move: The adaptive strategy initially trades more slowly than the optimal static trajectory. At $\hat{T}_*$, if the prices have moved in the traders' favor, the adaptive strategy accelerates, spending the investment gains on the impact costs.
14. Adaptation under Unfavorable Price Move: If the prices have moved against the trader, corresponding to positive values of $\hat{C}_0$, then the strategy decelerates to save impact costs in the remaining period. The values of $\bar{\kappa}$ become very large when $\hat{C}_0$ is large negative, corresponding to the instruction: "if you have gains in the first part of the trading, then finish the program immediately".

Continuous Response

1. $\mathbb{R}^1 \to \mathbb{R}^1$ Dependence on Brownian: Next Almgren and Lorenz (2007) illustrate a simple form of *continuous response* to trading gains or losses. In general one can specify any rule $\hat{v}(\hat{t})$ as a function of the price history $\hat{B}(s)$ for

$$0 \leq s \leq \hat{t}$$

   Rather than adjusting the rate $\hat{v}(\hat{t})$ directly it is more convenient to adjust $\bar{\kappa}$.
2. Trade Rate Explicit Functional Form: From

$$h(t, T, \kappa) = \frac{\sinh[\kappa(T - t)]}{\sinh[\kappa T]}$$

for

$$0 \leq t \leq T$$

on differentiating

$$x(s) = x(t)h(s - t, 1 - t, \kappa)$$

with respect to s and evaluating at

$$s = t$$

the following relationship between v and κ is obtained.

$$v(t) = x(t)\kappa(t) \coth[\kappa(t)(1 - t)]$$

For all choices of $\kappa(t)$ the trajectories hit

$$x = 0$$

at

$$t = 1$$

3. <u>Exponential Price Brownian Functional Form</u>: Determining the full optimal dependence of $\kappa(t)$ on $B(s)$ for

$$0 \leq s \leq t$$

is difficult. Thus the following relationship is considered:

$$\kappa(t) = ae^{bB(t)}$$

Thus the instantaneous urgency depends on the instantaneous price level. Other functional relationships for $\kappa(t)$ in terms of $B(t)$ are possible as well. Here $\kappa(t)$ is always positive, and is monotone in $B(t)$.

4. Corresponding Trade Rate Shortfall: From

$$v(t) = x(t)\kappa(t)\coth[\kappa(t)(1-t)]$$

one readily obtains $x(t)$ and finally the shortfall $\mathcal{C}$ by integration as in

$$\mathcal{C} = \sigma\int_0^T x(t)dB(t) + \eta\int_0^T v^2(t)dt$$

5. Need for Numerical Framework: However, because of the highly nonlinear dependence of $\kappa(t)$, and thus $v(t)$ and $x(t)$, on the Brownian motion $B(t)$, analytic solution of this stochastic integral is beyond reach.

Continuous Response Numerical Results

1. Price Move Brownian Bridge Construction: For numerical solutions one generates a fixed collection of sample paths using a Browning bridge construction with quasi-random variables.
2. Objective Value Function Numerical Evaluation: For any candidate values of a and b the stochastic integrals are evaluated numerically, and the sample mean E and the variance V are calculated. The objective function $E + \bar{\kappa}^2\mu^2V$ is then numerically minimized over a and b.
3. Generation of the Efficient Frontier: By solving for a series of values of

$$0 < \bar{\kappa} < \infty$$

the efficient frontier can again be traced for different values of μ, yielding similar results as in the single update framework.

4. Execution Cost Gain/Loss Adaptation: Again the optimal strategies are "aggressive in the money", having

$$b < 0$$

When the stock price goes down, an unexpected smaller shortfall is incurred, and a reaction occurs with increasing urgency $\bar{\kappa}(t)$ whereas for rising stock prices the trading is slowed down.

5. Exponential Urgency Response Trajectory Sample: As an illustration, Almgren and Lorenz (2007) generate optimal trading trajectories using the adaptation rule

$$\bar{\kappa}(t) = ae^{bB(t)}$$

with

$$a = 5.9$$

and

$$b = -1.7$$

for a static urgency

$$\bar{\kappa} = 6$$

As the stock price goes down the trading is accelerated compared to the optimal static trajectory, whereas for rising stock price it is slowed down.

Discussion and Conclusions

1. Rule Based Adaptive Scaling Strategies: The simple update rules presented in the previous sections demonstrate that price adaptive scaling strategies can lead to significant improvements over static trade schedules, and illustrate the importance of the market power parameter μ.
2. Dynamic Programming Based Optimal Trajectory: However neither of these rules is the fully adaptive optimal trading strategy. A fully optimal adaptive trading strategy would use stochastic dynamic programming to determine the trading rate as a general function of the continuous state variables such as the number of shares remaining, time remaining, current stock price, and trading gains or losses experienced to date.
3. Infeasibility of Mean Variance Optimization: One subtlety is that the mean-variance optimization cannot be used directly in this context; it involves the square of an expectation, which is not amenable to dynamics programming techniques.
4. Quadratic Utility Family of Optimization: However Li and Ng (2000) have shown how to embed mean-variance optimization into a family of optimizations that use the quadratic utility function.
5. MVO as a Family Member: The mean-variance solution is recovered as one element of this family. The need to solve this family of problems is an additional degree of complication.
6. HJB PDE Based Stochastic Control: The calculation uses tools of optimal stochastic control and requires the numerical solution of a highly nonlinear Hamilton-Jacobi-Bellman partial differential equation.
7. Adaptive Strategies as a Simplifier: Partial formulation of this problem, and the solution of the resulting equations, is an involved undertaking and the focus of a later chapter. The examples shown here show that even with very simple adaptive strategies substantial improvement is possible over static strategies.

References

- Almgren, R. F., and N. Chriss (1999): Value under Liquidation *Risk* **12 (12)** 61-63.
- Almgren, R. F., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.
- Almgren, R. F., and J. Lorenz (2007): Adaptive Arrival Price, in: *Algorithmic Trading III (B. R. Bruce, editor)* **Institutional Investor** 59-66.
- Huberman, G., and W. Stanzl (2005): Optimal Liquidity Trading *Review of Finance* **9 (2)** 165-200.
- Kahneman, D., and A. Tversky (1979): Prospect Theory: An Analysis of Decision under Risk *Econometrica* **47 (2)** 263-291.
- Kissell, R., and R. Malamut (2006): Algorithmic Decision-Making Framework *Journal of Trading* **1 (1)** 12-21.
- Li, D., and W. L. Ng (2000): Optimal Dynamic Portfolio Selection: Multi-period Mean-Variance Formulation *Mathematical Formulation* **10 (3)** 387-406.
- Perold, A. F. (1988): The Implementation Shortfall: Paper versus Reality *Journal of Portfolio Management* **14 (3)** 4-9.
- Shefrin, H., and M. Statman (1985): The Disposition to sell Winners too early and ride Losers too Long: Theory and Evidence *Journal of Finance* **40 (3)** 777-790.

Optimal Trading in a Dynamic Market

Introduction, Overview, and Motivation

1. Stock Market Liquidity and Volatility: This chapter considers the problem of mean-variance agency execution strategies when the market volatility and the liquidity vary randomly in time.
2. Cost/Strategy Dynamic Optimal Trajectory: Under specific assumptions for the stochastic processes satisfied by these parameters, a Hamilton-Jacobi-Bellman equation is presented for the optimal cost and the strategy.
3. Trajectory Adoption to Market Conditions: This equation is solved numerically, and optimal strategies are illustrated for varying risk aversion. These strategies adapt optimally to the instantaneous variations of market quality.

Limitations of Arrival Price Frameworks

1. Optimal Order Execution Trade Scheduling: A fundamental part of agency algorithmic trading in equities and other asset classes is trade scheduling. Given a trade target, that is a number of shares of a trade that must be bought or sold before a fixed time horizon, trade scheduling refers to how many shares that will be bought or sold by each time instant between the beginning of trading and the horizon.
2. Optimal Measure of Execution Quality: The optimal execution is done so as to optimize some measure of execution quality, usually measured as the final average execution price relative to some benchmark price.
3. Definition of the Arrival Price: One of the most popular benchmarks is the "arrival price", i.e., the price prevailing in the market the time the order was received into the

trading system. The difference between the execution price and this pre-trade price is the *implementation shortfall* (Perold (1988)) or *slippage*.

4. Balancing Transaction Costs and Variance: Grinold and Kahn (1995) and Almgren and Chriss (2000) suggested that the optimal trajectory could be determined by balancing the market impact cost, which leads toward slow trading, versus volatility risk, which leads toward rapid completion of the order.
5. Risk Aversion Based Efficient Frontier: This framework leads to an efficient frontier in which the trade schedule is selected from a one-parameter family based on a risk-aversion parameter that must be specified by the trading client.
6. Risk Aversion Based Front Loading: Optimal trading strategies are typically *front-loaded*. They execute as much as possible early in the program to reduce risk relative to the benchmark price. The degree of front-loading depends on the risk aversion parameter that must be specified by the trading client. The exact shape of the schedule depends on the form of the market impact model.
7. Market Price Based Benchmarks: The largest alternative category of benchmarks is composed of some form of average market price during the trading interval; usually either time-weighted average price (TWAP) or volume weighted average price (VWAP).
8. Close Tailing of the Benchmarks: For these benchmarks optimal strategy follows the benchmarks quiet closely, since deviation from the profile both increases the risk relative to the benchmark and the impact costs. Determining optimal response to the short-term price is an interesting topic for optimization, but is not the focus here.
9. Use of Arrival Price Frameworks: While other factors such as anticipated price drift, serial correlation or other short-term signals, and daily patterns are certainly important, this fundamental "arrival price" framework has proven remarkably robust and useful in designing practical trading systems.
10. Time Profiles of Liquidity/Volatility: A fundamental assumption of most of this work has been that the market parameters are constant, or at least have known predictable profiles. This assumption is reasonably accurate for large-cap US stocks.
11. Static Nature of the Framework: Under that assumptions optimal strategies are *static*; that is the trade schedule can be determined before the trading starts and is not

modified by the new information revealed by price moves during trading. Almgren and Chriss (2000) did consider a model in which the market parameters updated at a single time to one of a known set of possible new values.

12. Algorithmic Trading of Less Liquid Assets: over the last few years, a major push of providers of algorithmic trading services has been to extend their functionality to smaller stocks and less liquid assets.
13. Random Intra-day Volatility/Liquidity: A distinguishing feature of these assets is that their liquidity and volatility vary randomly in time. That is, there will be times during the trading day when trading becomes very expensive, and times when trading is cheap; similarly there will be times when delaying trading introduces large amounts of volatility risk and other times when the delay is relatively costless.
14. Optimal Mean-Variance Tradeoff: The modeling challenge is to determine optimal strategies that adapt to the instantaneous market state, while retaining the mean-variance trade-off inherent in the arrival price framework.
15. Continuous Time and State Treatment: Walia (2006) solved this problem in a discrete time discrete state model. Almgren (2009) provides a systematic mathematical solution to the problem in continuous time and continuous state.
16. Coordinated Liquidity/Volatility Joint Moves: The first section of this chapter presents the basic price and the impact models used, and presents the optimal trading problem. Also presented is the "coordinated variation" approximation in which the liquidity and the volatility vary together, which is very realistic and greatly simplifies the mathematical problem.
17. Dynamic HJB Optimal Cost Function: The second section uses a Hamilton-Jacobi-Bellman PDE describing the optimal cost function and the trade rate. The third section examines some aspects of the numerical solution to this PDE, and presents example solutions.

The Liquidation Problem

1. <u>The Continuous Holdings Rate Trajectory</u>: The trader begins trading at a time

$$t = 0$$

with a purchase order of X shares which must be completed by

$$t = T$$

The number of shares remaining to purchase at the time t is the remaining trajectory $x(t)$ with

$$x(0) = X$$

and

$$x(T) = 0$$

The rate of buying is

$$v(t) = -\frac{dx(t)}{dt}$$

2. <u>The Trajectory as a Random Variable</u>: Thus for a buy program

$$X > 0$$

$$x(t) \geq 0$$

and decreasing, and

$$v(t) \geq 0$$

– sell program may be modeled similarly. In general, the trajectory conditions $x(t)$ may be determined depending on price motions and market conditions discovered during trading, so it is a random variable.

3. The Arithmetic Brownian Price Dynamics: The price $S(t)$ follows the arithmetic Brownian motion

$$\Delta S(t) = \sigma(t)\Delta B(t)$$

$$S(0) = S_0$$

   where $B(t)$ is a standard Brownian motion, and the instantaneous volatility $\sigma(t)$ depends on time either deterministically or stochastically.

4. Volatility and Permanent Impact Parameters: Note that $\sigma(t)$ is an absolute volatility rather than fractional; it contains an implicit factor of the reference price S_0 It is possible to include the permanent impact into the price equation, but it is not central to the problem.
5. Execution Price - Incorporating the Temporary Impact: The price actually received on each trade is

$$\tilde{S}(t) = S(t) + \eta(t)v(t)$$

   where $\eta(t)$ is the coefficient of the temporary market impact, also time varying. Again $\eta(t)$ is an absolute coefficient rather than fractional.

6. More Elaborate Market Impact Models: Much richer market impact models have been considered in the literature (Gatheral (2010)), but this simple one is adequate to highlight the response to stochastic liquidity.
7. Estimation of $\sigma(t)$ and $\eta(t)$: Both $\sigma(t)$ and $\eta(t)$ are assumed to be observable in real-time with some degree of confidence. There is a variety of techniques available for doing this estimation.

8. Techniques for Estimation of $\sigma(t)$: For volatility $\sigma(t)$ there is an extensive literature on estimation using high-frequency market data (for example, Gatheral and Oomen (2010)). The primary focus there is to find effective means to filter out noise associated with market details such as bid and offer prices so as to obtain reliable estimates on time intervals that are as short as possible. Thus, for example, one could estimate $\sigma(t)$ by using market data from the preceding five minutes, which would typically contain hundreds of trades and potentially thousands of quote updates.
9. Techniques for Estimation of $\eta(t)$: Instantaneous liquidity, the inverse of $\eta(t)$ is more difficult to estimate, since it is an estimation of what *would* happen if one were to submit trades to the market rather than being an observable in itself. One proxy for the instantaneous trade history would be the realized trade volume over the last few minutes; if more people are trading actively in the market then one would be able to move a given number of shares with less slippage.
10. Trade Volume as Liquidity Proxy: A refined version of the above would be to measure the trade volume at or near the bid price if one is a buyer (or at the ask if one is a seller); large volume there would indicate the presence of a motivated seller and a good opportunity to go in as a buyer with low impact. Although these measures are not quantitatively very precise, they are often adequate to distinguish *good* opportunity from *bad*.
11. Persistence of the Market Properties: Both of these estimators rely on the presence of market properties (volatility an liquidity), so that information about the past provides reasonable forecasts for the future. Such persistence, at least across short horizons, is well documented (Bouchard, Farmer, and Lillo (2009) contain a review).
12. Time Dependent Liquidity/Volatility Patterns: Two broad classes of problems may be addressed. First is the case in which $\sigma(t)$ and $\eta(t)$ are both known non-random functions in time. This would accommodate the well-known intra-day profiles of volatility and liquidity; generally markets are more active in the mornings and in the close than in the middle of the day. This case is not the primary focus.
13. Stochastic Liquidity and Volatility Processes: The second case is when the volatility and the liquidity vary randomly through the day, so that $\sigma(t)$ and $\eta(t)$ follow some stochastic processes. This effect is very important in small and medium capitalization

stocks' algorithmic trading, and other assets that are less heavily traded than the large-cap US stocks.

Cost of Trading

1. Expression for the Transaction Cost: The *cost of trading* is the total cost paid to purchase X shares relative to the initial market value of XS_0

$$\mathbb{C} = \int_0^T \tilde{S}(t)v(t)dt - XS_0 = \int_0^T \sigma(t)x(t)dB(t) - \int_0^T \eta(t)v^2(t)dt$$

where

$$x(t) = \int_t^T v(t)dt$$

2. Dynamic Optimal Trading Cost Control: The cost $\mathbb{C}$ is a random variable, both because of the price uncertainty in $B(t)$ in the first and because of the liquidity uncertainty. The strategy $x(t)$ is to tailor the properties of this random variable to meet some optimal criterion.
3. Forward Time Cost of Trading: More generally, starting at time

$$t \geq 0$$

with $x(t)$ shares remaining to purchase, the cost of a strategy $x(s)$ on

$$t \leq s \leq T$$

is

$$\mathbb{C} = \int_t^T \sigma(s)x(s)dB(s) + \int_t^T \eta(s)v^2(s)ds$$

4. Trading Cost Expectation and Variance: The optimal trajectory is defined by the mean-variance criterion

$$\min_{x(s):\, t \leq s \leq T} \{\mathbb{E}[\mathbb{C}] + \lambda\mathbb{V}[\mathbb{C}]\}$$

where

$$\lambda \geq 0$$

is a risk-aversion coefficient. Note that

$$\mathbb{E}[\mathbb{C}] = \int_t^T \eta(s)v^2(s)ds$$

since the first term is an Ito's integral, and

$$\mathbb{V}[\mathbb{C}] = \int_t^T \sigma^2(s)x^2(s)ds + \{Terms\ from\ the\ Uncertainty\ of\ \eta(s)\ and\ \sigma(s)\}$$

5. Components of the Transaction Cost Variance: The first term in the variance contains the largest source of uncertainty, which corresponds to the price changes during execution. The other terms arise from the uncertainty in the market impact $\eta(s)$ that

will be paid on the transaction in the future, in the volatility $\sigma(s)$ that will be experienced at a later time, and in the trade strategy $v(s)$ itself if it is determined in response to uncertain market conditions.

6. Domination of the Market Volatility Term: Almgren (2009) argues that the first term in $\mathbb{V}[\mathbb{C}]$ above dominates the other terms.
7. Practical Use of Risk Aversion: The risk aversion coefficient λ is rarely defined in terms of fundamental investment preferences (Engle and Ferstenberg (2007)). Rather it is a parameter used to adjust the trajectories to a form that seems reasonable by other criteria such as representing a desired fraction of the market volume.

Constant Coefficients

1. Constant Volatility and Market Impact: The classic problem of Almgren and Chriss (2000) takes σ and η constant. Then for a strategy $x(t)$ that is fixed in advance and does not adapt to price motions

$$\mathbb{E}[\mathbb{C}] + \lambda\mathbb{V}[\mathbb{C}] = \int_t^T [\eta(s)v^2(s) + \sigma^2(s)x^2(s)]ds$$

2. Application of the Calculus of the Variations: Using the calculus of variations to minimize this over trajectories $x(s)$ gives the second order ODE

$$\frac{d^2x}{ds^2} = \kappa^2 x(s)$$

with

$$\kappa^2 = \frac{\lambda\sigma^2}{\eta}$$

3. Optimal Trading Rate and Trajectory: The solution is a combination of the exponentials $e^{\pm\kappa s}$

$$x(s) = x(t)\frac{\sinh[\kappa(T-s)]}{\sinh[\kappa(T-t)]}$$

$$v(s) = \kappa x(t)\frac{\cosh[\kappa(T-s)]}{\sinh[\kappa(T-t)]}$$

Thus $\frac{1}{\kappa}$ is the characteristic time scale of liquidation.

4. The Corresponding Cost of Trading: The strategy may also be expressed as a rule for

$$v(t) = \kappa x(t)\coth[\kappa(T-t)]$$

The corresponding cost function is

$$\mathbb{C}(x, t, \eta, \sigma) = \eta\kappa x^2 \coth[\kappa(T-t)] = \eta v x$$

5. Components of the Trading Cost: The total cost is equal to the impact cost component – neglecting the volatility term – incurred by trading x shares at a price concession given by the instantaneous velocity v. The actual trajectory slows down as the position size decreases, thus reducing market impact costs, but the total cost includes volatility risk as well as impact costs, giving the above value.
6. Non-dimensionalization of the Time Scales: The shape of the solution is governed by the non-dimensional quantity $\kappa(T-t)$ – the ratio of time remaining to the intrinsic time scale determined by the market's parameters and the trader's rick aversion.
7. Limit of Long Execution Time: In the infinite horizon limit

$$\kappa(T-t) \gg 1$$

the strategy has the limit

$$x(s) = x(t)e^{-\kappa(s-t)}$$

with

$$v(t) = \kappa x(t)$$

and the cost function

$$\mathbb{C} \to \eta\kappa x^2$$

Trading is substantially completed well before the expiration, and the precise value of T is not controlling.

8. <u>Risk Neutral Optimal Execution Trajectory</u>: Note that discounting has not been included, so the only motivation for rapid execution is risk aversion. In the limit of complete risk neutrality

$$\lambda \to 0$$

minimization of market impact costs would lead the trader to use all available time, and no infinite-horizon limit would exist; since

$$\kappa \to 0$$

the regime

$$\kappa T \gg 1$$

would never be achieved.

9. <u>Limit of Short Execution Time</u>: In the short horizon limit

$$\kappa(T-t) \ll 1$$

the strategy has the linear form

$$x(s) = x(t)\frac{T-s}{T-t}$$

$$v(s) = \frac{x(t)}{T-t}$$

and the cost function is essentially non-random with the value

$$\mathbb{C} \sim \frac{\eta x^2}{T-t}$$

$$\kappa(T-t) \to 0$$

If a time change is applied to match the market average profile then this is equivalent to the volume weighted average price (VWAP) execution.

10. Short-Term Limit – Higher Orders: In the same limit, the cost function has the higher order local behavior

$$\mathbb{C} \sim \frac{\eta x^2}{T-t} + \frac{\lambda \sigma^2 x^2}{3}(T-t) + \mathcal{O}((T-t)^3)$$

where ~ denotes asymptotic equivalence, that is, equal up to the terms that are asymptotically smaller than the displayed expressions in the given limit.

11. Short Term Limit Cost Components: The first term in this expression is the transaction cost associated with selling x shares at a price concession of

$$\eta v = \frac{\eta x}{T-t}$$

The second term is the risk penalty for holding an average of $\frac{x^2}{3}$ shares across time $T - t$.

12. Rolling Forward Dynamically Optimal Strategies: Whether adaptive strategies are better than a fixed one is a subtle question. Almgren and Chriss (2000) showed that if the strategy is re-evaluated at an intermediate time using the mean and variance measured at that time, then the optimal strategy is the remaining part of the initial strategy, and hence the optimal strategy is fixed. This is the context of Almgren (2009, 2012), since it is appropriate for dynamic programming.
13. Adaptive Strategies for *Large* Portfolios: In contrast, Almgren and Lorenz (2007), Lorenz (2008), and Lorenz and Almgren (2011) showed that adaptive strategies are optimal if the mean and the variance are measured at an initial time for portfolios that are large enough so that their impact is a substantial fraction of volatility. This framework is appropriate for *ex post* measurement of historical mean and variance across a large collection of trades. Tse, Forsyth, Kennedy, and Windcliff (2013) have given a fuller description of optimal solutions in the latter framework.
14. Cost Reductions from Adaptive Strategies: Schied and Schoneborn (2009) and Schied, Schoneborn, and Tehranchi (2010) showed that improvement from adaptivity depends on the risk profile; for example, it vanishes for a CARA utility function.

Coordinated Variation

1. Inverse Relation between Liquidity/Variance: Suppose $\sigma(t)$ and $\eta(t)$ vary inversely so that

$$\sigma^2(t)\eta(t) = constant = \bar{\sigma}^2\bar{\eta}$$

where $\bar{\sigma}$ and $\bar{\eta}$ are constant reference values.

2. Joint Arrival trading Time Model: For example, this relationship would be a natural consequence of a *trading time* model (Jones, Kaul, and Lipson (1994), Geman, Madan, and Yor (2001)) in which the single source of uncertainty is the arrival rate of trade events. If each trade event brings both a fixed amount of price variance, and the opportunity to trade a fixed number of shares for a particular cost, then one obtains the above relation.
3. Change of Drift/Wander Variables: Time may then be changed to an artificial variable defined by $\hat{t}(t)$ defined by

$$\Delta\hat{t} = \sigma^2(t)\Delta\hat{t}$$

In this time frame a modified Brownian motion $\hat{B}(\hat{t})$ results, with

$$\Delta\hat{B}(\hat{t}) = \sigma(t)\Delta B(t)$$

4. Change of Variables - The Trading Rate: The holdings are the same trajectory at different times, so

$$\hat{x}(\hat{t}) = x(t)$$

The trade rate is modified to

$$\hat{v}(\hat{t}) = -\frac{d\hat{x}}{d\hat{t}} = \frac{v(t)}{\sigma^2(t)}$$

5. Change of Variables: Transaction Cost: In terms of these new variables the trading cost is

$$\mathbb{C} = \int_0^{\hat{T}} \hat{x}(\hat{t})d\hat{B}(\hat{t}) + \bar{\sigma}^2\bar{\eta}\int_0^{\hat{T}} \hat{v}^2(\hat{t})d\hat{t}$$

where

$$\hat{T} = \hat{t}(T)$$

is the time horizon in the changed variable.

6. <u>Deterministic Liquidity and Volatility Profiles</u>: The above is easy to solve in two cases. First if the time varying volatility and liquidity have known non-random profiles, then the upper bound $\hat{T}$ may be computed exactly. The problem reduces exactly to the constant coefficient problem

$$\mathbb{E}[\mathbb{C}] + \lambda\mathbb{V}[\mathbb{C}] = \int_t^T [\eta(s)v^2(s) + \sigma^2(s)x^2(s)]ds$$

and the solution is the exponential functions computed there.

7. <u>The Corresponding Trade Rate/Time Scale</u>: The rule

$$v(t) = \kappa x(t) \coth[\kappa(T - t)]$$

becomes

$$v(t) = \kappa(t)x(t) \coth\left[\frac{\kappa(t)}{\sigma^2(t)} \int_t^T \sigma^2(s)ds\right]$$

where

$$\kappa(t) = \sqrt{\frac{\lambda\sigma^2(t)}{\eta(t)}}$$

is the time scale formed with the instantaneous values of the parameters.

8. Special Random Case - Infinite Horizon: If the coefficients vary randomly, then the problem is not the same as the constant-coefficient problem, because of the uncertainty in the end. But in the infinite horizon case

$$T = \infty$$

one also has

$$\hat{T} = \infty$$

under very mild assumptions on $\sigma(t)$. The trade rate is

$$v(t) = \kappa(t)x(t)$$

and the cost is

$$\mathbb{C} = \eta(t)\kappa(t)x^2 = x^2\sqrt{\lambda\bar{\sigma}^2\bar{\eta}}$$

9. Infinite Horizon Trade Rate/Cost: Somewhat surprisingly the optimal cost in the coordinated variation random market infinite horizon case does not depend on the instantaneous market state $\eta(t)$ and $\sigma(t)$ though the instantaneous trade rate does depend on the market state.
10. Trading Only Under Favorable Conditions: In effect, since volatility is low whenever the market impact is high, the strategy is always able to wait for favorable market conditions without incurring very much risk from the delay.
11. Absence of Variance/Liquidity Constraint: Thus the interesting problems come from two sources. First is the case of variation of the profiles of $\sigma(t)$ and $\eta(t)$ away from the "base case"

$$\sigma^2(t)\eta(t) = constant$$

with non-random coefficients. Kim and Boyd (2008) contain a fuller discussion of optimal trading with general market profiles.

12. Departure from "Trading Time" Approximation: For example, even within the *trading time* framework intra-day profiles may vary from the base case because different market participants are active at different times of the day. This leads to problems in the ODE's that are not the focus of this chapter.
13. Finite Horizon Adaptation to Randomness: With random coefficients, the proper handling of the uncertainty is needed as the end time is approached. For example, if liquidity is temporarily poor, is it worthwhile to wait for a better opportunity to trade, or is there a large risk that the opportunity will not come before expiration?

Rolling Time Horizon Approximate Strategy

1. Piecewise Constant Time Realizations: One way to determine a plausible strategy is to use

$$v(t) = \kappa x(t) \coth[\kappa(T - t)]$$

   to compute $v(t)$ using instantaneous values of $\eta(t)$, $\sigma(t)$, and hence $\kappa(t)$.
2. Explicitly Adapted On-Change Re-evaluation: That is, the assumption is that the values observed at each instant will remain constant throughout the liquidation period, and determine the optimal strategy using those values. When the values change, a stationary solution is re-computed.
3. Piecewise Adaptation Approach - Caveats: This strategy is strictly optimal only in the infinite horizon case, and only when the market parameters co-vary in the appropriate way. In general it is not optimal, but provides a reasonable solution that is easy to implement.

Small Impact Approximation

1. <u>Approximating the Trading Cost Variance</u>: In order to do dynamic programming when $\eta(t)$ and $\sigma(t)$ vary randomly, the variance term needs to be approximated. Almgren (2009) approximates it as

$$\mathbb{V}[\mathbb{C}] = \mathbb{E}\left[\int_t^T \sigma^2(s)x^2(s)ds\right]$$

   That is, the variance comes primarily by the price volatility represented by σ with lesser contributions from the uncertainty in $\eta(t)$ and $\sigma(t)$.
2. <u>The *Small Impact Approximation* Assumption</u>: When the market impact is small in absolute terms, it is nonetheless significant because it is always positive, but the uncertainty in the market impact is negligible compared to price volatility. This is called the *small-impact approximation*.
3. <u>"Market Power" Non-dimensional Quantity</u>: In the language of Almgren and Lorenz (2007) and Lorenz (2008) this is a small value of the "market power". This is also equivalent to the "small portfolio" approximation used in Lorenz and Almgren (2011) to neglect uncertainty in impact cost when the portfolio is small enough. Mathematically, the approximation relies on small values of "market power" parameter

$$\mu = \frac{\eta X / T}{\sigma\sqrt{T}}$$

   the amount by which trading moves the market compared to its intrinsic motion due to volatility across time T.
4. <u>The Corresponding Optimal Value Function</u>: The value function is then taken as

$$\mathbb{C}(t, x, \eta, \sigma) = \min_{x(s): t \le s \le T} \mathbb{E}\left[\int_t^T \lambda\sigma^2(s)x^2(s)ds + \int_t^T \eta(s)v^2(s)ds\right]$$

From this point on this approximation shall be made.

Dynamic Programming – Fully Co-ordinated Version

1. PDE Derivation and Numerical Solution: It is simplest to derive the PDE for the coordinated variation case directly, since that the only one that is solved here numerically. The extensions that are necessary to handle the general case are treated soon after.
2. Dynamics of Liquidity and Volatility: Since $\eta(t)$ and $\sigma(t)$ are positive it is convenient to write

$$\eta(t) = \bar{\eta} e^{\xi(t)}$$

and

$$\sigma(t) = \bar{\sigma} e^{-\frac{1}{2}\xi(t)}$$

where $\bar{\eta}$ and $\bar{\sigma}$ are typically values for η and σ, and $\xi(t)$ is a single non-dimensional variable indicating the *market state*. When $\xi(t)$ is large positive, the market is non-volatile and illiquid, and trading should be done more slowly; when $\xi(t)$ is large negative, the market is volatile and liquid, and trading should be done fast.
3. Dynamics of the Intrinsic Time Scale: The intrinsic time scale in the mean market is written as

$$\bar{\kappa} = \sqrt{\frac{\lambda \bar{\sigma}^2}{\bar{\eta}}}$$

and

$$\kappa(t) = \sqrt{\frac{\lambda\sigma^2(t)}{\eta(t)}} = \bar{\kappa}e^{-\xi(t)}$$

for the instantaneous value. The value function $c(t, x, \xi)$ then depends only on three variables.

4. <u>Coordinated Market State Evolution</u>: $\xi(t)$ is assumed to solve and SDE of the form

$$\Delta\xi(t) = a\big(\xi(t)\big)\Delta t + b\big(\xi(t)\big)\Delta B_L(t)$$

where $B_L(t)$ is a Brownian motion independent of the one driving the price motion.

5. <u>The Corresponding HJB Variational Increment</u>: Then by standard dynamic programming applied to

$$c(t, x, \eta, \sigma) = \min_{v(s), t \le s \le T} \mathbb{E}\left[\int_t^T \{\lambda\sigma^2(s)x^2(s) + \eta(s)v^2(s)\}\, ds\right]$$

with instantaneous trade rate

$$v = -\frac{dx}{dt}$$

as the control parameter, one writes

$$c(t, x, \xi) = \min_v \{\lambda\sigma^2x^2\Delta t + \eta v^2\Delta t + \mathbb{E}[c(t + \Delta t, x + \Delta x, \xi + \Delta\xi)]\}$$

giving the HJB PDE

$$0 = \frac{\partial c(t, x, \xi)}{\partial t} + \lambda\sigma^2x^2 + \min_v \left[\eta v^2 - v\frac{\partial c}{\partial x}\right] + a\frac{\partial c}{\partial \xi} + \frac{1}{2}b^2\frac{\partial^2 c}{\partial \xi^2}$$

6. Value Function Optimal HJB PDE: The minimum is clearly

$$v = \frac{1}{2\eta}\frac{\partial c}{\partial x}$$

and the PDE for $c(t, x, \xi)$ becomes

$$-\frac{\partial c(t, x, \xi)}{\partial t} = \lambda\sigma^2 x^2 - \frac{1}{4\eta}\left(\frac{\partial c}{\partial x}\right)^2 + a\frac{\partial c}{\partial \xi} + \frac{1}{2}b^2\frac{\partial^2 c}{\partial \xi^2}$$

7. Near Expiration Asymptotic Cost Behavior: The initial data for the HJB PDE above is in fact a local asymptotic condition and must be treated with some care. Near expiration, liquidation must happen on a linear trajectory

$$v = \frac{x}{T - t}$$

As with constant coefficients

$$\mathbb{C} \sim \frac{\eta x^2}{T - t} + \frac{\lambda\sigma^2 x^2}{3}(T - t) + \mathcal{O}((T - t)^3)$$

the cost comes primarily from market impact.

8. Approximating the Terminal Cost Estimate: To accurately approximate the cost one must account for the expected changes in the impact coefficient during the time

$$t \le s \le T$$

A simple application of the Ito's lemma shows that

$$\mathbb{E}[\eta(s)] = \eta(t)\left[1 + \left(a + \frac{1}{2}b^2\right)(s - t)\right]$$

$$s - t \to 0$$

and hence the average value of $\eta(s)$ for s between t and T is

$$\eta \sim \eta(t)\left[1 + \frac{1}{2}\left(a + \frac{1}{2}b^2\right)(T - t)\right]$$

$$T - t \to 0$$

9. <u>The Terminal Cost Estimate Approximation</u>: Using this value in

$$\mathbb{C} \sim \frac{\eta x^2}{T - t} + \frac{\lambda \sigma^2 x^2}{3}(T - t) + \mathcal{O}((T - t)^3)$$

$$\kappa(T - t) \to 0$$

with

$$\eta(t) = \bar{\eta} e^{\xi(t)}$$

one gets the cost expansion

$$\mathbb{C} \sim \frac{\bar{\eta} e^{\xi(t)} x^2}{T - t} + \frac{1}{2}\left(a + \frac{1}{2}b^2\right)\bar{\eta} e^{\xi(t)} x^2 + \mathcal{O}(T - t)$$

$$\kappa(T - t) \to 0$$

In the $\mathcal{O}(T - t)$ term there would appear both the risk contribution and a further expansion of the impact cost.

Log-Normal Model and Non-dimensionalization

1. Market State Explicit Drift/Wander: The next step is to assume that $\xi(t)$ evolves according to an Ornstein-Uhlenbeck mean-reverting process of zero mean. Thus

$$a(\xi(t)) = -\frac{\xi(t)}{\delta}$$

and

$$b(\xi(t)) = \frac{\beta}{\sqrt{\delta}}$$

2. Coordinated Market State Behavior: Here δ is a market relaxation time, and β is a *burstiness* parameter describing the dispersion of liquidity and volatility around their average levels. In the steady state $\xi(t)$ is normal with unconditional moments

$$\mathbb{E}[\xi(t)] = 0$$

and

$$\mathbb{V}[\xi(t)] = \frac{1}{2}\beta^2$$

3. Cost Function Dependence on x: Clearly the value function is strictly proportional to x^2 – the square of the number of shares remaining to execute. This is a consequence of the linear market impact model, since both the variance and the expected cost are quadratic in quantity.
4. Non-dimensionalization of the Cost Function: The non-dimensionalization may now be done using δ as the time scale and incorporating the factor x^2. One this defines

$$\tau = \frac{T - t}{\delta}$$

the time remaining to expiration as a multiple of the market relaxation time, and sets

$$c(t, x, \xi) = \frac{\bar{\eta} x^2}{\delta} u\left(\frac{T - t}{\delta}, \xi\right)$$

where $u(\tau, \xi)$ is a non-dimensional function of non-dimensional variables.

5. Non-dimensionalization of the HJB PDE: Then

$$-\frac{\partial c(t, x, \xi)}{\partial t} = \lambda \sigma^2 x^2 - \frac{1}{4\eta}\left(\frac{\partial c}{\partial x}\right)^2 + a\frac{\partial c}{\partial \xi} + \frac{1}{2} b^2 \frac{\partial^2 c}{\partial \xi^2}$$

becomes

$$\frac{\partial u(\tau, \xi)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi)}{\partial \xi} = e^{-\xi}[K^2 - u^2(\tau, \xi)] + \frac{1}{2}\beta^2 \frac{\partial^2 u(\tau, \xi)}{\partial \xi^2}$$

in which the non-dimensional risk aversion parameter K is given as

$$K = \bar{\kappa}\delta = \frac{Market\ Relaxation\ Time}{Trade\ Time\ Scale\ in\ Mean\ Market\ State}$$

6. The Corresponding Dimensional Trade Velocity: From

$$v = \frac{1}{2\eta}\frac{\partial c}{\partial x}$$

the dimensional trade velocity becomes

$$v = \frac{x}{\delta} e^{-\xi} u(\tau, \xi)$$

7. Terminal Asymptote Behavior of $u(\tau, \xi)$: Substituting the above expression into

$$\mathbb{C} \sim \frac{\bar{\eta} e^{\xi(t)} x^2}{T - t} + \frac{1}{2}\left(a + \frac{1}{2} b^2\right) \bar{\eta} e^{\xi(t)} x^2 + \mathcal{O}(T - t)$$

$$\kappa(T - t) \to 0$$

the initial condition is determined as

$$u(\tau, \xi) = \frac{e^{\xi}}{\tau} - \frac{1}{2}\left(\xi - \frac{1}{2} \beta^2\right) e^{\xi} + \mathcal{O}(\tau)$$

as

$$\tau \to 0$$

for each fixed ξ.

8. Terminal $u(\tau, \xi)$ Asymptote at $\xi < 0$: For

$$\xi < 0$$

when trading is fast, the region of approximate validity of this trading is limited by the rate of trading itself, and this expression should be replaced by

$$u(\tau, \xi) \sim K \coth\left(K \tau e^{-\xi}\right)$$

$$\xi \to -\infty$$

for

$$\tau > \mathcal{O}\left(e^{-\xi}\right)$$

9. <u>Terminal $u(\tau, \xi)$ Asymptote at $\xi > 0$</u>: For

$$\xi > 0$$

when trading is slow, the region of validity is

$$\tau \ll \mathcal{O}(1)$$

since the market itself changes on times of scale $\mathcal{O}(1)$. Almgren (2012) illustrates using pictorial summary the various asymptotic behavior of solutions to

$$\frac{\partial u(\tau, \xi)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi)}{\partial \xi} = e^{-\xi}[K^2 - u^2(\tau, \xi)] + \frac{1}{2}\beta^2 \frac{\partial^2 u(\tau, \xi)}{\partial \xi^2}$$

Constant Market

1. <u>Non-volatile Steady-State Market</u>: The steady-state market takes

$$\beta = 0$$

Along the line

$$\xi = 0$$

the PDE

$$\frac{\partial u(\tau,\xi)}{\partial\tau}+\xi\frac{\partial u(\tau,\xi)}{\partial\xi}=e^{-\xi}[K^2-u^2(\tau,\xi)]+\frac{1}{2}\beta^2\frac{\partial^2 u(\tau,\xi)}{\partial\xi^2}$$

reduces to the ODE

$$\frac{\partial u(\tau,\xi)}{\partial\tau}=K^2-u^2(\tau,\xi)$$

with

$$u(\tau)\sim\frac{1}{\tau}+\mathcal{O}(\tau)$$

as

$$\tau\to 0$$

whose solution is

$$u(\tau)\sim K\coth(K\tau)$$

On undoing the change of variables, this reduces exactly to

$$c(t,x,\eta,\sigma)=\eta\kappa x^2\coth[\kappa(T-t)]=\eta v(t)x$$

2. <u>The $\xi\to-\infty$ Case: Fast Trading</u>: To generalize the above relation one considers the limit

$$\xi\to-\infty$$

That is, the market impact is very small, and the volatility is very large, thus the optimal strategy is trade very quickly.

3. Fast Relative to Market Relaxation: Since the market relaxation time scales are fixed, fast trading means that the program is completed before the market parameters have had time to change. Thus the cost is the static cost

$$c(t, x, \eta, \sigma) = \eta\kappa x^2 \coth[\kappa(T - t)] = \eta v(t)x$$

using the instantaneous market parameters, which, in the transformed functions becomes

$$u(\tau, \xi) \sim K \coth\left(K\tau e^{-\xi}\right)$$

$$\xi \to -\infty$$

4. Relaxation to *Rolling Horizon* Strategy: The corresponding trade rate is the *rolling horizon* strategy, which is always an admissible, though sub-optimal, strategy. The expression for $u(\tau, \xi)$ accurately describes the optimal cost only in the indicated limit when indeed the market coefficients do not change substantially before trading is completed.

Long Time

1. Asymptotic $u(\tau, \xi)$ far from Expiration: As noted before, with coordinated variation, when time is far from expiry, the value of the function is

$$C = x^2\sqrt{\lambda\bar{\sigma}^2\bar{\eta}}$$

or

$$u(\tau, \xi) \to \infty$$

as

$$\tau \to \infty$$

in non-dimensional terms. Certainly

$$u = K$$

is a steady state solution of the PDE.

2. Validity of the Terminal Asymptote: And since the value function must be decreasing in τ it is clear that

$$u(\tau, \xi) \geq K$$

for all

$$\tau \geq 0$$

As a consequence the initial expression

$$u(\tau, \xi) = \frac{e^{\xi}}{\tau} - \frac{1}{2}\left(\xi - \frac{1}{2}\beta^2\right)e^{\xi} + \mathcal{O}(\tau)$$

as

$$\tau \to 0$$

for each fixed ξ can only be valid when

$$\frac{e^{\xi}}{\tau} \geq K$$

or

$$\tau \leq \mathcal{O}\left(e^{\xi}\right)$$

- a very thin region when ξ is negative.

3. Uniqueness of the Solution to the PDE: Provided that a unique solution $u(\tau, \xi)$ to

$$\frac{\partial u(\tau, \xi)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi)}{\partial \xi} = e^{-\xi}[K^2 - u^2(\tau, \xi)] + \frac{1}{2}\beta^2 \frac{\partial^2 u(\tau, \xi)}{\partial \xi^2}$$

exists, a standard verification argument establishes that this function does indeed give the optimal control to the original control problem.

4. Uniqueness of the Solution to the PDE: Since

$$\frac{\partial u(\tau, \xi)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi)}{\partial \xi} = e^{-\xi}[K^2 - u^2(\tau, \xi)] + \frac{1}{2}\beta^2 \frac{\partial^2 u(\tau, \xi)}{\partial \xi^2}$$

is a non-degenerate diffusion equation with lower-order terms, it certainly has smooth unique solutions locally in time if the solution at some positive time satisfies

$$u(\tau, \xi) < C e^{\alpha \xi^2}$$

5. Decomposition of the Initial Term: To understand the initial behavior near the initial term, writing

$$u(\tau, \xi) = \frac{e^{\xi}}{\tau}[1 + w(\tau, \xi)]$$

$w(\tau, \xi)$ for

$$\tau > 0$$

satisfies the PDE

$$\frac{\partial w(\tau,\xi)}{\partial \tau} + \frac{w(\tau,\xi)}{\tau}[1 + w(\tau,\xi)]$$
$$= K^2 \tau e^{-2\xi} - \left(\xi - \frac{1}{2}\beta^2\right) - \left(\xi - \frac{1}{2}\beta^2\right) w(\tau,\xi) - (\xi - \beta^2)\frac{\partial w(\tau,\xi)}{\partial \tau}$$
$$+ \frac{1}{2}\beta^2 \frac{\partial^2 w(\tau,\xi)}{\partial \xi^2}$$

as

$$\underset{\tau \to 0}{Limit}\, w(\tau,\xi) = 0$$

for each ξ.

6. Uniqueness of the Decomposed Solution: It is in this sense that $u(\tau,\xi)$ satisfies its PDE

$$\frac{\partial u(\tau,\xi)}{\partial \tau} + \xi \frac{\partial u(\tau,\xi)}{\partial \xi} = e^{-\xi}[K^2 - u^2(\tau,\xi)] + \frac{1}{2}\beta^2 \frac{\partial^2 u(\tau,\xi)}{\partial \xi^2}$$

and the singular boundary condition. Although Almgren (2012) does not formally present a proof for the existence and the uniqueness of the function $w(\tau,\xi)$ and hence of $u(\tau,\xi)$ there do not appear to be any obstacles.

7. Perturbation of the Decomposed Solution: A search for the perturbation expansion of the form

$$w(\tau,\xi) \sim \tau w_1(\xi) + \tau^2 w_2(\xi) + \cdots$$

$$\tau \to 0$$

readily determines

$$w_1(\xi) = -\frac{1}{2}\left(\xi - \frac{1}{2}\beta^2\right)$$

$$w_2(\xi) = \frac{1}{3}K^2e^{-2\xi} + \frac{1}{12}\left[\xi^2 + (2-\beta^2)\xi + \frac{1}{4}\beta^4 - 2\beta^2\right]$$

8. <u>Local Behavior Description for $u(\tau,\xi)$</u>: The construction of this asymptotic behavior is strong evidence that the solution exists and has the associated local behavior. Thus a description of the local behavior of $u(\tau,\xi)$ slightly fuller than

$$u(\tau,\xi) = \frac{e^{\xi}}{\tau} - \frac{1}{2}\left(\xi - \frac{1}{2}\beta^2\right)e^{\xi} + \mathcal{O}(\tau)$$

as

$$\tau \to 0$$

for each fixed ξ is

$$u(\tau,\xi) \sim \frac{e^{\xi}}{\tau} - \frac{1}{2}\left(\xi - \frac{1}{2}\beta^2\right)e^{\xi} + \tau\left\{\frac{1}{3}K^2e^{-2\xi} + \frac{1}{12}\left[\xi^2 + (2-\beta^2)\xi + \frac{1}{4}\beta^4 - 2\beta^2\right]\right\} + \mathcal{O}(\tau^2)$$

Dynamic Programming – Custom $\eta(t)$ and $\sigma(t)$

1. <u>Non-dimensional Liquidity and Volatility</u>: Since $\eta(t)$ and $\sigma(t)$ are positive it is convenient to use

$$\xi(t) = \log \frac{\eta(t)}{\bar{\eta}}$$

and

$$\zeta(t) = \log \frac{\sigma(t)}{\bar{\sigma}}$$

as state variables.

2. Mean Market State Time Scale: Here $\bar{\eta}$ and $\bar{\sigma}$ are typical values of $\eta(t)$ and $\sigma(t)$, and $\xi(t)$ and $\zeta(t)$ are non-dimensional values that fluctuate around zero. $\bar{\kappa}$ is written as

$$\bar{\kappa} = \sqrt{\frac{\lambda\bar{\sigma}^2}{\bar{\eta}}}$$

for the intrinsic time scale in the mean market state.

3. Evolution of $\xi(t)$ and $\zeta(t)$: $\xi(t)$ and $\zeta(t)$ are taken to evolve according to the stochastic differential equations (SDE) of the forms

$$\Delta\xi = a_\xi \Delta t + b_\xi \Delta\beta_L$$

and

$$\Delta\zeta = a_\zeta \Delta t + b_\zeta \Delta\beta_V$$

where a_ξ, b_ξ, a_ζ, and b_ζ are coefficients whose values may depend on $\xi(t)$ and $\zeta(t)$.

4. Correlated Liquidity/Volatility Brownian Processes: $\beta_L(t)$ and $\beta_V(t)$ are Brownian motions that are independent of the process $B(t)$ driving the price process, but possibly correlated with each other, with

$$\mathbb{E}[\Delta\beta_L\Delta\beta_V] = \rho\Delta t$$

5. Applying the Dynamic Programming Criterion: Then, by dynamic programming, it follows that

$$\mathbb{C}(t, x, \xi, \zeta) = \underset{v}{min}\{\lambda\sigma^2 x^2 \Delta t + \eta v^2 \Delta t + \mathbb{C}(t + \Delta t, x + \Delta x, \xi + \Delta\xi, \zeta + \Delta\zeta)\}$$

giving the PDE

$$\begin{aligned} 0 = \frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial t} + \lambda\sigma^2 x^2 + \underset{v}{min}\left[\eta v^2 - v\frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial x}\right] + a_\xi \frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial\xi} \\ + a_\zeta \frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial\zeta} + \frac{1}{2}{b_\xi}^2 \frac{\partial^2\mathbb{C}(t, x, \xi, \zeta)}{\partial\xi^2} + \frac{1}{2}{b_\zeta}^2 \frac{\partial^2\mathbb{C}(t, x, \xi, \zeta)}{\partial\zeta^2} \\ + \frac{1}{2} b_\xi b_\zeta \frac{\partial^2\mathbb{C}(t, x, \xi, \zeta)}{\partial\xi\partial\zeta} \end{aligned}$$

6. Optimality in the Trade Rate Space: The minimum is clearly

$$v = \frac{1}{2\eta}\frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial x}$$

and the PDE for $\mathbb{C}(t, x, \xi, \zeta)$ is

$$\begin{aligned} -\frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial t} \\ = \lambda\sigma^2 x^2 - \frac{1}{4\eta}\left[\frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial x}\right]^2 + a_\xi \frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial\xi} + a_\zeta \frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial\zeta} \\ + \frac{1}{2}{b_\xi}^2 \frac{\partial^2\mathbb{C}(t, x, \xi, \zeta)}{\partial\xi^2} + \frac{1}{2}{b_\zeta}^2 \frac{\partial^2\mathbb{C}(t, x, \xi, \zeta)}{\partial\zeta^2} + \rho b_\xi b_\zeta \frac{\partial^2\mathbb{C}(t, x, \xi, \zeta)}{\partial\xi\partial\zeta} \end{aligned}$$

7. Exogenous Expiration Trade Rate Asymptote: Near expiration the liquidation must happen on a linear trajectory

$$v = \frac{x}{T - t}$$

The cost comes entirely from market impact in the market conditions at that time, since volatility risk is negligible across a short time. Thus

$$\mathbb{C}(t, x, \xi, \zeta) \sim \frac{\eta x^2}{T - t}$$

$$\kappa(T - t) \to 0$$

applies, and

$$\mathbb{C}(t, x, \xi, \zeta) \sim \frac{\eta x^2}{T - t} = \frac{\bar{\eta} e^{\xi} x^2}{T - t}$$

$$(T - t) \to 0$$

8. De-dimensionalization of the PDE State Variables: To non-dimensionalize the cost function and the differential equation, a time scale needs to be defined, which also helps define a cost scale; the market parameters are already non-dimensionalized by their mean values.
9. Time Scale Choice - Liquidity Reversion: So far the only two time scales are the intrinsic liquidation time in the mean market state $\frac{1}{\kappa}$ and the imposed horizon T. Since both of these depend on a trader's preferences for a particular trade order, it will be more natural to use a time scale based on market dynamics.

Log-Normal Model

1. Ornstein-Uhlenbeck Mean-Reverting Dynamics: Here the assumption is that $\xi(t)$ and $\zeta(t)$ evolve according to the mean-reverting process of zero mean.
2. $\xi(t)$ and $\zeta(t)$ Drift/Wander: Accordingly

$$a_\xi = -\frac{\xi}{\delta_L}$$

$$a_\zeta = -\frac{\zeta}{\delta_V}$$

$$b_\xi = \frac{\beta_L}{\sqrt{\delta_L}}$$

$$b_\zeta = \frac{\beta_V}{\sqrt{\delta_V}}$$

3. Liquidity/Volatility Relaxation Time Scales: Here δ_L and δ_V are relaxation time scales for liquidity and volatility, and β_L and β_V are non-dimensional "burstiness" parameters.
4. $\xi(t)$ and $\zeta(t)$ Steady State: In the steady state, $\xi(t)$ and $\zeta(t)$ are normal with

$$\mathbb{E}[\xi(t)] \to 0$$

$$\mathbb{V}[\xi(t)] \to \frac{1}{2}{\beta_L}^2$$

$$\mathbb{E}[\zeta(t)] \to 0$$

$$\mathbb{V}[\zeta(t)] \to \frac{1}{2}{\beta_V}^2$$

as

$$t \to \infty$$

Thus β_L and β_V describe the liquidity and the volatility around their mean levels.

5. δ_L as the Reference Time Scale: One may non-dimensionalize using δ_L as the time scale. That is, defining

$$\tau = \frac{T - t}{\delta_L}$$

and setting

$$\mathbb{C}(t, x, \xi, \zeta) = \frac{\bar{\eta} x^2}{T - t} = u\left(\frac{T - t}{\delta_L}, \xi, \zeta\right)$$

where $u\left(\frac{T-t}{\delta_L}, \xi, \zeta\right)$ is a non-dimensional function of non-dimensional variables, one gets

$$\begin{aligned}\frac{\partial u(\tau, \xi, \zeta)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi, \zeta)}{\partial \xi} + \mu\zeta \frac{\partial u(\tau, \xi, \zeta)}{\partial \zeta} \\ &= K^2 e^{2\zeta} - e^{-\xi} u^2(\tau, \xi, \zeta) + \frac{1}{2} {\beta_L}^2 \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \xi^2} + \frac{1}{2} {\beta_V}^2 \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \zeta^2} \\ &+ \rho\sqrt{\mu}\beta_L\beta_V \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \xi \partial \zeta}\end{aligned}$$

6. Volatility Market State Time Scales: On writing

$$\mu = \frac{\delta_L}{\delta_V}$$

the multi-dimensional risk aversion parameter becomes

$$K = \bar{\kappa}\delta_L = \frac{Relaxation\,Time\,for\,Market\,Liquidity}{Trade\,Time\,Scale\,in\,Mean\,Market\,State}$$

Of course τ is the time remaining to expiration measured in units of the market relaxation time.

7. Initial Condition Re-cast: The initial condition becomes

$$u(\tau, \xi, \zeta) = \frac{e^{\xi}}{\tau}$$

$$\tau \to 0$$

8. The Corresponding Dimensional Trade Rate: From

$$v = \frac{1}{2\eta}\frac{\partial\mathbb{C}(t, x, \xi, \zeta)}{\partial x}$$

the dimensional trade velocity is

$$v = \frac{xe^{-\xi}}{\delta_L}u(\tau, \xi, \zeta)$$

9. Deterministic Liquidity and Volatility Processes: The constant volatility market takes

$$\beta_L = \beta_V = 0$$

Along the line

$$\xi = \zeta = 0$$

the PDE

$$\frac{\partial u(\tau,\xi,\zeta)}{\partial\tau}+\xi\frac{\partial u(\tau,\xi,\zeta)}{\partial\xi}+\mu\zeta\frac{\partial u(\tau,\xi,\zeta)}{\partial\zeta}$$
$$=K^2e^{2\zeta}-e^{-\xi}u^2(\tau,\xi,\zeta)+\frac{1}{2}{\beta_L}^2\frac{\partial^2u(\tau,\xi,\zeta)}{\partial\xi^2}+\frac{1}{2}{\beta_V}^2\frac{\partial^2u(\tau,\xi,\zeta)}{\partial\zeta^2}$$
$$+\rho\sqrt{\mu}\beta_L\beta_V\frac{\partial^2u(\tau,\xi,\zeta)}{\partial\xi\partial\zeta}$$

reduces to the ODE

$$\frac{\partial u(\tau,\xi,\zeta)}{\partial\tau}=K^2-u^2(\tau,\xi,\zeta)$$

with

$$u(\tau)\to\frac{1}{\tau}$$

as

$$\tau\to 0$$

whose solution is

$$u(\tau)=K\coth K\tau$$

10. Constant Coefficient Cost Function Reduction: On undoing the change of variables, this reduces exactly to

$$\mathbb{C}(t,x,\eta,\sigma)=\eta\kappa x^2\coth\kappa(T-t)=\eta v x$$

11. Low Market Impact/High Volatility Scenario: To generalize the above solution the limits

$$\xi \to -\infty$$

and

$$\zeta \to +\infty$$

are considered. That is, the market impact is temporarily very small and the volatility is very large; the optimal strategy would be to trade very quickly.

12. Consequence of the Fast Trading: Since the market relaxation times are fixed, fast trading means that the program is completed before the market parameters have had time to change.
13. The Corresponding Non-dimensional Cost: Thus the cost is the static cost

$$\mathbb{C}(t, x, \eta, \sigma) = \eta \kappa x^2 \coth \kappa (T - t) = \eta v x$$

using instantaneous market parameters, which in the transformed functions becomes

$$u(\tau, \xi, \zeta) \sim K e^{\zeta + \frac{1}{2}\xi} \coth K \tau e^{\zeta - \frac{1}{2}\xi}$$

$$\xi \to -\infty$$

$$\zeta \to +\infty$$

14. Comparison with the "Rolling Horizon" Approximation: The corresponding trade rate is the "rolling horizon" strategy seen earlier, which is always an admissible, although sub-optimal, strategy. The expression

$$u(\tau, \xi, \zeta) \sim K e^{\zeta + \frac{1}{2}\xi} \coth K \tau e^{\zeta - \frac{1}{2}\xi}$$

$$\xi \to -\infty$$

$$\zeta \to +\infty$$

accurately describes the optimal cost only in the indicated limit, when indeed the market coefficients do not change substantially before the trading is completed.

Coordinated Variation

1. Coordinated Variation Case Reduction: Rather than solve the full PDE

$$\frac{\partial u(\tau,\xi,\zeta)}{\partial \tau} + \xi \frac{\partial u(\tau,\xi,\zeta)}{\partial \xi} + \mu\zeta \frac{\partial u(\tau,\xi,\zeta)}{\partial \zeta} = K^2 e^{2\zeta} - e^{-\xi} u^2(\tau,\xi,\zeta) + \frac{1}{2}{\beta_L}^2 \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \xi^2} + \frac{1}{2}{\beta_V}^2 \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \zeta^2} + \rho\sqrt{\mu}\beta_L\beta_V \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \xi \partial \zeta}$$

in two space dimensions and one time dimension, more insight can be attained by considering the coordinated variation model described above.

2. Equal Liquidity/Volatility Time Scales: Thus the following assumptions are made on the stochastic processes. First the time scales of liquidity and volatility are assumed to be equal.

$$\delta_L = \delta_V = \delta$$

so

$$\mu = 1$$

3. Fully Correlated Liquidity/Volatility Brownians: The Brownian motions driving the liquidity and the volatility have perfect positive correlation

$$\rho = 1$$

4. Liquidity/Volatility Wander Intensity Ratio: For now the fluctuation magnitudes β_L and β_V are arbitrary. The setting is

$$\mathcal{y} = -\frac{\beta_L}{\beta_V} = \frac{Signed\ Fractional\ Variation\ of\ \sigma^2}{Signed\ Fractional\ Variation\ of\ \eta}$$

so that the coordinated variation case takes

$$\mathcal{y} = 1$$

The assumption here is that

$$\mathcal{y} > 0$$

5. Wander Intensity Scaled Volatility/Liquidity: Then, from

$$\Delta\xi = a_\xi \Delta t + b_\xi \Delta\beta_L$$

and

$$\Delta\zeta = a_\zeta \Delta t + b_\zeta \Delta\beta_V$$

$$a_\xi = -\frac{\xi}{\delta_L}$$

$$a_{\zeta} = -\frac{\zeta}{\delta_V}$$

$$b_{\xi} = \frac{\beta_L}{\sqrt{\delta_L}}$$

$$b_{\zeta} = \frac{\beta_V}{\sqrt{\delta_V}}$$

one gets

$$\Delta(\beta_V \xi - \beta_L \zeta) = -(\beta_V \xi - \beta_L \zeta) \frac{\Delta t}{\delta}$$

and hence, after at most an initial transient, the solutions satisfy

$$(\beta_V \xi - \beta_L \zeta) = 0$$

or

$$\zeta = -\frac{\mathscr{y}}{2} \xi$$

6. <u>The Coordinated Variation Cost PDE</u>: The PDE

$$\frac{\partial u(\tau, \xi, \zeta)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi, \zeta)}{\partial \xi} + \mu \zeta \frac{\partial u(\tau, \xi, \zeta)}{\partial \zeta}$$
$$= K^2 e^{2\zeta} - e^{-\xi} u^2(\tau, \xi, \zeta) + \frac{1}{2} {\beta_L}^2 \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \xi^2} + \frac{1}{2} {\beta_V}^2 \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \zeta^2}$$
$$+ \rho \sqrt{\mu} \beta_L \beta_V \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \xi \partial \zeta}$$

maybe re-cast as

$$\frac{\partial u(\tau,\xi,\zeta)}{\partial\tau} + \left(\xi\frac{\partial}{\partial\xi} + \zeta\frac{\partial}{\partial\zeta}\right)u(\tau,\xi,\zeta)$$

$$= K^2 e^{2\zeta} - e^{-\xi}u^2(\tau,\xi,\zeta) + \frac{1}{2}\left(\beta_L\frac{\partial}{\partial\xi} + \beta_V\frac{\partial}{\partial\zeta}\right)^2 u(\tau,\xi,\zeta)$$

Ignoring cross-variation (see the reasons below) the PDE for $u(\tau,\xi,\zeta)$ is, with

$$\beta = \beta_L$$

$$\frac{\partial u(\tau,\xi,\zeta)}{\partial\tau} + \xi\frac{\partial u(\tau,\xi,\zeta)}{\partial\xi} = K^2 e^{-\mathscr{y}\xi} - e^{-\xi}u^2(\tau,\xi,\zeta) + \frac{1}{2}\beta^2\frac{\partial^2 u(\tau,\xi,\zeta)}{\partial\xi^2}$$

7. <u>Change of Variable - Initial Condition</u>: The initial condition is still

$$u(\tau,\xi) = \frac{e^{\xi}}{\tau}$$

as

$$\tau \to 0$$

8. <u>Change of Variable - Wander Differential</u>: To trace the change of variables in detail, introduce $w(\tau,\xi,\zeta)$ with

$$u(\tau,\xi,\zeta) = w(\tau,\xi,\beta_V\xi - \beta_L\zeta)$$

so that

$$\left(\xi\frac{\partial}{\partial\xi} + \zeta\frac{\partial}{\partial\zeta}\right)u(\tau,\xi,\zeta) = \left(\xi\frac{\partial}{\partial\xi} + \chi\frac{\partial}{\partial\chi}\right)w(\tau,\xi,\chi)$$

and

$$\left(\beta_L \frac{\partial}{\partial \xi} + \beta_V \frac{\partial}{\partial \zeta}\right)^2 u(\tau, \xi, \zeta) = {\beta_L}^2 \frac{\partial^2 w(\tau, \xi, \chi)}{\partial \xi^2}$$

9. <u>Change of Variables Zero Wander</u>: Thus

$$\frac{\partial w(\tau, \xi, \chi)}{\partial \tau} + \left(\xi \frac{\partial}{\partial \xi} + \chi \frac{\partial}{\partial \chi}\right) w(\tau, \xi, \chi) = K^2 e^{2\zeta} - e^{-\xi} w^2(\tau, \xi, \chi) + \frac{1}{2} {\beta_L}^2 w(\tau, \xi, \chi)$$

and on the plane

$$\chi = 0$$

this reduces to

$$\frac{\partial u(\tau, \xi, \zeta)}{\partial \tau} + \xi \frac{\partial u(\tau, \xi, \zeta)}{\partial \xi} = K^2 e^{-y\xi} - e^{-\xi} u^2(\tau, \xi, \zeta) + \frac{1}{2} \beta^2 \frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \xi^2}$$

-

Asymptotic Behavior

1. <u>Short Time Frame Asymptotic Behavior</u>: To study the behavior as

$$\tau \to 0$$

for a fixed ξ one writes

$$u(\tau, \xi) = \frac{e^{\xi}}{\tau} + u_0(\xi) + \tau u_1(\xi) + \cdots$$

$$\tau \to 0$$

2. <u>Short Time Higher Order Dependence</u>: One finds that at $\mathcal{O}\left(\frac{1}{\tau}\right)$

$$u_0(\xi) = -\frac{1}{2}\left(\xi - \frac{1}{2}\beta^2\right)e^{\xi}$$

and at $\mathcal{O}(1)$

$$3u_1(\xi) = K^2 e^{-\mathscr{y}\xi} - e^{-\xi}{u_0}^2(\xi) - \xi {u_0}'(\xi) + \frac{1}{2}\beta^2 {u_0}''(\xi)$$

or

$$u_1(\xi) = \frac{1}{3}K^2 e^{-\mathscr{y}\xi} + \frac{1}{12}e^{\xi}\left[\left(\xi + 1 - \frac{1}{2}\beta^2\right)^2 - (1 + \beta^2)\right]$$

3. <u>Implications for the Solution Robustness</u>: Thus $u(\tau, \xi)$ has a regular expansion on the powers of τ. This reassures that the singular initial data is indeed enough to define the solution.
4. <u>Long Time Asymptotic PDE Dependence</u>: As

$$\tau \to \infty$$

presumably there is a steady state cost and a strategy in which the horizon is not controlling. The steady state solution $u(\xi)$ will be determined by the second order nonlinear ODE

$$\xi u' = K^2 e^{-\mathscr{y}\xi} u^2 + \frac{1}{2}\beta^2 u''$$

5. <u>Explicit Long Time Asymptote PDE</u>: In the coordinated variation case

$$\mathscr{y} = 1$$

this has a constant solution

$$u(\xi) = K$$

For

$$\mathscr{y} \neq 1$$

an explicit solution cannot be given, but for

$$\mathscr{y} > -1$$

the asymptotic behavior

$$u(\xi) \sim K e^{-\frac{1}{2}(\mathscr{y}-1)\xi}$$

as

$$\xi \to -\infty$$

can be identified, based on the balance

$$0 = K^2 e^{-\mathscr{y}\xi} - e^{-\xi} u^2$$

6. Low Impact High Volatility Simplification: As

$$\xi \to -\infty$$

with

$$\mathscr{y} > 0$$

one also has

$$\zeta \to +\infty$$

and thus the asymptotic solution

$$u(\tau, \xi, \zeta) \sim K e^{\zeta + \frac{1}{2}\xi} \coth\left(K\tau e^{\zeta - \frac{1}{2}\xi}\right)$$

$$\xi \to -\infty$$

$$\zeta \to +\infty$$

is valid, and simplifies to

$$u(\tau, \xi) \sim K e^{-\frac{1}{2}(\mathscr{y}-1)\xi} \coth\left(K\tau e^{-\frac{1}{2}(\mathscr{y}+1)\xi}\right)$$

$$\xi \to -\infty$$

7. Consistency with Long/Short Time Asymptote: At leading order this is consistent with both the short-term and the long-term behavior above.

Numerical Solution

1. Numerical Solution to the HJB: Since explicit analytical solutions cannot be given, Almgren (2009) resorts to numerical solutions to generate solution solutions to

$$\frac{\partial u(\tau,\xi,\zeta)}{\partial \tau} + \xi \frac{\partial u(\tau,\xi,\zeta)}{\partial \xi} = K^2 e^{-\mathscr{y}\xi} - e^{-\xi} u^2(\tau,\xi,\zeta) + \frac{1}{2}\beta^2 \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \xi^2}$$

for a range of the given parameters.

2. Choice for the Fluctuation Ratio: The coordination parameter $\mathscr{y}$ should be chosen as part of the market structure. Since there is no particular reason to choose other values

$$\mathscr{y} = 1$$

shall be considered.

3. Solution to the Generalized HJB: Similarly Almgren (2009) does not illustrate numerical solutions to the two-variable problem

$$\begin{aligned}\frac{\partial u(\tau,\xi,\zeta)}{\partial \tau} + \xi \frac{\partial u(\tau,\xi,\zeta)}{\partial \xi} + \mu\zeta \frac{\partial u(\tau,\xi,\zeta)}{\partial \zeta} \\ = K^2 e^{2\zeta} - e^{-\xi} u^2(\tau,\xi,\zeta) + \frac{1}{2}{\beta_L}^2 \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \xi^2} + \frac{1}{2}{\beta_V}^2 \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \zeta^2} \\ + \rho\sqrt{\mu}\beta_L\beta_V \frac{\partial^2 u(\tau,\xi,\zeta)}{\partial \xi \partial \zeta}\end{aligned}$$

since a truly complete study would also consider a broader range of market dynamics models.

4. Choice of the Burstiness Parameters: The burstiness parameter β is stock-specific. A large cap stock will have a β near zero, for a near-uniform profile. A small cap stock will have

$$\beta = 1$$

or larger. For the sample calculations below β is fixed at 2 – a relatively large value to better illustrate the effects of market variation.

5. Choice of the Risk Aversion: The risk aversion parameter K must range across non-negative values, since the actual choice of the trajectory will be determined by the

trader's risk preference. Values of K smaller than 1 are the most realistic, so that the algorithm has time to adapt to at least one market reversion time.

6. Technical Issues behind the Solution: Almgren (2009) briefly discusses a few issues with space and time discretization, and presents example solutions, which is covered below.

Time Discretization

1. Rationale for the Modified Euler Scheme: The first obstacle is that the initial condition is given as singular behavior. A simple modification to the Euler's forward scheme handles this problem. This is illustrated below using an ordinary differential equation (ODE).
2. The ODE and its Solution: Consider

$$\frac{du(t)}{dt} = -(u - a)(u - b)$$

$$u(t) \sim \frac{1}{t}$$

as

$$t \to 0$$

whose exact solution is

$$u(t) = \frac{ae^{-bt} - be^{-at}}{e^{-bt} - e^{-at}}$$

3. Local Expansion at the Origin: Either by expanding the solution, or directly from the ODE, the local expansion is determined as

$$u(t) \sim \frac{1}{t} + \frac{1}{2}(a+b) + \frac{t}{12}(a-b)^2 + \cdots$$

as

$$t \to 0$$

4. <u>Euler Scheme on Modified ODE</u>: For the numerics a forward Euler scheme is applied to

$$w(t) = tu(t)$$

which is regular near

$$t = 0$$

With

$$w'(t) = tu'(t) + u(t)$$

and denoting

$$u_n \approx u(t_n)$$

this gives

$$u_{n+1} = u_n + \frac{t_n}{t_{n+1}}(t_{n+1} - t_n){u_n}'$$

5. <u>Correction to the Euler Update</u>: Thus, a correction is applied to he Euler update formula, which becomes small as one moves away from the initial singular time and

$$\frac{t_n}{t_{n+1}} \to 1$$

6. Evolution on a Time Grid: The test is done on a regular grid with

$$t_n = n\Delta t$$

starting at

$$n = k \geq 1$$

k is chosen to satisfy the stability criterion for the forward Euler scheme.

7. The First Time Node: For an ODE

$$\frac{du}{dt} = f(u)$$

the stability criterion requires that

$$\Delta t < \frac{1}{\left|\frac{df(u)}{du}\right|}$$

In this case

$$\frac{df(u)}{du} \sim 2u \sim \frac{2}{t}$$

so the stability criterion translates to

$$\Delta t < \frac{t}{2}$$

or

$$t > 2\Delta t$$

Thus one can expect the scheme to be stable for

$$k \geq 2$$

8. Discretization Modification to the Scheme: Four cases, resulting from all combinations of the two parameters, are explored. First:
    a. The forward Euler discretization scheme can be applied directly to u
    b. The forward Euler discretization scheme can be applied directly to tu as seen above.
9. First Time Step Data Update: Second, for the data at the first time step:
    a.

$$u_k = \frac{1}{t_k}$$

    using the given initial conditions

    b.

$$u_k = \frac{1}{t_k} + \frac{1}{2}(a + b)$$

    using the local expansion

$$u(t) \sim \frac{1}{t} + \frac{1}{2}(a + b) + \frac{t}{12}(a - b)^2 + \cdots$$

    as

$$t \to 0$$

10. Improvements from Discretization/Data Modifications: Almgren (2009) carries a demonstration of the example solutions. The combination of improved initial data, with a time discretization that takes into account the initial singularity, yields far more accurate results than naïve discretization.

Space Discretization

1. Diffusion and Convection Terms Discretization: Almgren (2009) uses a 3-point standard discretization scheme for the diffusion term, and upwind differencing for the convection term (see, for example, Le Veque (1992)). A forward Euler time discretization scheme with the correction seen earlier is used; thus a small time step is used for stability.
2. Initial Time Node Discretization Scheme: For initial data, the asymptotic expressions

$$u(\tau, \xi) = \frac{e^{\xi}}{\tau} + u_0(\xi) + \tau u_1(\xi) + \cdots$$

$$\tau \to 0$$

$$u_0(\xi) = -\frac{1}{2}\left(\xi - \frac{1}{2}\beta^2\right)e^{\xi}$$

and

$$3u_1(\xi) = K^2 e^{-\mathscr{y}\xi} - e^{-\xi}{u_0}^2(\xi) - \xi {u_0}'(\xi) + \frac{1}{2}\beta^2 {u_0}''(\xi)$$

or

$$u_1(\xi) = \frac{1}{3}K^2 e^{-\mathscr{y}\xi} + \frac{1}{12}e^{\xi}\left[\left(\xi + 1 - \frac{1}{2}\beta^2\right)^2 - (1 + \beta^2)\right]$$

are used at an initial time

$$t = k\Delta t$$

3. Discretizing the Initial Trade Date: It is more convenient to discretize

$$v(\tau, \xi) = e^{-\xi}u(\tau, \xi)$$

rather than u directly; from

$$v = \frac{x}{\delta_L}e^{-\xi}u(\tau, \xi, \zeta)$$

this is the instantaneous trade rate, except for the dimensional factor. The PDE for v is easily derived from

$$\frac{\partial u(\tau, \xi, \zeta)}{\partial \tau} + \xi\frac{\partial u(\tau, \xi, \zeta)}{\partial \xi} = K^2 e^{-\mathscr{y}\xi} - e^{-\xi}u^2(\tau, \xi, \zeta) + \frac{1}{2}\beta^2\frac{\partial^2 u(\tau, \xi, \zeta)}{\partial \xi^2}$$

4. Left Far Field Boundary Condition: A finite spatial domain

$$-\Xi \le \xi \le \Xi$$

is used. At the left boundary

$$\xi = -\Xi$$

the far field solution

$$u(\tau, \xi) \sim K e^{-\frac{1}{2}(\mathscr{y}-1)\xi} \coth\left(K\tau e^{-\frac{1}{2}(\mathscr{y}+1)\xi}\right)$$

$$\xi \to -\infty$$

is used.

5. <u>Right Far Field Boundary Condition</u>: At the right boundary

$$\xi = +\Xi$$

the "natural" boundary conditions

$$\frac{\partial^2 v(\tau, \xi, \zeta)}{\partial \xi^2} = 0$$

are used. Since the convective term is flowing outwards, the effect of the boundary conditions is confined to a narrow boundary layer.

Almgren (2009, 2012) Sample Solutions

1. <u>Runs for $\mathscr{y} = 1$, $\beta = 1$, and $K = 0.1$</u>: Almgren (2009, 2012) illustrate the computed solution of the PDE for

$$\mathscr{y} = 1$$

$$\beta = 1$$

and

$$K = 0.1$$

As noted above, the natural log of $e^{-\xi}u$ - the dimensionless trade rate as a fraction of the shares remaining – is examined.

2. High Impact/Low Volatility Behavior: As expected, when τ is small, the trade rate becomes large like $\frac{1}{\tau}$. When ξ is large positive the market impact is high and the volatility is low, so the optimal strategy trades very slowly except near expiration.
3. Low Impact High Volatility Behavior: When ξ is large negative, the market impact is low and the volatility is high, so the optimal strategy trades rapidly. As

$$\tau \to \infty$$

the solution approaches the steady state

$$\log\left(e^{-\xi}u\right) = \log K - \xi$$

4. Realization of the Market State: Almgren (2009, 2012) also show a realization of the market state process $\xi(t)$ used for the trajectory simulations. With the "coordinated variation" approximation, the market moves back and forth between a high-activity regime with low impact and high volatility (small ξ) and a low-activity regime with high impact and low volatility (large ξ).
5. Multi-Market Cycle Simulation Span:

$$\beta = 1$$

has been assumed so that the root mean square fluctuation of $\xi(t)$ is $\frac{1}{2}$. The mean reversion time is

$$\delta = 1$$

so that with a time

$$T = 10$$

several market cycles are experienced.

6. Response Dynamics across Risk Aversion: The optimal trading trajectories are examined for several values of the risk aversion parameter K and are compared with the non-adaptive trajectories computed in the mean market state. From these examinations the dynamic nature of the response is very clear.
7. Trading Slow Down/Speed Up: For instance, in the simulation, Almgren (2009, 2012) shows that around

$$t = 1$$

the market state is poor, so all trajectories trade slowly and fall behind the static ones. Around

$$t = 1.5$$

there is a brief burst of liquidity, and all the trajectories accelerate in response.
8. Impact of Urgency of the Trajectory: The trajectories with lower urgency have more shares remaining to trade, so they are able to react more than the high-urgency trajectories, which have completed a substantial fraction of the goal by that time. Thus the lowest urgency trajectory is able to adapt and is able to benefit from an eventual large and prolonged burst of liquidity.
9. "Rolling Horizon" vs. Fast Trading: For large risk aversion (fast trading), this approximate strategy is almost identical to the optimum.
10. "Rolling Horizon" vs. Slow Trading: For smaller risk-aversion (slow trading), the rolling horizon strategy almost rigidly follows a straight line trajectory, while the true optimum is able to adapt to the varying market state even when its profile is linear.

11. Validity of the "Rolling Horizon" Approximation: In general the rolling horizon strategy seems to be an adequate approximation when the risk aversion is relatively high.

References

- Almgren, R. F., and N. Chriss (2000): Optimal Execution of Portfolio Transactions *Journal of Risk* **3 (2)** 5-39.
- Almgren, R. F., and J. Lorenz (2007): Adaptive Arrival Price, in: *Algorithmic Trading III (B. R. Bruce, editor)* **Institutional Investor** 59-66.
- Almgren, R. F. (2009): Optimal Trading in a Dynamic Market.
- Almgren, R. F. (2012): Optimal Trading with Stochastic Liquidity and Volatility *SIAM Journal of Financial Mathematics* **3 (1)** 163-181.
- Bouchaud, J. P., J. D. Farmer, and F. Lillo (2009): How Market slowly Digest Changes in Supply and Demand, in: *Handbook of Financial Markets: Dynamics and Evolution* 57-160 **North Holland** San Diego.
- Engle, R., and R. Ferstenberg (2007): Execution Risk: It is the same as Investment Risk *Journal of Portfolio Management* **33** 34-44.
- Gatheral, J (2010): No Dynamic Arbitrage and Market Impact *Quantitative Finance* **10** 749-759.
- Gatheral, J., and R. C. A. Oomen (2010): Zero Intelligence Realized Variance Estimation *Finance and Stochastics* **14** 249-283.
- Geman, H., D. B. Madan, and M. Yor (2001): Time Changes for Levy Processes *Mathematical Finance* **11 (1)** 79-96.
- Grinold, R. C., and R. N. Kahn (1995): *Active Portfolio Management* **Probus Publishing**.
- Jones, C. M., G. Kaul, and M. L. Lipson (1994): Transactions, Volume, and Volatility *Review of Financial Studies* **7** 631-651.

- Le Veque, R. J. (1992): *Numerical Methods for Conservation Laws 2nd Edition* **Birkhauser** Basel.
- Lorenz, J. (2008): *Optimal Trading Algorithms* Ph. D. **ETH Zurich**.
- Lorenz, J., and R. F. Almgren (2011): Mean-Variance Optimal Adaptive Execution *Applied Mathematical Finance* **18** 395-422.
- Perold, A. F. (1988): The Implementation Shortfall: Paper versus Reality *Journal of Portfolio Management* **14 (3)** 4-9.
- Schied, A., and T. Schoneborn (2009): Risk Aversion and Dynamics of Optimal Liquidation Strategies in Illiquid Markets *Finance and Stochastics* **13 (2)** 181-204.
- Schied, A., T. Schoneborn, and M. Tehranchi (2010): Optimal Basket Liquidation for CARA Investors is Deterministic *Applied Mathematical Finance* **17** 471-489.
- Tse, S. T., P. A. Forsyth, J. S. Kennedy, and H. Windcliff (2013): Comparison between Mean-Variance Optimal and Mean Quadratic Variation Optimal Trading Strategies **20 (5)** 415-449.
- Walia, N. (2006): *Optimal Trading: Dynamic Stock Liquidation Strategies* Senior Thesis **Princeton University**.

1. Tight Skew: α_T
2. Loose Skew: α_L
3. Tight Width: ω_T
4. Loose Width: ω_L
5. Algorithmically generated Ideal Mid Cash Price: Π_{Ideal}
6. Position: P (expressed in cumulative net position per unit under consideration – firm/desk/trader)
7. Position Pivot: P_{Pivot}. Dimensionless ontological view of the scaling position metric – roughly equivalent to the Reynolds' number of market making position units. Expressed in currency units.
8. Risk: R (expressed in cumulative net risk per unit under consideration – firm/desk/trader)
9. Risk Pivot: R_{Pivot}. Dimensionless ontological view of the scaling risk metric – roughly equivalent to the Reynolds' number of market making risk units. Expressed in PV01 currency units.

Framework Glossary

1. Equilibrium quantity: Quantity that only changes with the macro drivers/factors, and not the technical factors. Typically stable, but jumpy and undergoes changes when drivers shift – and introduces perturbations on the disequilibrium quantities.
2. Disequilibrium quantity: Quantity that changes with the technical, transient factors.

Width/Skew/Size Estimation Models

1. Tight Models:

- Tight models estimate the market making quantities on a trader/firm/desk independent manner.
- They estimate the "secular" market making parameters – width, skew, and size for either the Market Making Outputs or the Axe Outputs – estimate them based on classes of input parameters.
- For each input parameter class, the following are needed:
    a. A proxy that serves as a quantitative estimate of the desired parameter class.
    b. Segmentation of the proxy over the sub-classified parameter set.

2. Input Class => Risk Profile:
    - Captures all the cumulative risk components => the credit/solvency, market, and liquidity risk behind the issue.
    - Proxy => CDS Spread, rating, bond basis
    - Sub-classification => Issue, issuer, and sector.
3. Input Class => Liquidity:
    - Captures the frequency and volume of the trade flow of a given issue, and the ease of getting in and getting out at the given side.
    - Proxy:
        i. Aggregated periodic (e.g., daily) volume for each side (buy/sell).
        ii. Aggregated periodic (e.g., daily) notional for each side (buy/sell).
    - Sub-classification => Issue, issuer, sector, and the instrument universe.
4. Firm/Desk/Trader level parameters: These provide aggregated controls for trading.
    - Net Position => vital metric for inventory control.
    - Risk limits => to control/manage exposure to specific granules – issue, issuer, tenor, sector, unit etc.
5. Monitor Mobility: Certain measures such as PV01 based risk, inventory, etc. are more easily human-monitored, so they are done daily. Others (such as tenor 01s) are less easily monitored, so they are done infrequently.

Market Making System SKU

1. Intra day Curve Generation Scheme
2. Mid Price Estimation Models
    i. Accommodate different mid price estimation models, and their respective parameters
3. Algorithmic Quote Construction => used for generating venue/ECN independent width/skew/size [composed of tight/loose components]. Broadly speaking achieves the following:
    i. Specific parameters to control skew for targeted alpha generation strategies
    ii. Accommodate different width and size estimation models, and their respective parameters
    iii. Venue-independent base quote synthesis/construction
    iv. Circuit breaker heuristics
    v. Policy driven/policy enforcement/policy control applied at this level
4. Quote Management: Publishing/tailoring the constructed quote towards specific venues (possibly with order routing applied at this stage).
    i. Venue specific rules (and thereby external vendor incorporations, like Broadway etc. at this stage.

Market Making Parameter Types

1. <u>Model Parameters</u>: Parameters for generation of algorithmic generation of width, skew, and size.
2. <u>Quote Generation Control Parameters</u>
3. <u>Quote Heuristics Control</u>
4. <u>Quote Management Control</u>

Intra-day Pricing Curve Generation Schemes

1. Issue Benchmark Bonds: The following set of threshold criteria are used to determine the issuer specific benchmark bonds:

   ii. Threshold of daily TRACE volume/number of trades
   iii. Threshold of outstanding notional
   iv. Only senior obligations
   v. Some combination of the following threshold of the ratios:
      - $\frac{CUMULATIVEDAILYISSUETRACEVOLUME}{OUTSTANDINGNOTIONAL}$
      - $\frac{CUMULATIVEDAILYISSUETRACEVOLUME}{CUMULATIVEDAILYISSUERTRACEVOLUME}$

2. Benchmark bonds basis tracking: Track the bid side and ask side credit basis of the benchmark bonds from each TRACE print, using EMA VWAP/TWAP from the intra-day rates/credit curves. This will be the attempt to estimate the mid credit basis for the, and it is generally well behaved.

- Need to find a way to accommodate the institutional closing CDS mid marks and the benchmark bonds into the credit curve construction – these are highly valid points.

3. Liquid vs. illiquid: Typical liquid securities' quote may be proxied out of print (or at least EMA'd). Intra-day quote generation, however, is materially important for illiquid securities.
4. Intra-day credit curve generation inputs: Need a way to generate the credit curve from
   i. The CDS marks
   ii. The basis-adjusted benchmark bonds
   iii. It always needs to be used in conjunction with tension splines.
   iv. Also need intra-day TRACE series to update the basis (direct or EMA) – will use this to establish the intra-day relationship between the CDS nodes and the TRACE cut-off threshold).
5. Intra-day credit curve updating:

a. Use the relationship grid between CDS 5Y, the off-tenors, and the benchmark bonds
b. Any change in any of them automatically re-adjusts using the set relationships.
c. CDS Curves are trader set; bond basis are EMA'd from the TRACE series using the prior credit curve
d. Relationships are either reviewed daily EOD

6. Live updating of bond prices: Use the live curve (either pure CDS, or a mixture of CDS/bond instruments) to extract the basis of each print, and then EMA that to generate the bond live prices.

Mid-Price Models

i. Definition: Computed theoretical mid-price, as to where the next print should be – assuming zero transaction costs, zero position/risk constraints, and infinite liquidity. Mid Price is an ***Equilibrium Quantity***.
ii. Estimation parameters: Typical mid price estimation parameters are: the IR curve, the survival curve, and the recovery curve. The other possible drivers are: funding curve – typically for long position, and repo curve – typically for shorts.

Width Models

1. Tight Width: Computed theoretical width, after accounting for the issue liquidity and the issue riskiness. Tight width is the first in the set of disequilibrium quantities. Tight width is:
    a. Proportional to issue risk (combination of credit and market risk – not counter party risk).
    b. Inversely proportional to liquidity

Skew Models

1. Tight Skew: This measure how far the last print has been OFF from the theoretical mid price. Thus Tight Skew is representative of the alpha potential – for a theoretical mid price that chases the print in a sequence, the tight skew is zero.
2. Tight Bid Skew and Tight Ask Skew: This is an alternative SKU – instead of tight width and tight skew cognitive view, tight bid/ask skew parameters are determined only from their corresponding liquidity and flow metrics (i.e., bid/ask liquidity metrics).
3. Loose Skew: Simply put, loose skew is:

$$\alpha_L = \max\left(\frac{P}{P_{Pivot}}, \frac{R}{R_{Pivot}}\right)$$

4. Heuristic Checks on Loose Skew: Following checks applied to round out quoting:
    1. Ceiling/floor applied
    2. Maximum cutoff for width
    3. Best right skew – bid becomes ask.
    4. Best left skew – ask becomes bid.

Size Models

1. Tight bid size/tight ask size: Basically, tight size is inversely proportional to tight width, to within normalized bounds.

Heuristics Control

1. Can Buy/Can Short: Can But/Can Short => whether the bid/ask stays within the LONGABLE/SHORTABLE cutoff.

2. ECN Threshold Cross: Check to see if there is a cross between the published bid/ask and a given ECN's bid/ask.

Published Market Quote Picture

1. Bid/Ask Sizes: Truncated to their appropriate rounding.
2. Bid Price: $\Pi_{Ideal} - \frac{1}{2}\omega_L \alpha_L$
3. Ask Price: $\Pi_{Ideal} + \frac{1}{2}\omega_L \alpha_L$
4. Bid/Ask Prices rounded downward/upward to their appropriate increments.

Flow Analysis

1. Dimensionless flow classifier: If the metric (ADV etc.) is greater than a specific threshold, then the flow becomes "turbulent", else it is "laminar".
2. Flow Potential: Skew of all kinds is related to the flow driver/equilibration strength.