

# Transaction Cost Analytics in DRIP

Lakshmi Krishnamurthy

**v2.19**, 27 August 2016

**Optimal Execution of Portfolio Transactions**

**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 , they define opportunity costs as the expected costs of 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 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 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 units of a security that they want to completely liquidate before time . 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 into units of length

and defines the discrete times

for

The *trading trajectory* is defined to be the list where is the number of units that the seller plans to hold at time .

1. Outright/Re-balanced Trajectories: The initial holdings is

and liquidation at time requires

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

the initial position, and

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

and

1. Outstanding Holdings/Incremental Trade Lists: Equivalently, a strategy may be specified using the “trade list” where

is the number of units that the seller will sell between times and . Clearly, and are related by

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

and

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

**Price Dynamics**

1. Exogenous/Endogenous Price Move Factors: Suppose that the initial security price is so that the initial market value of the position is . The securities’ price evolves according to two exogenous factors – volatility and drift, and one endogenous factor – market impact.
2. Market Forces vs. Trading Impact: Volatility and drift are assumed to the result of market forces that occur randomly and independent of the trading.
3. Earlier Literature on Market Impact: 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. Origin of the Market Impact: 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. Definition of Temporary Market Impact: *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. Definition of Permanent Market Impact: *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. Price Evolution Stochastic Difference Equation: Almgren and Chriss (2000) assume that the security price evolves according to the discrete random walk

for

1. Glossary of the Equation Terms: Here represents the volatility of the asset, ’s are draws from independent random variables each with zero mean and unit variance, and the permanent impact function is a function of the *average rate* of trading

during the interval to .

1. 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.
2. 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 in

scale with . 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 between times and , 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 is sufficiently large, the execution price may steadily decrease between and 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 , the temporary drop in the average price per share caused by trading at an average rate during one time interval.
4. Net Price Received at Execution: Give this, the actual price per share received on sale is

but the effect of does not appear in the next “market” price .

1. Choice of Market Microstructure: The functions in

and in

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 sold in each time interval times the effective price per share received on that sale. It is readily computed as
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 represents the total impact from the volatility.
5. The Permanent Market Impact Term: The permanent market impact term 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 is the price drop due to selling, acting only on the units sold during the period.
7. The Total Cost of Trading: The *total cost of trading* is the difference 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 for the expected short-fall and 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

The units of are in dollars, and the units of are dollars squared.

1. Distribution of Implementation Short-fall: The distribution of the short-fall is Gaussian if is Gaussian, in any case if is large, it is very nearly Gaussian.
2. Almgren and Chriss Minimizer Utility: Almgren and Chriss (2000) devote much of their paper to finding trajectories that minimize for various values of . They demonstrate that for each value of there corresponds a unique trading trajectory such that 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, has the form

in which the constant has units of ($/share)/share.

1. Corresponding Execution Time Security Price: With this form, each units sold depresses the price per share by regardless of the time taken to sell units.

readily yields

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

becomes

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

where is the sign function.

1. Estimating the Fixed Costs of Execution: The units of are $/share, and those of are ($/share)/(share/time). A reasonable estimate for is the fixed cost of selling, such as half of bid-ask spread plus premium.
2. Estimating the Linear Impact Coefficient: It is more difficult to estimate 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

to be most doubtful.

1. Total Temporary Impact Function: The linear model

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

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

and

the expectation of the impact costs

becomes

in which

1. Strictly Convex Nature of : Clearly is a strictly convex function as long as

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

1. and Computation Illustration: To illustrate, Almgren and Chriss (2000) compute and 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.
2. 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

and

1. Minimum Impact and : From

and

one has

and from

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

remains finite, and and have finite limits.

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

which results in

and

1. Minimum Variance Limits: The trajectory has the smallest possible variance –equal to zero – because of the way time has been discretized in the model above. If is large and hence is short, then on the full initial portfolio, one takes a price hit that can be arbitrarily large.
2. 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

That is, for a given maximum level of variance

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

1. Convex Objective Function and Domain: Since is convex, the set

is convex – it is a sphere – and since is strictly convex, there is a unique minimizer .

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

has higher expected costs than for the same or lower variance, and can never be more efficient.

1. Efficient Frontier of Optimal Strategies: Thus the family of all possible efficient (optimal) strategies is parametrized by a single variable representing all possible maximum levels of variance in transaction costs. This family is referred to as *the efficient frontier of optimal trading strategies*.
2. Introducing KKT Type Constraint Multipliers: The constrained optimization problem

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

1. Frontier as a Function of : If

is strictly convex, and the above minimizer has a unique solution . As varies, sweeps out the same one parameter family, and thus traces out an efficient frontier.

1. as a Risk Aversion Parameter: The Parameter has a direct financial interpretation. It is already apparent from

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

1. as an Efficient Frontier Curvature: In fact, is the curvature – second derivative – of a smooth utility function, as will be made more precise eventually.
2. Solution given and : For given values of the parameters, problem

can be solved by various numerical techniques depending on the functional forms chosen for and . 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. Optimal Solution in Trajectory Space: With from

and from

and assuming that does not change sign, the combination

is a quadratic function of the control parameters ; it is strictly convex for

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

for

1. Combinations of Linear Difference Equations: Then

is equivalent to

with

1. Abstracted and Re-factored Parameter Set: Note that

is a linear difference equation whose solution may be written as a combination of the exponentials where satisfies

The tilde’s on and denote an correction; as

one has

and

1. Trading Trajectory/Trade List Solutions: The specific solution with

and

is a trading trajectory of the form

and the associated trade list is

where and are the hyperbolic sine and cosine functions, and

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

1. Monotonicity of the Trading Trajectory: One has

as long as

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 .

1. 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.
2. Approximation under Small Time Step: For a small time step one has the approximate expression

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

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

and

which reduce to

in the limits

**The Half-Life of a Trade**

1. Definition of the Half-Life: Defining

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

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

the trader will still liquidate his position on a time scale . The half-life is the intrinsic time scale of the trade.

1. Half Life Smaller than : For a given the ratio

tells us what factors constrain the trade. If

then the intrinsic half-life of the trade is small compared to the imposed time ; this happens because temporary costs are very small, because volatility is very large, or because of high risk aversion.

1. Impact of Small Half-Life: In this case the bulk of the trading will be done well in advance of the time . Viewed on a time scale the trajectory will look like a minimum variance solution
2. Very High Half Life Limit: Conversely if

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

one approaches the straight line minimum cost strategy

1. 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 is held constant.
2. 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.
3. 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.
4. 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.
5. Size Dependent Temporary Impact Parameter: A simple practical solution to this problem is to choose different values of - 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

They also illustrate the trajectories corresponding to a few sample points on the frontier.

1. Trajectory corresponding to Positive : Their first trajectory has

– 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.

1. Trajectory corresponding to Zero : Their second trajectory has

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.

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

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.

1. 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.
2. Trajectory corresponding to Negative : Finally their trajectory has

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 where is the total wealth. This function may be characterized by its risk-aversion coefficient
3. Approximation in Estimating the : If the initial portfolio is fully owned, then as the transfer of assets happens from the risky stock into the alternative riskless investment, remains roughly constant, and one may take to be a constant throughout the trading period. If the initial portfolio is highly leveraged, then the assumption of constant is an approximate one.
4. Formulation of the Optimal Execution Strategy: For short time horizons and small changes in the higher derivatives of may be neglected. Thus choosing an optimal execution strategy is equivalent to minimizing the scalar function

The units of are ; one is willing to accept an extra square of variance if it reduces the expected cost by .

1. Constructing Family of Optimal Paths: The combination is precisely the one used to construct the efficient frontier seen earlier; the parameter , 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.
2. Static Nature of Optimal Path: Returning now to an important point raised earlier, the computation of optimal strategies by minimizing 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.
3. 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

and ends at a time

the optimal strategy computed at

is simply a continuation from

to

of the optimal strategy computed at time

1. Proof Steps: General/Specific Functions: 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.
2. Proof Steps: Function Time Shift: First suppose that at time , where

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

with replaced by , replaced by , and replaced by . Using the subscript to denote the strategy computed at time one would have

and the trade lists

1. Proof Step: Recovering Optimal Solutions: It is then apparent that if is the optimal solution from

with

then

and

where

and

are the strategies from

and

1. Proof Step: Non-linear Impact: For general non-linear impact functions and the optimality condition

is replaced by a second-order *non-linear* difference relation. The solution beginning at a given time is determined by the two boundary values and

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

1. 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.
2. 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

the value-at-risk of defined is defined to be the level of transaction costs by the trading strategy that will not be exceeded percent of the time. Put another way, it is the percentile level of transaction costs for total costs of trading .

1. 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 from the mean of the inverse of the cumulative normal distribution function, and the value at risk for the strategy is given by
2. Relation to Implementation Short-fall: That is, with a probability the trading strategy will not lose more than of its market value in trading. Borrowing from the language of Period (1988), the implementation shortfall of execution will not exceed more than a fraction of the time. A strategy is efficient if it has the minimum possible value at risk for the confidence level .
3. Execution Trajectory Optimized for VaR: Note that is a complicated non-linear function of composing ; it can be easily evaluated for any given trajectory, but finding the minimizing trajectory directly is difficult.
4. 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 corresponding to a given value of . 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.
5. Almgren-Chriss Optimal VaR Illustration: Almgren and Chriss (2000) produce an illustration of the above, using the square root of variance in the -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 they simply find a tangent to the curve where the slope is .
6. 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 is now held constant, necessarily changes.
7. 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.
8. Liquidity Adjusted Value at Risk: Despite this shortcoming, Almgren and Chriss (2000) the smallest possible value of 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)).
9. 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.
10. 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 to Mapping: A crucial insight is that the curve defining the efficient frontier is a smooth convex function mapping the levels of variance to the corresponding minimum mean transaction cost levels.
4. Region around the Naïve Strategy: Write for the mean and variance around the naïve strategy. Regarding as a point on the smooth curve defined by the frontier, evaluated at is equal to zero. Thus for near one has

where

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

1. 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.
2. 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.
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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 and that they initially have one million shares, for an initial portfolio size of million. The stock will have annual volatility, expected annual rate of return, a bid-ask spread of , and a median daily trading volume of million shares.

**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 (2000): Value Under Liquidation *Risk* **12 (12)**.
* 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.
* 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.
* 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.
* 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.
* 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**.

# Algorithmic Market Making

### Symbology Glossary

1. Tight Skew: 
2. Loose Skew: 
3. Tight Width: 
4. Loose Width: 
5. Algorithmically generated Ideal Mid Cash Price: 
6. Position: (expressed in cumulative net position per unit under consideration – firm/desk/trader)
7. Position 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: (expressed in cumulative net risk per unit under consideration – firm/desk/trader)
9. Risk 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:
  + - * 1. A proxy that serves as a quantitative estimate of the desired parameter class.
        2. Segmentation of the proxy over the sub-classified parameter set.

1. 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.

1. 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:
  + 1. Aggregated periodic (e.g., daily) volume for each side (buy/sell).
    2. Aggregated periodic (e.g., daily) notional for each side (buy/sell).
* Sub-classification => Issue, issuer, sector, and the instrument universe.

1. 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.

1. 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
     1. 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:
     1. Specific parameters to control skew for targeted alpha generation strategies
     2. Accommodate different width and size estimation models, and their respective parameters
     3. Venue-independent base quote synthesis/construction
     4. Circuit breaker heuristics
     5. 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).
     1. Venue specific rules (and thereby external vendor incorporations, like Broadway etc. at this stage.

## Market Making Parameter Types

1. Model Parameters: Parameters for generation of algorithmic generation of width, skew, and size.
2. Quote Generation Control Parameters
3. Quote Heuristics Control
4. Quote Management Control

## 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:

* + 1. Threshold of daily TRACE volume/number of trades
    2. Threshold of outstanding notional
    3. Only senior obligations
    4. Some combination of the following threshold of the ratios:
* 
* 

1. 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.

1. 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.
2. Intra-day credit curve generation inputs: Need a way to generate the credit curve from

The CDS marks

The basis-adjusted benchmark bonds

It always needs to be used in conjunction with tension splines.

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).

1. Intra-day credit curve updating:
   * + - 1. Use the relationship grid between CDS 5Y, the off-tenors, and the benchmark bonds
         2. Any change in any of them automatically re-adjusts using the set relationships.
         3. CDS Curves are trader set; bond basis are EMA’d from the TRACE series using the prior credit curve
         4. Relationships are either reviewed daily EOD
2. 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

* + 1. 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***.
    2. 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:
   1. Proportional to issue risk (combination of credit and market risk – not counter party risk).
   2. 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:



* 1. 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: 
3. Ask Price: 
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.