Optimal Liquidation

Robert Almgren* and Neil Chriss†

January 14, 1998

Abstract

We consider the problem of portfolio liquidation with the aim of minimizing a combination of volatility risk and transaction costs arising from permanent and temporary market impact. For a simple linear cost model, we explicitly construct the *efficient frontier* in the space of time-dependent liquidation strategies, which have minimum expected cost for a given level of uncertainty. We consider the risk-reward tradeoff both from the point of view of classic mean-variance optimization, and from the standpoint of Value at Risk. This analysis leads to general insights into optimal portfolio trading, and to several applications including a definition of *liquidity-adjusted value at risk*.

*The University of Chicago, Department of Mathematics, 5734 S. University Ave., Chicago IL 60637; almgren@math.uchicago.edu

†Morgan Stanley Dean Witter and Courant Institute of Mathematical Sciences; neilc@cims.nyu.edu

Contents

Broker dealers are frequently faced with the task of buying or selling large blocks of a single stock or large baskets of multiple stocks. For example, program trading desks perform transition trades for pension funds facilitating the move from one manager's positions to another's. Clients engage broker-dealers to perform such trades and must evaluate performance against a suitable benchmark. Thus broker dealers face the problem of minimizing transaction costs of large trades that take place over fixed time periods, while buy side participants face the problem of devising reasonable performance benchmarks for these services.

These problems, two sides of the same coin, have an underlying structure given by the interaction between two ingredients:

- *Market impact.* If a large trade is executed too rapidly, costs will be incurred as the trades move the market in an adverse direction.

- *Volatility risk.* On the other hand, if the trade is executed too slowly, then the position is subject to risk during the time that the shares remain in the portfolio.

These two quantities must be played off against each other by taking account of the desired performance characteristics of the various participants.

In this paper we define the notion of an efficient or optimal liquidation strategy for a basket of securities, and study the notion under a simple model of price movements. The main results of our papers are the following:

- A method for determining the optimal method for liquidating a basket of securities.

- A method for evaluating performance by computing the expected value of a trading strategy and its standard deviation.

- The introduction of an "efficient frontier" of optimal trading strategies, each offering a different risk/reward tradeoff.

- General insights into the relation between utility function and portfolio trading independent of specific features of the stock price movements and transaction cost model.

In addition, we offer insight into the question of the magnitude of the transaction costs associated with liquidation under extreme market conditions. This has several interesting applications to risk management, including offering a simple definition of what we call L-VaR (liquidity-adjusted value at risk).

The central point of our discussion is the observation that the cost of trading—the difference between the initial market value and the value realized after liquidation—is a random variable, whose mean and variance at the initial time depend on the liquidation strategy to be followed. This observation allows us to introduce an efficient frontier of trading strategies and define the concept of a risk/reward tradeoff for trading strategies: reward is low transaction costs, and risk is the level of variability of transaction costs.

This is not a paper about transaction cost models. We do not propose to give a method for parametrizing the transaction cost models stated in the subsequent sections. Rather, we give a complete analysis of the consequences of a given transaction cost model on the financial economics of portfolio trading. In fact, the entire analysis carried out here can be redone for almost any transaction cost function one can write down, and the basic conclusions will remain (see the Appendix).

The approach we have taken toward the problem of optimal liquidation is new. The problem of optimal trading has been clearly formulated by Perold (1988). Previous analytical work has typically focused on a single point along our efficient frontier; see, for example, Bertsimas and Lo (1998) and Subramanian (1997b, 1997a). Empirical studies of market impact have been carried out by Kraus and Stoll (1972), Holthausen, Leftwich, and Mayers (1987), and Chan and Lakonishok (1993, 1995). Vayanos (1997) has considered a model of market microstructure and its consequences for optimal trading strategies.

1 Trading Model

Below we give a formal definition of a trading strategy for the liquidation of a single stock or other risky asset. Our definition is in terms of selling a (presumably) large block of stock, but is equally relevant to other assets such as currencies or bonds. In fact, the method generalizes to any trade (buy or sell) in which both transaction costs and volatility are an issue. We present this in a discrete-time framework, but the extension to continous-time is immediate.

1.1 Trading strategy

Suppose that we hold a block of X shares of a single stock or other risky asset, that we want to completely liquidate by time T. We divide T into N intervals of length τ, and define the discrete times $t_k = k\tau$, for $k = 0, \ldots, N$. We define a *trading strategy* to be the list $x_0, \ldots, x_N$, where x_k is the number of shares that we will hold at time t_k. Our initial holding is $x_0 = X$, and liquidation at time T requires $x_N = 0$.

We may also specify our strategy by giving the "trade list" $n_1, \ldots, n_N$, where $n_k = x_{k-1} - x_k$ is the number of shares that we sell between times t_{k-1} and t_k. Clearly, x_k and n_k are related by

$$x_k \;=\; X - \sum_{j=1}^{k} n_j \;=\; \sum_{j=k+1}^{N} x_j, \qquad k = 0, \ldots, N.$$

We consider more general programs of simultaneously buying and selling several securities in Section 5. For notational simplicity, we have taken all the time intervals to be of equal length τ, but this restriction is not essential.

Since the length τ of the time interval is a rather arbitrary quantity, it will be more convenient to work with the instantaneous *rate* of trading, in shares per unit time; this will also make it easier to pass to the continuous-time limit. During the interval t_{k-1} to t_k, this rate is

$$v_k \;=\; \frac{n_k}{\tau} \;=\; \frac{1}{\tau}(x_{k-1} - x_k), \qquad k = 1, \ldots, N.$$

Note that each n_k or v_k is a "backwards" difference. Thus x_k and v_k represent the same piece of information: our choice at time t_{k-1} of how many shares we wish to hold at time t_k.

We shall take the trading strategy to be *deterministic*, even in the presence of random uncertainty in market events during the liquidation. That is, we imagine that we determine the entire strategy at the beginning of trading, and we evaluate the risk and reward of this strategy without permitting the strategy to evolve in time. Although this may appear to be a rather extreme hypothesis, we shall justify it more fully below.

1.2 A model for stock price movements

Suppose that the initial stock price is S_0, so that the initial market value of our position is XS_0. The stock's price evolves according to three influences: volatility, drift, and market impact. Volatility and drift are assumed to be the result of market forces that occur randomly and independently of our trading. Market impact is a direct result of our trading. As market participants begin to detect the volume we are selling (buying) they naturally adjusts their bids (offers) downward (upward). Later we will discuss two distinct types of market impact: temporary and permanent. Our discussions largely reflect the work of Kraus and Stoll (1972), and the subsequent work of Holthausen, Leftwich and Mayers (1987, 1990), and Chan and Lakonishok (1993, 1995).

Temporary impact means those due to temporary imbalances in supply in demand caused by our trading. Such transaction costs are one-time costs applicable to single trades. Permanent costs, on the other hand, refer to a shift in the equilibrium price of the stock under consideration due to trading, which last at least for the life of our liquidation.

We assume that the stock price evolves according to the discrete random walk

$$\begin{aligned} S_k &= S_{k-1} + \sigma\tau^{1/2}\xi_k + \mu\tau - \tau\, g(v_k) \\ &= S_0 + \sigma \sum_{j=1}^{k} \tau^{1/2}\xi_j + \mu t_k - \sum_{j=1}^{k} \tau g(v_j), \end{aligned} \tag{1}$$

for $k = 1, \dots, N$. Here σ represents the volatility of the asset, μ is an expected growth rate, the ξ_j are draws from independent random variables each with zero mean and unit variance, and $g(v)$ is a function of the rate of trading v. We assume that the proceeds from any sales are invested in some alternative investment that receives the riskless rate of return. In such a case we take μ to be the *excess* return over the return on the alternative investment.

In the absence of market impact, our model says that S_k executes a arithmetic Brownian random walk, so that at time t_k its mean is $S_0 + \mu t_k$ and its variance is $\sigma^2 t_k$.

Arithmetic vs. geometric Brownian motion The random walk above would be a standard *geometric* Brownian motion with constant drift $\bar{\mu}$ and volatility $\bar{\sigma}$, if we defined $\bar{\mu}$ and $\bar{\sigma}$ such that

$$\mu = S\,\bar{\mu}, \qquad \sigma = S\,\bar{\sigma},$$

and kept $\bar{\mu}$ and $\bar{\sigma}$ constant as S evolved. Because we are interested in short-term "trading" horizons rather than longer-term "investment" horizons, the total changes in S will be small, and we may assume that μ and σ are constant rather than $\bar{\mu}$ and $\bar{\sigma}$.

That is, the difference between arithmetic and geometric Brownian motions is negligible over short time horizons. Moreover the arithmetic model is analytically simpler. Note that our μ and σ must be divided by some reference price S in order to give rates of return in the standard sense. As a practical matter, to determine μ and σ, start with the standard fractional volatility and drift as parametrized in an ordinary Brownian motion and multiply by the "reference price," the price of the security under liquidation at the start of liquidation. We give an example in Section 4.

1.3 Permanent market impact

The function $g(v)$ above represents the *permanent* impact (of trading) on the price of the stock. Here *permanent* refers to the life of the liquidation at hand.

The specific form of the transaction cost function is embodied in the following assumption: as long as we are actively trading the stock, market participants will not bid in substantial volumes at the equilibrium price of the stock.

For now, we shall assume that the impact function is *linear* in our rate of trading. That is, the other participants will bid low (for a sell program) or offer high (for a buy program) in proportion to the average rate we are trading. The mathematical formulation of this assumption is that $g(v)$ has the form

$$g(v) = \gamma v. \tag{2}$$

Here $g(v)$ represents the drop in the price per share per unit time as a response to our selling at a rate of v shares per unit time. The constant γ has units of (\$/share)/share. Thus each n shares that we sell depresses the price per share by γn, regardless of the time we take to sell the n shares.

Substituting linear permanent impact (2) in equation (1) we obtain the price equation:

$$\begin{aligned} S_k &= S_{k-1} + \sigma \tau^{1/2} \xi_k + \mu \tau - \gamma n_k \\ &= S_0 + \sigma \sum_{j=1}^{k} \tau^{1/2} \xi_j + \mu t_k - \gamma (X - x_k). \end{aligned} \tag{3}$$

In the absence of fluctuation and drift, the stock price is a linear function of our instantaneous holdings. This is consistent with the partial equilibrium approach of Jarrow (1992), though he considers more general nonlinear dependence on our portfolio holding.

1.4 Temporary market impact

We imagine the trader plans to sell a certain number of shares n_k between times t_{k-1} and t_k, but may place the order in several smaller slices. If the total number of shares n_k is sufficiently large, the execution price may steadily decrease between t_{k-1} and t_k, partially due to using up the orders at the bid faster than new orders arrive. Thus the trader will suffer some temporary costs that will be offset in the next period when new orders have arrived.

We model this effect by introducing a temporary price impact function $h(v)$, which is the temporary drop in price per share caused by trading at rate v. Then the actual price per share received on sale k is

$$\tilde{S}_k = S_{k-1} - h(v_k),$$

but the effect of $h(v)$ does not appear in the next "market" price S_k.

For now, we consider only the linear model

$$h(v) = \epsilon + \eta v. \tag{4}$$

The units of ϵ are \$/share, and those of η are (\$/share)/(share/time). A reasonable estimate for ϵ is the fixed costs of selling, such as half the bid-ask spread plus fees. It is more difficult to estimate η since it depends on internal and transient aspects of the market microstructure. It is in this term that we would expect nonlinear effects to be most important, and the approximation (4) to be most doubtful. We consider general cost functions in the Appendix.

The linear model for transaction costs (4) is often called a *quadratic* cost model because the total cost incurred by trading n shares in a single unit of time is

$$n\,h\left(\frac{n}{\tau}\right) = \epsilon n + \frac{\eta}{\tau} n^2.$$

That is, the drop in price *per share sold* is linear in the number of shares sold, so the total cost is quadratic in n.

With both linear cost models (2,4), the actual average price we receive on the sale between t_{k-1} and t_k is

$$\begin{aligned}\tilde{S}_k &= S_{k-1} - \epsilon - \eta v_k \\ &= S_0 + \sigma \sum_{j=1}^{k-1} \tau^{1/2} \xi_j + \mu t_{k-1} - \gamma (X - x_{k-1}) - \epsilon - \eta v_k. \end{aligned} \tag{5}$$

1.5 Capture and cost of trading strategies

We define the *capture* of a strategy to be the value we receive upon liquidating the entire sell program. This is the sum of the product of the number of shares $n_k = \tau v_k$ that we sell in each time interval times the effective price per share $\tilde{S}_k$ received on that sale.

Write $X\bar{S}$ for the capture of a strategy, so that $\bar{S}$ is the average price per share received. Using the interchange of sums formula $\sum_{k=1}^{N} \sum_{j=1}^{k-1} = \sum_{j=1}^{N-1} \sum_{k=j+1}^{N}$, we readily compute

$$\begin{aligned} X\bar{S} = \sum_{k=1}^{N} n_k \tilde{S}_k = X S_0 + \sigma \sum_{k=1}^{N} \tau^{1/2} x_k \xi_k + \mu \sum_{k=1}^{N} \tau x_k \\ - \gamma \sum_{k=1}^{N} \tau x_k v_k - \epsilon X - \eta \sum_{k=1}^{N} \tau v_k^2. \end{aligned} \tag{6}$$

The first term XS_0 on the right of (6) is the initial market value of our position; each additional term represents a gain or a loss during the liquidation, and each has a simple interpretation. At step k we hold x_k shares of stock. Each effect that moves the price by δS_k at time t_k changes the market value of our position by $x_k\, \delta S_k$, and the total change is $\sum x_k\, \delta S_k$.

The first term of this type is $\sum \sigma \tau^{1/2} \xi_k x_k$, representing the total effect of volatility. The second is $\sum \mu \tau x_k$, representing the expected total return. And the market impact term $-\sum \gamma \tau v_k x_k = -\sum \gamma n_k x_k$ represents the loss in value of our total position, caused by the permanent price drop associated with selling a small piece of the position.

Of the two terms corresponding to the instantaneous cost, the first, $-\epsilon X$, is the accumulation of the fixed costs, which depend simply on the total amount we sell. The final term, $-\sum \eta \tau v_k^2$, represents the cost of market impact on each individual trade. As observed above, the total cost incurred by a linear price impact function is quadratic in the number of shares traded.

Using the summation by parts formula

$$\sum_{k=1}^{N} \tau x_k v_k \;=\; \sum_{k=1}^{N} x_k (x_{k-1} - x_k) \;=$$
$$= \tfrac{1}{2} \sum_{k=1}^{N} \left(x_{k-1}^2 - x_k^2 - (x_k - x_{k-1})^2 \right) \;=\; \tfrac{1}{2} X^2 - \tfrac{1}{2} \sum_{k=1}^{N} \tau^2 v_k^2,$$

we may put (6) into the simpler form

$$X\bar{S} \;=\; X S_0 \;+\; \sigma \sum_{k=1}^{N} \tau^{1/2} x_k \xi_k \;+\; \mu \sum_{k=1}^{N} \tau x_k$$
$$-\; \tfrac{1}{2}\gamma X^2 \;-\; \epsilon X \;-\; \left(\eta - \tfrac{1}{2}\gamma\tau \right) \sum_{k=1}^{N} \tau v_k^2. \quad (7)$$

Surprisingly, except for a term of size $\mathcal{O}(\tau)$ in the last sum, the contribution of permanent market impact γ depends only on the initial holding X. That is, *the effect of linear permanent impact may be valued independently of the liquidation path.*

The *total cost of trading* is the difference $XS_0 - X\bar{S}$ between the initial book value and the capture. This is the standard *ex-post* measure of transaction costs used in performance evaluations, and is essentially what Perold (1988) calls *implementation shortfall.*

Under our assumptions, this cost is a random variable whose expected value and variance, measured at the initial time, depend on the trading strategy $x = (x_1, \ldots, x_n)$ used to liquidate the position. Write $E(x)$ and $V(x)$ for respectively the expected value and variance of the total cost of trading strategy x. We will informally call $E(x)$ the *expected value* of the strategy and $V(x)$ the *variance* of the strategy.

From (7) we compute

$$E(x) \;=\; -\mu \sum_{k=1}^{N} \tau x_k + \tfrac{1}{2}\gamma X^2 \;+\; \epsilon X \;+\; \left(\eta - \tfrac{1}{2}\gamma\tau \right) \sum_{k=1}^{N} \tau v_k^2 \quad (8)$$
$$V(x) \;=\; \sigma^2 \sum_{k=1}^{N} \tau x_k^2. \quad (9)$$

The units of E are \$; the units of V are $\$^2$. To illustrate, let us compute these quantities for the two most extreme strategies: sell at a constant rate, and sell to minimize variance without regard to transaction costs.

Minimum impact The most obvious strategy is to sell at a constant rate over the whole liquidation period. Thus, we take each

$$v_k = \frac{X}{T} \qquad \text{and} \qquad x_k = \left(1 - \frac{k}{N}\right) X, \qquad k = 0, \dots, N. \tag{10}$$

From (8) we readily compute

$$E = -\tfrac{1}{2}\mu T X \left(1 - \frac{1}{N}\right) + \tfrac{1}{2}\gamma X^2 \left(1 - \frac{1}{N}\right) + \epsilon X + \eta \frac{X^2}{T} \tag{11}$$

and from (9),

$$V = \tfrac{1}{3}\sigma^2 X^2 T \left(1 - \frac{1}{N}\right)\left(1 - \frac{1}{2N}\right). \tag{12}$$

This strategy minimizes market impact costs; if μ is negligible then it minimizes overall expected costs. E and V have finite limits as the number of trading periods $N \to \infty$.

Minimum variance The other extreme is to sell our entire position in the first time step. We then take

$$x_0 = X, \qquad x_1 = \cdots = x_N = 0$$
$$v_1 = \frac{X}{\tau}, \qquad v_2 = \cdots = v_N = 0,$$

which give

$$E = \epsilon X + \eta \frac{X^2}{T} N, \qquad V = 0. \tag{13}$$

This strategy has the smallest possible variance; by the way that we have discretized time in our model, this minimum is zero. Its expected costs are extremely high if N is large. That is, as we increase the number of time steps, we also increaes the rate of trading within the first step and hence the overall costs.

To summarize, $E(x)$ and $V(x)$ are the expectation and variance of the final cost of trading, assuming that the strategy $x = (x_0, \dots, x_N)$ chosen at the beginning is adhered to throughout liquidation. The particular realization of the cost depends on the realization of the price of the stock at each step of the strategy. As these prices are at least partially random, the entire cost of trading is random.

The distribution of the net cost is exactly Gaussian if the ξ_k are Gaussian; in any case if N is large it is very nearly Gaussian. Since a Gaussian distribution is completely described by its mean and variance, the two quantities $E(x)$ and $V(x)$ contain all the information we need in order to evaluate different liquidation strategies. We shall need to show along the way that it is not possible to improve performance by adapting the strategy x to market movements.

2 The Efficient Frontier and Optimal Trading

Equations (8,9) indicate that the choice of a particular trading strategy determines both the expected cost and the level of uncertainty attached to that strategy. In general, one can decrease the variance of the costs only by increasing the expected level of the same costs, and conversely. For example, if we "dump" our entire block into the market at one time (as we examined above), we are exposed to very little volatility risk, but we incur large transaction costs. On the other hand, if we sell slowly, we minimize transaction costs, but we are exposed to volatility during the liquidation process.

2.1 The definition of the frontier

With this in mind, we examine the proper definition of an "optimal" trading strategy, and compute optimal trading strategies directly from the definition.

First observe that a rational trader will always seek to minimize the expectation of cost for a given level of variance. Thus in analogy with modern portfolio theory we define a trading strategy to be *efficient* or *optimal* if there is no strategy which has lower variance for the same level of expected transaction costs, or, equivalently, no strategy which has a lower level of expected transaction costs for the same level variance.

Concentrating on the latter definition, we see that we may construct efficient strategies by solving the constrained optimization problem

$$\min_{x:V(x)=V_*} E(x). \tag{14}$$

That is, for a given level of variance V_*, we find a strategy that has minimum expected level of transaction costs. This minimum exists uniquely whenever E and V are convex functions; this is the case for the linear models above, and for more general transaction cost models as we consider in the Appendix.

Write $x_*(V_*)$ for a solution to (14). Regardless of our preferred balance of risk and return, every other solution x which has $V(x) = V_*$ has higher expected costs than $x_*(V_*)$ for the same variance, and can never be efficient. Thus, the family of all possible efficient (optimal) strategies is parameterized by the single variable V_*, representing all possible levels of variance in transaction costs. We call this family *the efficient frontier of optimal trading strategies.*

We solve the constrained optimization problem (14) by introducing a Lagrange multiplier λ. For a given value of λ, we solve the unconstrained problem

$$\min_x \big(E(x) + \lambda V(x) \big) \tag{15}$$

and call the solution $x_*(\lambda)$. As λ varies, x_* sweeps out the same one-parameter family as for the constrained optimization problem, and thus traces out the efficient frontier.

The parameter λ has a direct financial interpretation. It is already apparent from (15) that λ is a measure of our risk-intolerance, that is, how much we penalize variance relative to expected cost. In fact, λ is the curvature (second derivative) of a smooth utility function, as we shall make precise in Section 3.

2.2 Explicit construction of optimal strategies

With $E(x)$ and $V(x)$ from (8,9), the combination $U(x) = E(x) + \lambda V(x)$ is a convex quadratic function of the control parameters $x_1, \ldots, x_{N-1}$. Therefore it has a unique global minimum, at a value determined by setting its partial derivatives with respect to the control variables to zero. We readily calculate

$$\frac{\partial U}{\partial x_k} \;=\; 2\eta\tau \left\{ -\frac{\mu}{2\eta} \;+\; \frac{\lambda\sigma^2}{\eta}\, x_k \;-\; \left(1 - \frac{\gamma\tau}{2\eta}\right) \frac{x_{k-1} - 2x_k + x_{k+1}}{\tau^2} \right\}$$

for $k = 1, \ldots, N-1$. Then $\partial U / \partial x_k = 0$ is equivalent to

$$\frac{1}{\tau^2}\Big(x_{k-1} - 2x_k + x_{k+1} \Big) \;=\; \tilde{\kappa}^2 \, (x_k - \bar{x}), \tag{16}$$

with

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

and

$$\bar{x} = \frac{\mu}{2\lambda\sigma^2}. \tag{17}$$

The quantity $\bar{x}$ is the optimal level of stock holding for a time-independent portfolio optimization problem.

The solution to the linear difference equation (16) may be written as $\bar{x}$ plus a combination of the exponentials $\exp(\pm\kappa t_k)$, where κ satisfies

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

The specific solution satisfying the boundary conditions $x_0 = X$ and $x_N = 0$ is

$$x_k = \bar{x} + \frac{\sinh(\kappa(T - t_k))}{\sinh(\kappa T)}(X - \bar{x}) - \frac{\sinh(\kappa t_k)}{\sinh(\kappa T)}\,\bar{x}. \tag{18}$$

for $k = 0, \dots, N$. This solution depends on the input parameters only through the combinations $\bar{x}$ and κ.

We readily calculate the associated velocity

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

for $k = 1, \dots, N$, with $t_{k-\frac{1}{2}} = \left(k - \frac{1}{2}\right)\tau$. We have $v_k > 0$ for each k as long as $0 \le \bar{x} \le X$. Thus the solution decreases *monotonically* from its initial value to zero; an optimal strategy never tells us to buy as part of a liquidation program.

For small τ we have the approximate expression

$$\kappa \sim \tilde{\kappa} + \mathcal{O}(\tau^2) \sim \sqrt{\frac{\lambda\sigma^2}{\eta}} + \mathcal{O}(\tau), \qquad \tau \to 0. \tag{19}$$

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

The behavior expressed by the difference equation (16) and the solutions (18) is relaxation towards the static optimum $\bar{x}$ at a rate κ, subject to the constraints imposed by the initial and final values. The inverse $1/\kappa$ is the time required for the solution to move a factor of e towards or away from $\bar{x}$. We characterize solutions by comparing this time to the total time for liquidation, that is, by examining the ratio $T/(1/\kappa) = \kappa T$ for fixed $\bar{x}$.

If $\kappa T \gg 1$, then either temporary costs are very small, volatility is extremely large, or we are very sensitive to volatility. Our strategy is dominated by the need to reduce volatility risk. We initially sell very rapidly from our initial position down to the optimum level $\bar{x}$; we wait at $\bar{x}$ for most of the time, and near the end sell rapidly to achieve liquidation $x = 0$ at $t = T$.

Conversely, if $\kappa T \ll 1$, then either temporary costs are very large, volatility is small, or we are risk-indifferent. Our strategy is dominated by the need to minimize market impact costs. In the limit $\kappa T \to 0$, we approach the straight line strategy.

We have thus explicitly determined the one-parameter family of solutions $x_*(\lambda)$. It is possible to evaluate $E(\lambda) = E(x_*(\lambda))$ and $V(\lambda) = V(x_*(\lambda))$ in closed forms, since they are just sums of exponentials, but the resulting expressions are very complicated and not very enlightening. It is, however, trivial to evaluate the sums numerically.

Risk-neutral strategy In the risk-neutral limit $\lambda \to 0$ we have

$$x_k = X\left(1 - \frac{t_k}{T}\right) + \frac{1}{4}\frac{\mu}{\eta - \frac{1}{2}\gamma\tau} t_k(T - t_k), \tag{20}$$

which is the linear strategy (10) plus a small quadratic correction proportional to the expected return μ. This is the strategy we call “naïve” below.

2.3 Structure of the frontier

An example of the efficient frontier is shown in Figure 1. The plot was produced using parameters chosen as in Section 4. Each point of the frontier represents a distinct strategy for optimally liquidating the same basket. The tangent line indicates the optimal solution for risk parameter $\lambda = 10^{-6}$. The trajectories corresponding to the indicated points on the frontier are shown in Figure 2.

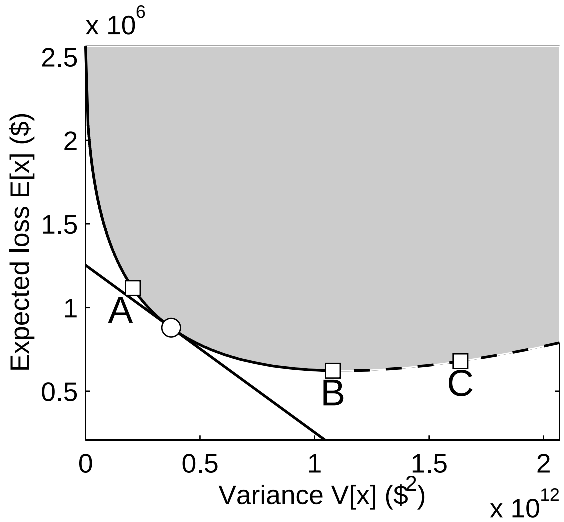

Figure 1: The efficient frontier for parameters as in Table 1. The shaded region is the set of variances and expectations attainable by some time-dependent strategy. The solid curve is the efficient frontier; the dashed curve is strategies that have higher variance for the same expected costs. Point B is the "naïve" strategy, minimizing expected cost without regard to variance. The straight line illustrates selection of a specific optimal strategy for $\lambda = 10^{-6}$. Points A,B,C are strategies illustrated in Figure 2.

Trajectory A has $\lambda = 2 \times 10^{-6}$; it 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.

Trajectory B has $\lambda = 0$. We call this the naïve strategy, since it represents the optimal strategy corresponding to simply minimizing expected transaction costs without regard to variance. For a stock with zero drift and linear transaction costs as defined above, it corresponds to a simple linear reduction of holdings over the liquidation period. Since drift is generally not significant over short trading horizons, the naïve strategy is very close to the linear strategy, as in Figure 2. We demonstrate below that this is *never* an optimal strategy, because one can obtain substantial reductions in variance for a relatively small increase in transaction costs.

Trajectory C has $\lambda = -2 \times 10^{-7}$; it would be chosen only by a trader who likes risk. He postpones selling, thus incurring both higher expected trading costs due to his rapid sales at the end, and higher variance during

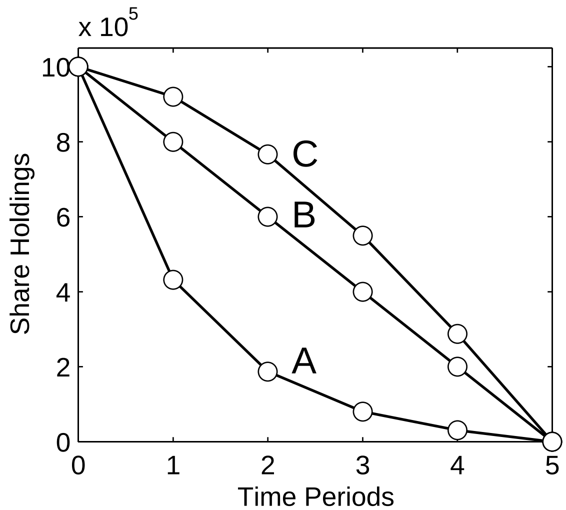

Figure 2: The trajectories corresponding to the points shown in Figure 1. (A) $\lambda = 2 \times 10^{-6}$, (B) $\lambda = 0$, (C) $\lambda = -2 \times 10^{-7}$.

the extended period that he holds the stock.

General conclusions about trading The structure of the efficient frontier lends general insight into the nature of trading large baskets, based on the observation that the curve defining the efficient frontier is a smooth convex function $E(V)$ mapping levels of variance to the corresponding minimum mean transaction cost levels.

Let us denote by (E_0, V_0) the point corresponding to the naïve strategy, that is, the global minimum of E. By smoothness, we have $dE/dV = 0$ there For (E, V) near (E_0, V_0), we have

$$E - E_0 \approx \frac{1}{2}(V - V_0)^2 \left. \frac{d^2E}{dV^2} \right|_{V=V_0},$$

where $d^2E/dV^2\,|_{V_0}$ is positive by convexity. Any reduction in the level of uncertainty of transaction costs comes at the price of a general increase in the level of costs, *but* at the naïve strategy, a first-order decrease in variance can be obtained for only a second-order increase in costs.

Thus, unless you are absolutely risk neutral, it is always advantageous to trade "to the left" of the naïve strategy, or

A risk-averse trader should never use the naïve strategy.

3 The Risk/Reward Tradeoff

We now consider how to choose among the various efficient strategies the one to execute. This amounts to finding a way to convert a dollar of expected transaction cost into a unit of variance and vice-versa. We do this in one of two ways, either by direct analogy with modern portfolio theory employing a utility function, or by a novel approach: value-at-risk.

3.1 Utility function

Suppose we measure utility by a smooth concave function $u(w)$, where w is our total wealth. This function may be characterized by its risk-aversion coefficient $\lambda_u = -u''(w)/u'(w)$. As we transfer our assets from the risky stock into the alternative investment, w remains roughly constant, and thus we may take λ_u to be constant throughout our liquidation period.

For short time horizons and small changes in w, higher derivatives of $u(w)$ may be neglected. Thus choosing an optimal liquidation strategy is equivalent to minimizing the scalar function

$$U_{\text{util}}(x) = \lambda_u V(x) + E(x). \tag{21}$$

The units of λ_u are $\$^{-1}$: we are willing to accept an extra square dollar of variance if it reduces our expected cost by λ_u dollars.

The combination $E + \lambda V$ is precisely the one we used to construct the efficient frontier in Section 2; the parameter λ, introduced artificially in as a Lagrange multiplier, has a precise definition as a measure of our aversion to risk. Thus the construction used above to construct the efficient frontier also gives us the optimal point for any given utility function.

Now we can consider whether we may gain by adapting the strategy to market events as they unfold. Suppose that we stop halfway through the liquidation to reconsider our strategy for the rest of the program. Clearly the variance of the remaining part will be smaller than at the beginning, since some of what might have happend has already happened. And the expectation of cost may be larger or smaller than its starting value, depending on how the price has moved. However, the optimal strategy for the remaining time is just the same as the initial optimal strategy, as may be seen by examining the exact solution.

The exact solution is controlled by the parameters κ and $\bar{x}$. These depend on the properties of the stock random walk and on our risk-aversion factor λ, neither of which changes as the liquidation proceeds.

These parameters do *not* depend either on our initial holdings or on the number of periods remaining. Since the solution to a second-order difference equation is uniquely determined by its two boundary values, the solution for the remaining time is the same as the remaining part of the overall solution.

Note that this argument does not consider whether our estimation of the problem parameters has changed. If a highly unlikely event suddenly occurs, such as the stock price dropping 10%, then we may well wish to raise our estimate of the volatility, or we may wish to increase our risk-aversion λ. But barring such events, our estimations of future motions are not affected by stock movements. This argument also does not apply if price movements are serially correlated, as in Bertsimas and Lo (1998).

3.2 Value at Risk

The concept of value at risk is traditionally used to measure the greatest amount of money (maximum profit and loss) a portfolio will sustain over a given period of time under "normal circumstances," where "normal" is defined by a confidence level.

Given a trading strategy $x = (x_1, \ldots, x_N)$, we define the value at risk of x, $\mathrm{Var}_p(x)$, to be the level of transaction costs incurred by 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 the total cost of trading x.

Under the arithmetic Brownian motion assumption, total costs (market value minus capture) are normally distributed with known mean and variance. Thus the confidence interval is determined by the number of standard deviations λ_v from the mean by the inverse cumulative normal distribution function, and the value-at-risk for the strategy x is given by the formula:

$$\mathrm{Var}_p(x) \;=\; \lambda_v \sqrt{V(x)} \;+\; E(x);. \tag{22}$$

That is, with probability p the trading strategy will not lose more than $\mathrm{Var}_p(x)$ of its market value in trading. Borrowing from the language of Perold (1988), the implementation shortfall of the liquidation will not exceed $\mathrm{Var}_p(x)$ more than a fraction p of the time.

From this point of view or optimal (or efficient) liquidation, a strategy x is efficient if it has the minimum possible value at risk for the confidence level p.

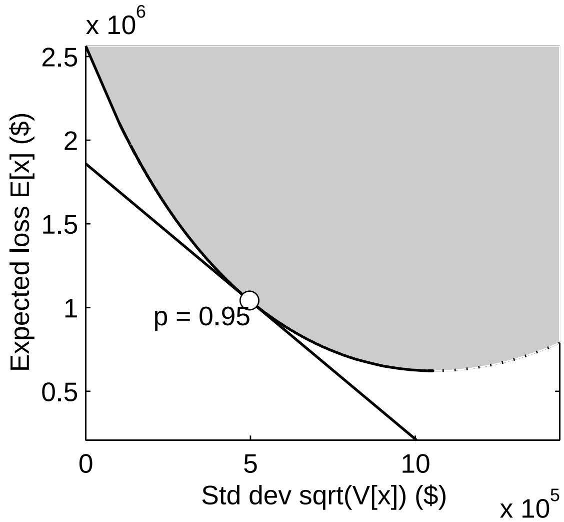

Figure 3: The efficient frontier for parameters as in Table 1, in the plane of $V^{1/2}$ and E. The point of tangency is the optimal value at risk solution for a 95% confidence level, or $\lambda_v = 1.645$.

Note that $\mathrm{Var}_p(x)$ is a complicated nonlinear function of the x_j composing x: we can easily evaluate it for any given trajectory, but finding the minimizing trajectory directly is difficult. But once we have the one-parameter family of solutions which form the efficient frontier, we need only solve a one-dimensional problem to find the optimal solutions for the value-at-risk model, that is, to find the value of λ_u corresponding to a given value of λ_v. Alternatively, we may characterize the solutions by a simple graphical procedure, or we may read off the confidence level corresponding to any particular point on the curve.

Figure 3 shows the same curve as Figure 1, except that the x-axis is the square root of variance rather than variance. In this coordinate system, lines of optimal VaR have constant slope, and for a given value of λ_v, we simply find the tangent to the curve where the slope is λ_v.

In the discrete-time model, the efficient frontier intersects the $V = 0$ axis at a finite height given by (13). In the plane of $\sqrt{V}$ and E, its slope there is finite, and equal to

$$\lambda_v^{\max} = \frac{1}{\sigma\tau^{3/2}}\left(2\eta X - \gamma X\tau + \mu\tau^2 \right).$$

If the confidence level p is large enough that the risk-intolerance parameter λ_v is larger than $\lambda_v^{\max}$, then the optimal strategy is is the minimum-variance strategy of complete liquidation in the first time interval. This behavior is clearly an artifact of our discrete-time model.

Now the question of reevaluation is more complicated and subtle. If we reevaluate our strategy halfway through the liquidation process, we may well decide that the new optimal strategy is not the same as the original optimal one. The reason is that we now hold λ_v constant, and so λ_u necessarily changes. This is a well-known defect of the value-at-risk approach, as recognized by Artzner *et al* (1997a, 1997b). We regard it as an open problem to formulate suitable measures of risk for time-dependent strategies.

4 Numerical Examples

In this section we compute some numerical examples for the purpose of exploring the qualitative properties of the efficient frontier. Throughout the examples we will assume we have a single stock with current market price $S_0 = 50$, and that we initially hold one million shares. Moreover, the stock will have 30% annual volatility, a 10% expected annual return of return, a bid-ask spread of 1/8 and a median daily trading volume of 5 million shares.

With a trading year of 250 days, this gives daily volatility of $0.3/\sqrt{250} = 0.019$ and expected fractional return of $0.1/250 = 4 \times 10^{-4}$. To obtain our absolute parameters σ and μ we must scale by the price, so $\sigma = 0.019 \cdot 50 = 0.95$ and $\mu = (4 \times 10^{-4}) \cdot 50 = 0.02$. Table 1 summarizes this information.

Suppose that we want to liquidate this position in one week, so that $T = 5$ (days). We divide this into daily trades, so τ is one day and $N = 5$.

4.1 Choice of model parameters

We now choose parameters for the temporary cost function (4)

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

We choose $\epsilon = 1/16$, that is, the fixed part of the temporary costs will be one-half the bid-ask spread. For η we will suppose that for each one percent of the daily volume we trade, we incur a price impact equal to the bid-ask spread. For example, trading at a rate of 5% of the daily volume

Initial stock price:	S_0	=	50 \$/share
Initial holdings:	X	=	10^6 share
Liquidation time:	T	=	5 days
Number of time periods:	N	=	5
30% annual volatility:	σ	=	0.95 (\$/share)/day$^{1/2}$
10% annual growth:	μ	=	0.02 (\$/share)/day
Bid-ask spread = 1/8:	ϵ	=	0.0625 \$/share
Daily volume 5 million shares:	γ	=	2.5×10^{-7} \$/share^2
Impact at 1% of market:	η	=	2.5×10^{-6} (\$/share)/(share/day)
Static holdings 11,000 shares:	λ_u	=	10^{-6}/\$
VaR confidence p = 95%:	λ_v	=	1.645

Table 1: Parameter values for our test case.

incurs a one-time cost on each trade of 5/8. Under this assumption we have $\eta = (1/8)/(0.01 \cdot 5 \times 10^6) = 2.5 \times 10^{-6}$.

For the permanent costs, a common rule of thumb is that price effects become significant when we sell 10% of the daily volume. If we suppose that "significant" means that the price depression is one bid-ask spread, and that the effect is linear for smaller and larger trading rates, then we have $\gamma = (1/8)/(0.1 \cdot 5 \times 10^6) = 2.5 \times 10^{-7}$. Recall that this parameter gives a fixed cost independent of path.

In Figure 1 we have chosen $\lambda = \lambda_u = 10^{-6}$. We may interpret this number in terms of the number of shares $\bar{x}$ that we are comfortable holding; this indicates how much diversification we require in our portfolio, if our initial holdings were our total worth. Then (17) gives approximately $\bar{x} = 1,100$ shares, or 0.11% of our initial portfolio. We expect this fraction to be very small since our optimal strategy drives us towards complete liquidation.

For these parameters, we have from (19) that for the optimal strategy, $\kappa \approx 0.6/\text{day}$, so $\kappa T \approx 3$. Since this value is near one in magnitude, the behavior is an interesting intermediate in between the naïve extremes.

For the value-at-risk representation, we assume a 95% desired confidence level, giving $\lambda_v = 1.645$.

	Min. VaR (L-VaR)			Naïve strategy			Static VaR		
	$\sqrt{V}$	E	VaR	$\sqrt{V}$	E	VaR	$\sqrt{V}$	E	VaR
0.25%	0.886	2.249	3.706	1.044	2.122	3.839	2.121	-0.100	3.389
0.5%	0.742	1.365	2.585	1.048	1.122	2.846	2.121	-0.100	3.389
1%	0.497	1.043	1.860	1.058	0.622	2.362	2.121	-0.100	3.389
2%	0.176	0.962	1.250	1.078	0.372	2.145	2.121	-0.100	3.389
$T = 1$	0.440	2.675	3.398	0.465	2.655	3.420	0.949	-0.020	1.540
$T = 2$	0.559	1.475	2.395	0.659	1.397	2.481	1.342	-0.040	2.167
$T = 5$	0.497	1.043	1.860	1.058	0.622	2.362	2.121	-0.100	3.389
$T = 10$	0.040	1.246	1.312	1.580	0.329	2.927	3.000	-0.200	4.735

Table 2: Expected costs, variance of cost, and value-at-risk for time dependent strategies, in millions of dollars. We show them for the strategy which minimizes VaR and for the naïve strategy which minimizes expected costs. For comparison we also show the static VaR values; static VaR depends on the time of holding but is independent of liquidity.

4.2 Effect of parameters

Temporary cost function The most important parameter in determining the path is η, the velocity-dependent part of the temporary cost function. We have selected a certain percentage of the daily volume, at which the temporary cost is equal to the bid-ask spread; above we selected that level to be 1%. Increasing this percentage level has roughly the same effect on the shape of the efficient frontier as does reducing the daily volume. That is, the smaller the value of this percentage, the more the price is sensitive to our trading; and, loosely speaking, the less liquid it is.

In Figure 4 we illustrate the effect of changing this percentage from 0.0025% to 2%, while keeping the desired value-at-risk parameter constant at $p = 95\%$. As this percentage increases, or, loosely speaking, as liquidity increases, the optimal trajectory shifts away from the naïve strategy in the direction of instant liquidation. The expected cost decreases as market impact is reduced.

Table 2 shows the variance, expected costs, and VaR for the minimum-VaR strategy and for the naïve minimum-expectation strategy.

Time to liquidation In Figure 5 we show the effect of changing the time allowed for liquidation between 1 and 10 days. We keep the number of

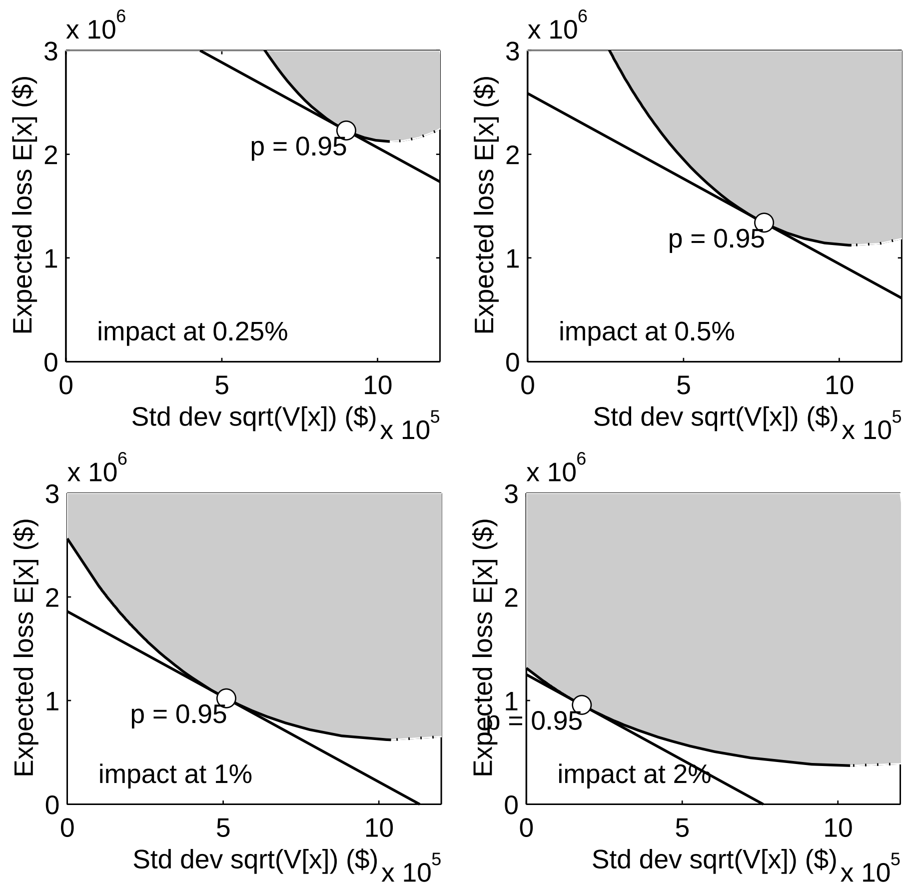

Figure 4: The effect of varying the temporary market impact function, viewed as a proxy for liquidity. From left to right and top to bottom the four frames represent the liquidation of the stock explained in Table 1 with increasing liquidity, controlled by the rate of trading at which temporary market impact occurs. The top left frame represents an illiquid stock: trading at a rate of 1/4% of the median daily volume results in a temporary cost of the bid-ask spread. The bottom right frame represents a more liquid stock: trading at a rate of 2% of the median daily volume results in a temporary cost of the bid-ask spread. Note the two-fold effect of increasing liquidity (Table 2): first, expected transaction costs move down from over 2 million dollars (6%) to just under 1 million dollars (2%). In addition, the "distance" of the 95% confidence level strategy to the naïve strategy increases: the marginal importance of not trading the naïve strategy increases as the liquidity of the position increases.

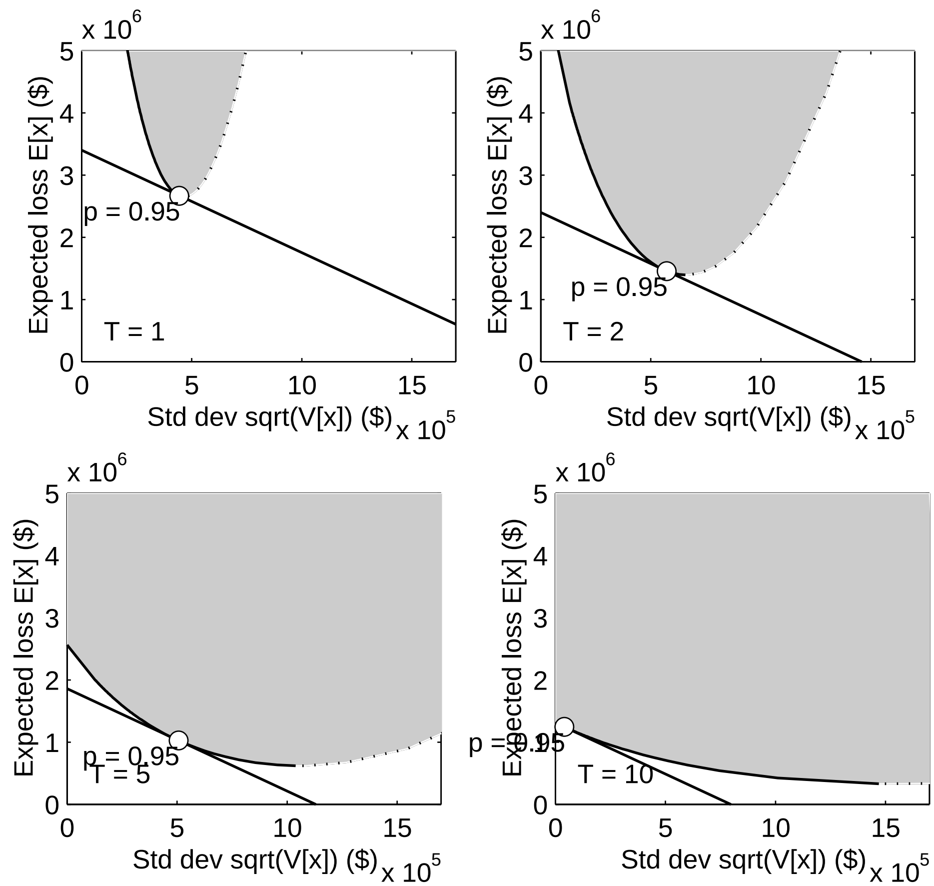

Figure 5: From left to right and top to bottom the four frames represent the liquidation of the stock in Table 1 with time to liquidation increasing (while holding number of trading periods constant). Increasing time to liquidation effectively increases the liquidity of the position, and this is borne out in the simultaneous decreasing level of transaction costs and decreasing variance of costs as time to liquidate increases. As T increases the line of constant slope slides downwards and to the left, indicating that the minimum value of VaR decreases as liquidation time T increases (Table 2).

periods constant at 5, so as the total time is increased, the length of each period increases. As the liquidation time increases, the optimal strategy shifts towards complete liquidation in the first period. The numbers are shown in Table 2.

5 Multiple-Stock Portfolios

With m stocks, our position at each moment is a column vector $x_k = (x_{1k}, \ldots, x_{mk})^{\mathrm{T}}$, where $^{\mathrm{T}}$ denotes transpose. The initial value $x_0 = X = (X_1, \ldots, X_m)^{\mathrm{T}}$, and our rate of selling is the column vector $v_k = (x_{k-1} - x_k)/\tau$. If $x_{jk} < 0$, then stock j is held short at time t_k; if $v_{jk} < 0$ then we are selling stock j between t_{k-1} and t_k.

5.1 Trading model

We assume that the column vector of stock prices S_k follows a multi-dimensional arithmetic Brownian random walk. Its dynamics is again written as in (1), but now $\xi_k = (\xi_{1k}, \ldots, \xi_{rk})^{\mathrm{T}}$ is a vector of r independent Brownian increments, with $r \le m$, σ is an $m \times r$ volatility matrix, and $\mu = (\mu_1, \ldots, \mu_m)^{\mathrm{T}}$ is the vector of expected growth rates. $C = \sigma\sigma^{\mathrm{T}}$ is the $m \times m$ symmetric positive definite variance-covariance matrix.

The geometric model would let μ and C depend on the time step k as

$$\mu_{jk} = S_{jk}\,\bar{\mu}_j, \qquad C_{ijk} = S_{ik}\,S_{jk}\,\bar{C}_{ij},$$

where $\bar{\mu}$ and $\bar{C}$, rather than μ and C, stay constant as S evolves. As for a single stock, we assume the changes are small enough that this effect can be neglected, but we must not forget to insert the factors of S to compute μ and C in terms of fractional rates of return.

The permanent impact $g(v)$ and the temporary impact $h(v)$ are vector functions of a vector. For now, we consider only the linear model

$$g(v) = \Gamma v, \qquad h(v) = \epsilon + H v,$$

where Γ and H are $m \times m$ matrices, and ϵ is an $m \times 1$ column vector. The ij element of Γ and of H represents the price depression on stock i caused by selling stock j at a unit rate. We require that H be positive definite, since if there were a nonzero v with $v^{\mathrm{T}} H v \le 0$, then by selling at rate v we would obtain a net benefit (or at least lose nothing) from instantaneous market impact. We do not assume that H and Γ are symmetric.

The market value of our initial position is $X^{\mathrm{T}}S_0$. The loss in value incurred by a liquidation profile $x_1, \ldots, x_N$ is calculated just as in (7), and we find again, as in (8,9),

$$\begin{aligned}
E[x] &= -\sum_{k=1}^{N} \tau\mu^{\mathrm{T}} x_k + \epsilon^{\mathrm{T}} X + \sum_{k=1}^{N} \tau\, x_k^{\mathrm{T}} \Gamma v_k + \sum_{k=1}^{N} \tau\, v_k^{\mathrm{T}} H v_k \\
&= -\sum_{k=1}^{N} \tau\mu^{\mathrm{T}} x_k + \epsilon^{\mathrm{T}} X + \tfrac{1}{2} X^{\mathrm{T}} \Gamma_S X \qquad (23) \\
&\qquad + \sum_{k=1}^{N} \tau\, v_k^{\mathrm{T}} \Big(H_S - \tfrac{1}{2}\tau\Gamma_S \Big) v_k + \sum_{k=1}^{N} \tau\, x_k^{\mathrm{T}} \Gamma_A v_k \\
V[x] &= \sum_{k=1}^{N} \tau x_k^{\mathrm{T}} C x_k. \qquad (24)
\end{aligned}$$

We use the subscripts $_S$ and $_A$ to denote symmetric and anti-symmetric parts respectively, so $H = H_S + H_A$ and $\Gamma = \Gamma_S + \Gamma_A$ with

$$H_S = \tfrac{1}{2}(H + H^{\mathrm{T}}), \qquad \Gamma_S = \tfrac{1}{2}(\Gamma + \Gamma^{\mathrm{T}}), \qquad \Gamma_A = \tfrac{1}{2}(\Gamma - \Gamma^{\mathrm{T}}).$$

Note that H_S is positive definite as well as symmetric.

Despite the multidimensional complexity of the problem, the set of all outcomes is completely described by these two scalar functionals. The utility function and value at risk objective functions are still given in terms of E and V by (21,22).

5.2 Optimal trajectories

Determination of the optimal trajectory for the portfolio is again a linear problem. We readily find that stationarity of $E + \lambda V$ with respect to variation of x_{jk} gives the multidimensional extension of (16)

$$\frac{x_{k-1} - 2x_k + x_{k+1}}{\tau^2} = \lambda \tilde{H}^{-1} C (x_k - \bar{x}) + \tilde{H}^{-1} \Gamma_A \frac{x_{k-1} - x_{k+1}}{2\tau},$$

for $k = 1, \ldots, N-1$. Here the optimal static portfolio is

$$\bar{x} = \frac{1}{2\lambda} C^{-1} \mu$$

and the symmetric transaction cost matrix is

$$\tilde{H} = H_S - \tfrac{1}{2}\tau\Gamma_S.$$

We shall assume that τ is small enough so that $\tilde{H}$ is positive definite and hence invertible.

Since $\tilde{H}^{-1}C$ is not necessarily symmetric and $\tilde{H}^{-1}\Gamma_A$ is not necessarily antisymmetric, despite the symmetry of $\tilde{H}$, it is convenient to define a new solution variable y by

$$y_k = \tilde{H}^{1/2}(x_k - \bar{x}).$$

We then have

$$\frac{y_{k-1} - 2y_k + y_{k+1}}{\tau^2} = \lambda A\, y_k + B\,\frac{y_{k-1} - y_{k+1}}{2\tau},$$

in which

$$A = \tilde{H}^{-1/2} C \tilde{H}^{-1/2}, \qquad \text{and} \qquad B = \tilde{H}^{-1/2}\Gamma_A \tilde{H}^{-1/2}$$

are symmetric positive definite and antisymmetric, respectively.

5.3 Explicit solution for diagonal model

To write explicit solutions, we make the *diagonal* assumption that trading in each stock affects the price of that stock only and no other prices. This corresponds to taking Γ and H to be diagonal matrices, with

$$\Gamma_{jj} = \gamma_j, \qquad H_{jj} = \eta_j. \tag{25}$$

We require that each $\gamma_j > 0$ and $\eta_j > 0$. With this assumption, the number of coefficients we need in the model is proportional to the number of stocks, and their values can plausibly be estimated from available data. For Γ and H diagonal, $E[x]$ decomposes into a collection of sums over each stock separately, but the covariances still couple the whole system.

In particular, since Γ is now symmetric, we have $\Gamma_A = 0$ and hence $B = 0$; further, $\tilde{H}$ is diagonal with

$$\tilde{H}_{jj} = \eta_j\left(1 - \frac{\gamma_j \tau}{2\eta_j}\right).$$

We require these diagonal elements to be positive, which will be the case if $\tau < \min_j(2\eta_j/\gamma_j)$. Then the inverse square root is trivially computed.

For $\lambda > 0$, λA has a complete set of positive eigenvalues which we denote by $\tilde{\kappa}_1^2, \ldots, \tilde{\kappa}_m^2$, and a complete set of orthonormal eigenvectors which form the columns of an orthogonal matrix U. The solution in the diagonal case is a combination of exponentials $\exp(\pm\kappa_j t)$, with

$$\frac{2}{\tau^2}\Big(\cosh(\kappa_j\tau)-1\Big) \;=\; \tilde{\kappa}_j^2.$$

With $y_k = Uz_k$, we may write

$$z_{jk} \;=\; \frac{\sinh(\kappa_j(T-t_k))}{\sinh(\kappa_j T)}\, z_{j0} \;+\; \frac{\sinh(\kappa_j t_k)}{\sinh(\kappa_j T)}\, z_{jN},$$

in which the column vectors z_0 and z_N are given by

$$z_0 = U^{\mathrm{T}} y_0 = U^{\mathrm{T}}\tilde{H}^{1/2}(X-\bar{x}), \qquad z_N = U^{\mathrm{T}} y_N = -U^{\mathrm{T}}\tilde{H}^{1/2}\bar{x}.$$

Undoing the above changes of variables, we have finally

$$x_k \;=\; \bar{x} \;+\; \tilde{H}^{-1/2} U z_k.$$

With multiple stocks, it is no longer true that each component of our holdings x is monotonically decreasing, even if each component of $\bar{x}$ is between zero and the corresponding component of X.

Risk-neutral strategy We again determine the risk-neutral strategy by taking $\lambda \to 0$ in the above expressions; we find

$$x_k \;=\; X\left(1-\frac{t_k}{T}\right) \;+\; \tfrac{1}{4}\tilde{H}^{-1}\mu\, t_k(T-t_k).$$

In the case $m = 1$, it is easy to see that all of these formulas reduce to those of Section 2.2.

5.4 Example

We now briefly consider an example with only two stocks. For the first stock we take the same parameters as for the example of Section 4. We choose the second stock to be more liquid and less volatile, with a moderate amount of correlation. These parameters are summarized in Table 3. From this market data, we determine the model parameters just as in Section 4.

Our initial holdings are 10 million shares in each stock; we take a time horizon $T = 5$ days and give ourselves $N = 5$ periods. Figure 6

$$
\begin{array}{rl}
\text{Share price} & \begin{pmatrix} \$50 \\ \$100 \end{pmatrix} \\
\text{Daily volume} & \begin{pmatrix} 5 \\ 20 \end{pmatrix} \text{million} \\
\text{Annual variance} & \begin{pmatrix} 30\% & 10\% \\ 10\% & 15\% \end{pmatrix} \\
\text{Annual growth} & \begin{pmatrix} 10\% \\ 10\% \end{pmatrix}
\end{array}
$$

Table 3: Parameters for two-stock example.

shows the efficient frontier in the (V, E)-plane. The three trajectories corresponding to the points A, B, C are shown in Figure 7.

For these example parameters, the trajectory of stock 1 is almost identical to its trajectory in the absence of stock 2 (Section 4). Increasing the correlation of the two stocks increases the interdependence of their trajectories; we expect that relaxing the assumption of diagonal transaction costs would have the same effect. We shall explore the detailed structure of the multidimensional case in future work.

6 Applications

The existence of the efficient frontier along with the ability to compute optimal strategies provides us with several immediate applications of the theory. The first, liquidity adjusted VaR or L-VaR, is a simple definition of a liquidity adjusted value at risk measure that directly generalizes the familiar notion of VaR for portfolios. Performance benchmarks, on the other hand, are concerned with the capture of a trade, and provide client's with a method for measuring the performance of a trader relative to a benchmark that takes their utility function into account.

6.1 Application 1: Liquidity-Adjusted Value at Risk

The value at risk of a portfolio measures the p-th percentile unrealized profit and loss of a portfolio for a particular holding period. See Duffie and Pan (1997) and the reference therein for a detailed discussion of the topic. The main ingredients, however, are

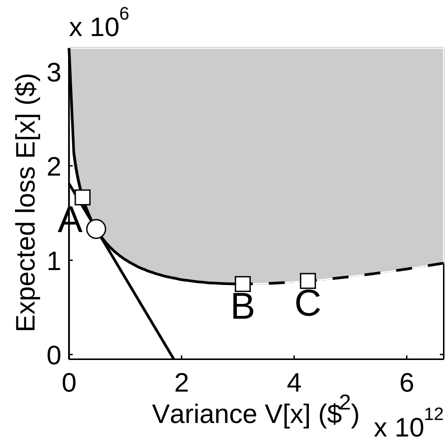

Figure 6: Efficient frontier for the two-stock example. The straight line illustrates the optimal point for $\lambda = 10^{-6}$; the three points A, B, and C illustrate optimal strategies for different values as illustrated in Figure 7.

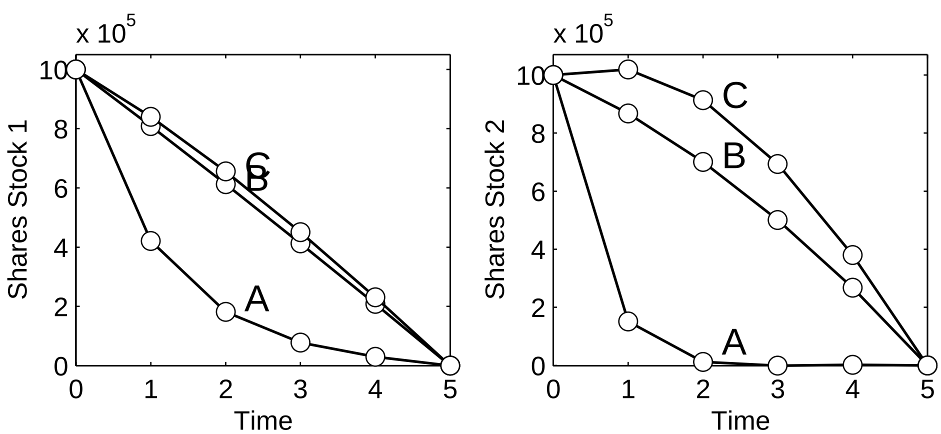

Figure 7: Optimal trajectories for the specific points of Figure 6, for (A) $\lambda = 2 \times 10^{-6}$, (B) the naïve strategy with $\lambda = 0$, (C) $\lambda = -5 \times 10^{-8}$.

- A holding period for the portfolio
- A measure of the mean and variance of the portfolio's value for the holding period,
- A confidence level p.

Given these ingredients, the value at risk of the portfolio is the p-th percentile unrealized profit and loss for the portfolio over the specified holding period. That is, as the portfolio's future value is random with known mean and variance, the profit and losses p-th percent may be computed. This is the most amount of money that will be lost p percent of the time. For example, 99% ten-day value at risk represents the most amount of money that will be lost 99% of the time in holding the given portfolio for 10 days. Put another way, only 1% of the time can we expected to have a ten day loss that exceeds the 99% value at risk.

One criticism frequently lobbied against value at risk is that it does not explicitly account for liquiditity in its computation. Rather, liquidity is often handled in an *ad hoc* manner by adjusting the holding period for the type of instruments being traded. For example, very liquid spot F/X may be treated with a one-day value at risk, while high yield debt with severe liquidity problems may require a full two-week value-at-risk. This definition improves upon no liquidity adjustments but fails to provide satisfactory results for portfolios of mixed composition wherein the range of liquidities is sufficiently great that the various consituents of the portfolio fall into differen "liquidity buckets." Our definition applies equally to all baskets, though our examples below are computed for a single stock.

6.1.1 Definition of L-VaR

The framework for studying liquidation that we have introduced supplies a ready solution to the problem of defining L-VaR for stocks or portfolios. Given a portfolio P, a confidence level p and a holding period T, we define the L-VaR of P for the holding period T and confidence level p to be the VaR of the strategy which is optimal for confidence level of p and liquidation time T.

6.1.2 Examples

To illustrate the computation of L-VaR we refer to Figures 4 and 5, and the computation of the $\mathrm{Var}_p(x)$ discussed in section 3.2. In particular we will compute the L-VaR of each of the eight “95% confidence level optimal” strategies in Figure 4 and 5. The results are tabulated as part of Table 2.

We start by considering the ordinary definition of VaR. The stock in figure 4 is held for 5 days, has a market value of \$50 and we are holding 1 million shares for a total market value of \$50 million. With a 30% annual volatility and 10% expected return this comes to (at the 95% confidence level) a VaR 6.78% or \$3.390M.

The stock in figure 5 has four different VaR’s, reflecting the four different holding periods of 1 day, 2 days, 5 days and 10 days. Table 2 displays the L-VaR numbers for the eight strategies and compares them to the VaR.

For a five day holding period under various liquidities we find that L-VaR decreases as liquidity increases. This makes sense. We also note that for almost all but the most illiquid stock, the L-VaR is less than the ordinary VaR. Why?

The reason is that with ordinary VaR the assumption is that the trader holds the portfolio for the entire holding period. Thus, market moves that occur at any time within the holding period affect the entire portfolio. With L-VaR, we assume liquidation is occuring at all times during the holding period. Thus, the portfolio position is successively diminishing throughout time, and thus market moves affect a successively smaller number of shares of the portfolio.

6.2 Application 2: Performance Benchmarks

In an agency trade, a broker dealer performs a program trade for a client such as a money manager or pension fund. The cost of the trade is usually on a per share basis, and the broker dealer does the trade in the client’s account. Thus the broker dealer assumes no risk for the implementational shortfall. Often the client uses a benchmark to evaluate the performance of the trading deks, such as the volume weighted average price (VWAP) of the shares traded. The problem with this sort of benchmark is that it fails to account for the utility function of the client. Typically, the trade is assigned to the trading desk on a “best effort” basis without a clear definition of the goal of the trade. As we see from the efficient frontier, best effort can produce very different results depending on the nature of the stocks traded and the utility of the trader. A client that wants a great deal of certainty in the level of transaction costs will minimize value at risk to a high degree of confidence. This will select strategies that have higher expected costs (lower capture) but smaller variance. Likewise, clients seeking lower expected costs but willing to assume more risk, will seek to minimize value at risk.

7 Conclusions

We have considered the problem of choosing optimal liquidation strategies for a large position in one or more stocks, by balancing the risk associated with holding stocks longer than necessary against the certain costs of market impact incurred by liquidating too rapidly.

The central feature of our analysis has been the construction of an *efficient frontier* in a two-dimensional plane whose axes are the expectation of total cost and its variance. Regardless of an individual's tolerance for risk, the only strategies which are candidates for being optimal are found in this one-parameter set. For linear impact functions, we give complete analytical expressions for the strategies in this set.

Then considering the details of risk aversion, we have shown how to select an optimal point on this frontier either by classic mean-variance optimization, or by the more modern concept of Value at Risk. These solutions are easily constructed numerically, and easily interpreted graphically by examination of the frontier.

Several conclusions of practical importance follow from this analysis:

- First, we observe that because the set of attainable strategies, and hence the efficient frontier, are generally a *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, a first-order reduction in variance can be obtained at the cost of only a second-order increase in expected cost.

- We 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. We define a quantity called *liquidity-adjusted Value at Risk* (L-VaR); for a given time horizon, this is the minimum VaR of any liquidation strategy.

Finally, let us point out one subtlety of this problem, which suggests directions for future research. The strategies we have considered here are *not adapted* to the random motions of the stock during the liquidation period. That is, at the beginning of liquidation, the trader assesses the risk and cost associated with a given strategy, assuming the strategy were carried to completion without responding to market events.

We have argued (Section 3.1) that this assumption is correct for optimal strategies selected by a classical mean-variance criterion; that is, the trader would never want to change his strategy as long as his estimation of the market parameters has not changed.

However, for strategies selected according to the criterion of Value at Risk, the situation is much more complicated. If the trader reevaluates his strategy half-way through the liquidation, he generally will wish to choose a different strategy for the remaining time. However, viewed from the initial time, the analysis we have proposed is the best that can be done. Value at Risk has several clear limitations as a mathematical tool (Artzner, Delbaen, Eber, and Heath 1997b), and we hope in future work to formulate a more robust notion of risk for time-dependent strategies.

Appendix: Extensions

Continuous time

The continuous-time limit of the above models is easily constructed by taking $\tau \to 0$, so that all the sums become integrals. The optimal strategies are easily found, either by taking the limits of the solutions above, or by applying the calculus of variations to the continuous time problem.

The essential assumption in taking this limit is that the cost functions $g(v)$ and $h(v)$ are well defined in terms of the trading *rate*. It is not obvious that this is the case.

General cost functions

Although above, for the sake of simplicity and concreteness, in the above analysis we have assumed linear permanant and temporary transaction cost functions, our main conclusions are independent of the nature of the cost functions. We now make this explicit by considering the general forms of the model. We shall consider only the continuous-time model, for both a single stock and a portfolio.

For general vector-valued cost functions $g(v)$ and $h(v)$, the variance of our strategy is still given by (24), but the expectation (23) becomes

$$E[x] \;=\; -\sum_{k=0}^{N} \tau\,\mu^{\mathrm{T}} x_k \;+\; \sum_{k=1}^{N} \tau\, x_k^{\mathrm{T}} g(v_k) \;+\; \sum_{k=1}^{N} \tau\, v_k^{\mathrm{T}} h(v_k). \tag{26}$$

The optimality condition becomes

$$\begin{aligned} 2\lambda C x_k \;-\; \mu \;+\; g(v_k) \;+\; \frac{1}{\tau}\Big(\nabla g(v_{k+1})^{\mathrm{T}} x_{k+1} - \nabla g(v_k)^{\mathrm{T}} x_k\Big) \\ +\; \frac{1}{\tau}\Big(h(v_{k+1}) - h(v_k)\Big) \;+\; \frac{1}{\tau}\Big(\nabla h(v_{k+1})^{\mathrm{T}} v_{k+1} - \nabla h(v_k)^{\mathrm{T}} v_k\Big) \;=\; 0. \end{aligned}$$

for $k = 1,\dots,N-1$, where the gradient matrices ∇g and ∇h are the usual Jacobians. In the linear case, we have $\nabla g = \Gamma$ and $\nabla h = H$.

The *diagonal* assumption, expressed by (25) for the linear model, now asserts that the jth cost components $g_j(v)$ and $h_j(v)$ depend only on the j velocity component v_j, so that g and h have the forms

$$g(v) \;=\; \big(g_1(v_1),\dots,g_m(v_m)\big)^{\mathrm{T}}, \qquad h(v) \;=\; \big(h_1(v_1),\dots,h_m(v_m)\big)^{\mathrm{T}}.$$

That is, each is simply a collection of m scalar functions of one variable. Under this assumption, (26) decomposes into an independent sum for each component. Although this assumption is often reasonable, in this section we shall not make it, for the sake of generality.

A reasonable model must choose $g(v)$ and $h(v)$ so that $E(x)$ is a *convex* function, but precise formulation of this condition is somewhat difficult, especially in the discrete-time case. It is easier in the continuous-time limit; loosely speaking, we find that we need the scalar function $v^{\mathrm{T}} h(v)$ to be convex, perhaps non-strictly.

The condition to be imposed on g is a little more subtle: we need each component $g_j(v)$ to be convex when the corresponding component x_j is positive, and concave when x_j is negative. If we assume that optimal liquidation does not tell us to take a short position in a stock in which we are initially long (this is not always true for multiple-stock portfolios), then this becomes a well-defined condition on $g(v)$. For example, in the single-stock case, we may choose $\delta_+, \delta_- \ge 0$ and take

$$g(v) \;=\; \begin{cases} \gamma v + \delta_+ v^2, & v > 0, \\ \gamma v - \delta_- v^2, & v < 0. \end{cases}$$

References

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1997a). A characterization of measures of risk. Talk presented at the Columbia/JAFEE Conference on the Mathematics of Finance, April 6–7 1997.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1997b). Thinking coherently. *Risk 10*(11), 68–71.

Bertsimas, D. and A. W. Lo (1998). Optimal control of liquidation costs. *J. Financial Markets*. To appear.

Chan, L. K. C. and J. Lakonishok (1993). Institutional trades and intraday stock price behavior. *J. Financial Econ. 33*, 173–199.

Chan, L. K. C. and J. Lakonishok (1995). The behavior of stock prices around institutional trades. *J. Finance 50*, 1147–1174.

Duffie, D. and J. Pan (1997). An overview of value at risk. *J. Derivatives* (Spring).

Holthausen, R. W., R. W. Leftwich, and D. Mayers (1987). The effect of large block transactions on security prices: A cross-sectional analysis. *J. Financial Econ. 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. *J. Financial Econ. 26*, 71–95.

Jarrow, R. A. (1992). Market manipulation, bubbles, corners, and short squeezes. *J. Fin. Quant. Anal. 27*, 311–336.

Kraus, A. and H. R. Stoll (1972). Price impacts of block trading on the New York Stock Exchange. *J. Finance 27*, 569–588.

Perold, A. F. (1988). The implementation shortfall: Paper versus reality. *J. Portfolio Management 14*(Spring), 4–9.

Subramanian, A. (1997a). The liquidity discount. Center for Applied Mathematics, Cornell University, Working paper June 1997.

Subramanian, A. (1997b). Optimal liquidation for a large investor. Center for Applied Mathematics, Cornell University, Working paper May 1997.

Vayanos, D. (1997). Strategic trading in a dynamic noisy market. Economic Theory Workshop.