# Optimal Execution of Portfolio Transactions: the Effect of Information Leakage

Adriana M. Criscuolo and Henri Waelbroeck

October 7, 2009

**Abstract**
In a seminal paper, Almgren and Chriss found that to optimize the risk-adjusted liquidation value of a stock it is optimal to trade fastest at the beginning of a trade. This result did not consider the possibility that such an aggressive trade start might leak information about the larger hidden order, leading the market to adjust prices preemptively. In this article, we propose a theoretical framework for the optimal trade-scheduling problem that explicitly accounts for the effect of information leakage on trading cost. The model is based on Information Arbitrage Theory, which posits that information leakage leads to expectations of future imbalances, which are in turn incorporated into the stock's price through arbitrage trades. Using numerical methods, we derive optimal solutions and compare the optimal cost to that of commonly used alternatives. We find that the Almgren Chriss solution is more expensive and causes more permanent impact than an optimal solution that accounts for the effect of information leakage.

PRELIMINARY DRAFT
Comments appreciated

**Table of Contents**

## Introduction

In this paper, we consider the optimal execution of a portfolio transaction. There are many dimensions to this problem that are potentially important to the institutional trader: the real-time fluctuations of liquidity, the news stream, and short-term alpha. In response to these variables, traders must select a trading velocity or aggression level. Also relevant are the use of limit prices and the selection of a trading strategy across an increasingly rich set of options available to institutional traders including block trading systems, dark pools, etc. We limit the scope of the problem by adopting the definition of optimal execution proposed by Almgren and Chriss in their landmark paper (Almgren, R. and Chriss, N., 2000): optimal execution is the execution velocity profile that minimizes the sum of trading cost and risk while completing the trade in a given amount of time, averaging over all other factors.

The optimization of the risk-adjusted cost requires a model for market impact. Market impact has been analyzed for different authors as a function of time and trade size. See for example (Bertismas, D., and Lo, A., 1998), (Almgren, R. and Chriss, N., 2000), (Almgren, R., Thum, C., Hauptmann E., Li, .H, 2005), (Obizhaeva,A., and Wang, J., 2006).

Almgren and Chriss found optimal velocity profiles assuming that market impact could be decomposed into a sum of permanent and temporary components. The permanent impact component of the security price was assumed to depend linearly on the number of shares traded in each temporary interval and the temporary impact component was a linear function of the trading velocity in that interval . It follows from these assumptions that total permanent impact is independent of the trade schedule. Temporary impact is dependent on the trade scheduling and in particular, on the amount of time allowed to execute the trade. These assumptions about linearity are based on theoretical arguments, but are not well supported in actual trading data. Authors acknowledge that *temporary* impact is likely non-linear in the trading velocity (Almgren, 2003), (Almgren, R., Thum, C., Hauptmann E., Li, .H, 2005); the optimization method has subsequently been adjusted for non-linear phenomenological models of temporary impact (Loeb, 1983; Lillo, F., Farmer, J. D., and Mantegna, R.N., 2003). Recently, Almgren has shown that the institution's buying power can be used to minimize the variance on the execution results by adapting the execution speed to price formation, when price formation is driven by the superposition of random Brownian motion and market impact.

In all these studies, the key assumptions are that price is driven by an arithmetic Brownian motion and that the effect of trading on price is stationary, i.e., the increment to permanent impact from one interval to the next is independent of time and the temporary impact is a correction that depends only on the current trading velocity. The assumption of stationary process leads directly to the observation that the permanent impact should be linear on the number of traded shares. This linearity has significant practical implications: it affects the long-term effect on the value of a stock after an institution's trade is complete, so it affects the nominal value of shares still held by a fund well after a trade is done. For example, if a mutual fund is holding 10% of the stock in a large company and wishes to reduce this position to 8%, the permanent impact resulting from the sale of 2% of the company's stock can affect the nominal value of the fund more than the implementation shortfall from executing the sale itself. If we believe that permanent impact *per unit of time* is a linear function of only the trading velocity, we would have to conclude that this loss of nominal value is inevitable and dependent *only on the total executed number of shares*. On the other hand, if permanent impact were a non-stationary process, it would also depend on the trading schedule. An institution could reduce this nominal value loss through an appropriate choice of trading strategy.

In this paper, we will develop a theory for market impact based on the assumption that traders drive prices towards a value that accounts for the information revealed in the order flow. Market participants observe the order flow imbalance caused by a large institutional order, and formulate expectations about the size of the trade and its information content. The fair value of the stock's price, given the observable imbalance, should reflect these expectations. Arbitrage programs provide a mechanism to drive prices towards this level: if the current price does not fully account for the information revealed in the order flow, arbitrage programs can trade on the discrepancy, with the result of moving the price towards fair value. Vice versa, the first trader to realize that an institutional trade is complete has an incentive to take a position on the opposite side, driving the price back towards its fair value with *no* further expectations of future imbalances. This mechanism drives price reversion after a trade is complete. We refer to strategies that detect the presence of informed institutional orders and trade on perceived mis-pricings as information arbitrage strategies. Information Arbitrage Theory aims to explain market impact as the discovery of the fair price given all revealed information about the presence of a hidden order. The first author to analyze the importance of information of future prices was Kyle in (Kyle, 1985). In previous work Farmer et al. (Farmer, J. D., Gerig, A., Lillo F., Waelbroeck H., 2009) showed that information arbitrage theory predicts that implementation shortfall for a trade executed at a constant velocity should be a concave function of trade size with an exponent between 0.5 and 0.6. This result for implementation shortfall is in agreement with most phenomenological models including the Barra model (Torre, 1997). See also, (Chan,L.K.C., and Lakonishok, J., 1993), (Chan, L.K.C., and Lakonishok, J., 1995), (Almgren, R., Thum, C., Hauptmann E., Li, .H, 2005), (Moro, E.,Moyano, L.G.,Vicente, J., Gerig, A., Farmer, J.D.,Vaglica, G.,Lillo,F.,and Mantegna, R.N., 2009). The theory also predicts a concave shape of the permanent impact as a function of the trade size; we will review empirical evidence for this in the concluding section.

Our objective is to use information arbitrage theory to address the optimal trade execution problem. This requires extending the theory to the case where the execution velocity is not constant.

In the following sections, we will propose two models that lead to optimal varying trading velocities. We begin Section 1 by reviewing information arbitrage theory and its prediction of a concave impact function given a constant velocity. In Subsection 1.4, we introduce a velocity profile in the cost function but ignore how observations of the change in velocity can affect expectations of future order flow. We find trading solutions that minimize the total trading cost without risk. In Section 2, we propose a second model that fully accounts for the effect of changing trading velocity on expectations of future order flow. We are able to find optimal solutions for the risk adjusted cost function by using a simulated annealing optimization algorithm. We then consider optimization with respect to the volume-weighted average price objective, with and without risk aversion, and show that trade optimization with respect to the two benchmarks (implementation shortfall and VWAP) is a frustrated problem. We conclude with a discussion of the implications of this paper for optimal trade execution in Section 3.

## 1. Foundations of the theory

This first section is based on the ideas developed in (Farmer, J. D., Gerig, A., Lillo F., Waelbroeck H., 2009)and (Almgren, R. and Chriss, N., 2000). Here we summarize the principal concepts and equations from those references while adding new ideas that we will combine to construct the optimal trajectories of trading.

One of our purposes in this paper is to understand the manner in which information leakage, revealing the presence of a hidden order, causes market impact. The most basic mechanism for that leakage to take place is through the imbalance of the order flow defined as the difference between buy and sell transactions according to the Lee and Ready tick test.

We will assume that a large institutional order is executed using an algorithm that routes slices to the market over time. Because this is not displayed, we will refer to it as a "hidden order." The execution of large institutional orders causes self-correlations in the order flow (Bouchoud, J-P., Farmer, J. D., Lillo, F.). Order flow imbalances can be observed in publicly available market data. This is to some extent inevitable; market structure developments in the past decade and advanced trading algorithms reduce observable imbalances but do not completely eliminate them. An often-mentioned example is the growth of dark pools of liquidity. Dark pools enable institutional traders to intercept some market orders before they reach the displayed markets, resulting in executions at midpoint prices – however, by intercepting market orders on one side and not the other, these systems do indirectly cause an imbalance in the order flow that reaches the displayed markets. Other systems such as Liquid Net and Pipeline offer the opportunity to execute large blocks in off-market crosses but such crossing opportunities are not always available. Efforts that are more recent, aim to minimize the observable effects of algorithmic trading by dynamically switching between different algorithmic trading styles (Stephens & Waelbroeck, 2009).

### 1.1 Detection of hidden orders

To handle the realistic situation where information about the presence of a hidden order is uncertain, we postulate that market participants are able to observe statistically significant aggregate imbalances in the order flow by following a random walk process.

We will consider an anonymous marketplace where the only information available to market participants is the tape and the sign of each transaction as either buyer-initiated (+sign) or seller-initiated (- sign). The choice of a trading trajectory affects how an order becomes detectable; a higher participation rate will cause detection to occur sooner. In this section, we aim to make this more precise by quantifying how much the execution velocity affects the detection process.

For simplicity, we suppose that the institution always sells or buys (never changes sign) the same number of shares $n$ in each transaction. Market participants observe prints on the market data feed; they cannot detect whether an individual print originates from a hidden order or not, but will detect the aggregate effect of the hidden order over trading segments comprising several hidden order transactions. We will call them "***detectable segments***," and the index $i$ will represent the $i^{th}$ *detectable* segment. The natural unit for trading time is transaction time as an integer counter of tape prints. Denote $l_i$ the last print in a detectable $i-$ segment. This is the corresponding transactional step number of the market. We define an expected institutional participation rate of $\pi_i$ during the $i$ –segment; i.e. the institution is expected to trade one of every ${\pi_i}^{-1}$ market transactions during that segment.

We suggested that the market behaves according to a random walk model. In general, to detect an imbalance one must identify a hidden order with a participation rate $\pi$ in a background of random buy and sell transactions. As before, we will consider the units of transaction time which increments by one unit on every tape –reported transaction. For a market observation period of $\tau$ transactions, the expected imbalance will be $\tau\pi,$ with a standard deviation of $\sqrt{\tau}$. When a hidden order is active we expect to find a sustained imbalance over several segments, each with a duration $\tau_i \stackrel{\text{def}}{=} l_i - l_{i-1}$. Therefore, an imbalance is expected to be observable or detectable if $(l_i - l_{i-1})\pi_i \geq \sqrt{l_i - l_{i-1}}$ . Then, the minimum expected transactional market period for detection of the $i^{th}$ segment is:

$$\tau_i = \frac{1}{{\pi_i}^2} . \qquad (1)$$

This means that, during the $i-$segment, the market is expected to report a minimum of $\pi_i^{-2}$ transactions of which ${\pi_i}^{-1}$ correspond to the institution. The interval vector $\{\pi_i\}_{i=1}^N$ represents the information leakage.

## 1.2 Information Arbitrage Theory with constant trading velocity

We assumed the market is influenced by the presence of an institution, which is trading a large hidden order through an order splitting strategy. Other market participants are able to detect evidence of the hidden order in market data but they do not know its size. However, they will observe imbalances in the order flow and formulate expectations about the size of the hidden order and its effect on prices.

We called a "detectable segment" the minimal segment of the timeline that is required to detect the presence of a hidden order at a given statistical significance level.

Let us call $i$ the $i^{th}$ *detectable* segment and $N$ the last one. At the $i-$segment, a key question for the statistical arbitrage trader is to determine whether it will be the last one ($i = N$) or not. We denote $p^+(i+1|i)$ the probability at the $i-$segment that the hidden order is not yet completed, and $p^-(i = N|i) \stackrel{\text{def}}{=} 1 - p^+(i+1|i)$ the complementary probability.

Information Arbitrage Theory is built on four basic hypotheses:

> The expected trading velocity or participation rate $\pi$ is a constant and the hidden order is detected on average every $\pi^{-2}$ market transactions. This fixes the discretization of the time scale.
>
> The second assumption is related to the structure of the distribution functions. We model the activity of hidden orders as a Markovian random process; therefore, it will be completely determined if we give the probability of the first event and the transitional probabilities such that they satisfy the Chapman – Kolmogorov equation (Lax, M., Cai, W., Xu M., 2006).

Let us denote the detection of the hidden order at the $i-$segment as the event $(+, i)$ and the non-detection of the hidden order as the event $(-, i)$. To assure that the $i-$segment symbolizes a segment within some institutional activity was detected by the arbitrageurs, we will define

$$p\,(a, i+1|+, i) = \begin{cases} p^+(i+1|i), & if\ a = + \\ p^-(i = N|i), & if\ a = - \end{cases} \quad \text{and} \quad p\,(a, i+1|-, i) = \begin{cases} 0, & if\ a = + \\ 1, & if\ a = - \end{cases}$$

Market participants, who are able to detect the presence of a hidden order, will formulate expectations about the remaining size. To quantify those expectations, we will define the probability that the hidden order stops at $i$ (or $N = i$) to be $p(N = i) \equiv p_i$ .

Then,

$$p^{+}(i+1|i) \stackrel{\text{def}}{=} \frac{p(N \geq i+1)}{p(N \geq i)} = \frac{\sum_{k=1}^{\infty} p_{i+k}}{\sum_{k=0}^{\infty} p_{i+k}} \quad (2).$$

The Chapman–Kolmogorov condition,

$$p(a, i+2|a'', i) = \sum_{a'} p\left(a, i+2\middle|a', i+1\right)\ p\left(a', i+1|a'', i\right),$$

is satisfied for all $i \geq 1$ for any arbitrary form of the $p_i$ distribution.

The empirical data show (Gopikrishnan, P.,Plerou, V., Gabaix, X., Stanley, H. E., 2000), (Gabaix, X., Gopikrishnan, P.,Plerou, V. and Stanley, H.E., 2006) that the hidden order sizes have a Pareto distribution with parameter $\alpha \in [1.5, 1.6]$,

$$p_i = \frac{1}{\zeta(\alpha+1, 1)\ i^{\alpha+1}}. \quad (3)$$

In this case, formula $(2)$ reduces to

$$p^{+}(i+1|i) = \frac{\zeta(\alpha+1, i+1)}{\zeta(\alpha+1, i)}, \quad (4)$$

where $\zeta(\alpha+1, i)$ is the Hurwitz Zeta function (Wikipedia, the free encyclopedia), (Abramowitz, M., and Stegun, I. A., 1972).

Nevertheless, unless otherwise specified, we will work with a general distribution $p_i$ .The total distribution function of our model will be the product of a Gaussian distribution $G$, which describes a Brownian motion, and  the markovian distribution $p_{i+k}$ of hidden order sizes.

By definition, the expected value of any  observable $X_N$ defined at the end of the trade given at least $i$ detectable segments is

$$E_i\,[X_N] \stackrel{\text{def}}{=} \frac{\sum_{k=0}^{\infty} p_{i+k}\ X_{i+k}}{\sum_{k=0}^{\infty} p_{i+k}}. \quad (5)$$

Here, $N$ is a variable or observable that takes the values $i+k$,  for $k = 0, 1, 2, \dots, \infty$.  The above equation represents our interest of weighting the values of the  variable $X_{i+k}$ as if $i+k$ were the last detectable step.

The Gaussian mean $\langle X_{i+k} \rangle_G$ makes a statistical average of the different possibilities for the values of $X_{i+k}$ at a fixed $i+k$ –segment. We will use the notation

$$E_i[\langle X_N \rangle_G] = \frac{\sum_{k=0}^{\infty} p_{i+k}\ \langle X_{i+k} \rangle_G}{\sum_{k=0}^{\infty} p_{i+k}}, \quad (6)$$

for the total expected value of an observable $X$conditioned on that, we detected a minimum of $i$ segments.

The third assumption is the efficiency condition for the short term: arbitrage opportunities disappear quickly due to the efficiency of the market. This concept translates to the equation:

$$p^{+}(^{i+1}/_{i})\,\langle r_i^{+}\rangle_G + p^{-}(^{i=N}/_{i})\,\langle r_i^{-}\rangle_G = 0, \qquad (7a)$$

where $\langle r_i^{+,-}\rangle_G$ are the corresponding Brownian expected price returns from $i$ to $i+1$. By definition, we write:

$$\langle r_i^{+}\rangle_G = \langle\tilde{S}_{i+1}\rangle_G - \langle\tilde{S}_i\rangle_G, \quad (7b)$$

$$\langle r_i^{-}\rangle_G = \langle S_i\rangle_G - \langle\tilde{S}_i\rangle_G. \quad (7c)$$

$\langle\tilde{S}_i\rangle_G$ is the expected cash price per share paid by the institution in the $i-$segment. If the hidden order is detected and ends at $i=N$, $\langle S_i\rangle_G = \langle S_N\rangle_G$ is the post-reversion price after the $i-$segment. If it does not stop, we will interpret $\langle S_i\rangle_G$ as the price per share only reflecting the permanent impact component at the $i-$segment.

In what follows, we will simplify the notation understanding that $X_i$ always means $\langle X_i\rangle_G$ for all observable $X_i$, which is any linear combination of the security price.

The above formulae $(7)$ and $(2)$ imply that

$$\tilde{S}_i = S_i + \left(\tilde{S}_{i+1} - \tilde{S}_i\right)\sum_{k=1}^{\infty}\frac{p_{i+k}}{p_i}. \quad (8)$$

Therefore, from equations $(6)$ and $(8)$, we get

$$E_i(S_N) = \tilde{S}_i. \quad (9)$$

Note that equation $(9)$ uses a simplified notation and should be read as $E_i(\langle S_N\rangle_G) = \langle\tilde{S}_i\rangle_G$.
Rewriting equation (8) as $(7c)$, we obtain

$$r_i^{-} = -r_i^{+}\sum_{k=1}^{\infty}\frac{p_{i+k}}{p_i}, \quad (8a)$$

The fourth assumption is the breakeven condition. That is, the expected average price of the trade will be equal to the expected post-reversion value of the stock or

$$S_N = \frac{1}{N}\sum_{k=1}^{N}\tilde{S}_k \quad (10),$$

for every trade size N. It was shown in (Farmer, J. D., Gerig, A., Lillo F., Waelbroeck H., 2009) that the efficiency condition $(7)$ implies that the hidden orders must break even on average. Unless we believe that the institutions systematically lose money for certain trade sizes, it follows that the hidden orders should break even on average for every trade size as stated in $(10)$.
As a consequence of breakeven, equation (9) becomes:

$$\tilde{S}_i = E_i\left(\frac{1}{N}\sum_{k=1}^{N}\tilde{S}_k\right), \qquad for\ \ i \geq 1. \quad (11)$$

This expectation can be solved for $\tilde{S}_i$ as shown in (Farmer, J. D., Gerig, A., Lillo F., Waelbroeck H., 2009). Here, we reproduce the derivation for completeness.

Taking the expectation value in the right-hand-side of (11) $(10)$ over hidden order sizes of at least $i$ segments, we have

$E_i\left(\frac{1}{N}\sum_{k=1}^{N}\tilde{S}_k\right) = \sum_{l=0}^{\infty} p_{i+l}\, \frac{1}{i+l}\, \sum_{k=1}^{i+l}\tilde{S}_k\, \frac{1}{\sum_{k=0}^{\infty} p_{i+k}}$.

This implies that

$$\tilde{S}_{i+1} - \tilde{S}_i = \frac{p_i}{\sum_{i+1}}\left(\tilde{S}_i - \frac{1}{i}\sum_{k=1}^{i}\tilde{S}_k\right),\ \ for\ \ i \geq 1, \qquad (12)$$

where we denoted $\sum_i \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} p_{i+k} = p(N \geq i)$ .

By definition of $r_k^+$, in $(7b)$, we can change variables from $\widetilde{S}_i$ to $r_k^+$ using

$$\tilde{S}_i = S_1 + \sum_{k=1}^{i-1} r_k^+.\ \ (13)$$

Inserting (13) in (12),

$$r_{i+1}^+ = \frac{i}{i+1}\frac{p_{i+1}}{p_i}\frac{\sum_i}{\sum_{i+2}} r_i^+. \quad (14)$$

In addition, it is reduced by recursion to,

$$r_i^+ = \frac{1}{i}\frac{p_i}{p_1}\frac{\sum_1\ \sum_2}{\sum_i\ \sum_{i+1}}\ r_1^+, \quad i \geq 1.\ (15)$$

Inserting (8) in (12), we find

$$S_i = \frac{1}{i}\sum_{k=1}^{i}\tilde{S}_k \quad for\ i \geq 1. \qquad (16)$$

Introducing (16) in the sequence $(i+1)S_{i+1} - i\,S_i$ , and using $(7c)$ we get

$$S_{i+1} = S_i - \frac{1}{i} r_{i+1}^- . \quad (17)$$

Applying $(8a)$ and$(15)$, equation (17) expands to

$$S_{i+1} = S_i + \frac{1}{i(i+1)}\,\frac{1}{p_1}\,\frac{\sum_1 \sum_2}{\sum_{i+1}}\ r_1^+\ ,\ \ i \geq 1. \qquad (18)$$

Note that this is not a stationary process: the incremental permanent impact from one segment to the next depends explicitly on $i$. Market expectations depend on the information revealed in the order flow in the first $i$ segments. In this sense, the predictions of the information arbitrage theory will be fundamentally different from those of prior impact models.

### 1.3. Temporary and permanent market impact

The information arbitrage theory explains how price formation is affected by arbitrage activity coming from the predictability of order flow.

The following step is to introduce the concepts of temporary and permanent impact.
As in (Almgren, R. and Chriss, N., 2000) and to find the relations with (Farmer, J. D., Gerig, A., Lillo F., Waelbroeck H., 2009) theory, *permanent impact of the security price* is defined by

$$S_{i+1} - S_i \overset{\text{def}}{=} -\gamma_{i+1} n_{i+1}, \qquad i \geq 0 \qquad (19)$$

where $n_{i+1}$ is the number of shares being sold $(n_{i+1} > 0)$ or bought $(n_{i+1} < 0)$ by the institution in the transactional $i + 1 - \text{segment}$.

Note that equation $(19)$ corresponds to the Brownian expected values of the prices; therefore, the noise term is null.

From (18) and (19), we obtain the *permanent impact function per share*

$$\gamma_{i+1} = \frac{(-1)}{i(i+1)} \frac{1}{p_1} \frac{\sum_1 \sum_2}{\sum_{i+1}} \frac{r_1^+}{n_{i+1}} \ , \ i \geq 1 \qquad (20)$$

Note that $\gamma_{i+1} > 0$ either a sell $(r_1^+ < 0)$ or a buy $(r_1^+ > 0)$.

Given a Pareto distribution of the order sizes as $(3)$, Eq .( 18) results in a prediction of a concave shape for the market impact function that is well approximated in the asymptotic limit by a power law impact function with exponent $\alpha - 1$. For the observed exponent $\alpha = 1.5$, this implies that the market impact function is a square root of time, which for constant trading velocity is the same as square root of executed shares. So information arbitrage theory predicts the correct form of the market impact function, in agreement with phenomenological models of market impact and pre-trade transaction cost analysis models.

Moreover, following (Almgren, R. and Chriss, N., 2000), we define the *temporary impact of the security price* as the difference

$$\tilde{S}_{i+1} - S_i \overset{\text{def}}{=} -h_{i+1}. \quad (21)$$

As in previous results, expression (21) indicates that the temporary impact reflects expectations of future order flow imbalance. Trading activity leads to expectations of future imbalance between supply and demand; market participants price in these expectations and subsequent impact results.

Matching (17) with $(7c)$, we obtain

$$\tilde{S}_{i+1} + r_{i+1}^- = S_i - \frac{1}{i} r_{i+1}^- . \quad (22)$$

Comparing with (21), we arrive at the equation that establishes the relation between the temporary impact and the reversion,

$$h_{i+1} = \frac{(i+1)}{i} r_{i+1}^{-}. \quad (23)$$

This, using $(8a)$ and (15), can be written as

$$h_{i+1} = -\frac{1}{i\, p_1} \frac{\Sigma_1 \Sigma_2}{\Sigma_{i+1}} r_1^{+}, \qquad i \geq 1. \qquad (24)$$

Note that $h_{i+1} > 0$ for a sell and $h_{i+1} < 0$ for a buy.
In contrast with (Almgren, R. and Chriss, N., 2000), temporary impact is a function of $i$. If we use the Pareto distribution $(3)$, it will be as a $\zeta(\alpha+1, i+1)$ Hurwitz Zeta Function. This is a consequence of a theory with assumptions I to IV.

In addition, using (24), we find a relation between permanent and temporary impact functions of the security price as

$$\gamma_{i+1} n_{i+1} = \frac{1}{(i+1)} h_{i+1}. \quad (25)$$

When hidden orders are Pareto distributed, the permanent and temporary impacts behave as a Hurwitz Zeta function. Note that the connection between both is due to the fact that they are related via the efficiency condition equation (7).

### 1.4 The Utility Function

One of the problems we aim to solve in this paper is that of finding the trading trajectory that minimizes the risk-adjusted cost function $U \stackrel{\text{def}}{=} E + \lambda V$. We denote $E$ for the expected implementation shortfall, $V$ for its variance and $\lambda$ is a risk parameter. $U$ is subject to constraints fixing the total number of shares to be executed

$$X = \sum_{i=1}^{N} n_i, \quad (a)$$

and the total transaction time *available for the trade,*

$$T = \sum_{i=1}^{N} \tau_i\,. \quad (b)$$

Here, $X > 0$ for a sell and $X < 0$ for a buy.

To address the optimal execution problem we need to extend the Information Arbitrage Theory to the case where the execution velocity is not constant. Ultimately, we will need an assumption regarding how the market participants form expectations about the information content of an order that is executing with a varying velocity. We will address this in detail in section 2. However, here we will introduce the set $\vec{\pi} = \{\pi_i\}_{i=1}^{N}$ of varying participation rates without altering the hypotheses II to IV.
In eq. $(9)$, we showed that the security price at the $i$ −segment is the expected value of the post-reversion price of the stock. Equations $(13)$ and $(15)$ combined show that the security price does not depend on the velocities:

$$\tilde{S}_i = S_1 + \sum_{k=1}^{i-1} \frac{1}{k} \frac{p_k}{p_1} \frac{\Sigma_1 \ \Sigma_2}{\Sigma_k \ \Sigma_{k+1}} \ r_1^+. \qquad (26)$$

There is no rate of trading in equation $(26)$, because we assumed that the size distribution could be modeled as a distribution over the number of detectable segments. The varying velocity implicitly appears through the equation $N = \pi X$, which indicates that a higher rate $\pi$ implies more detectable intervals $N$ for the same number of total shares $X$. Nevertheless, $\pi_i$ will explicitly appear through the detection process and through the constraints $(a)$ and $(b)$. From section 1.1, we are able to write the expected value of the number of transacted shares at the $i$ −segment

$$n_i = \frac{n}{\pi_i}, \qquad 1 \leq i \leq N, \qquad (27)$$

where $n$ is the number of shares traded in a single transaction and assumed to be a constant parameter (>0 for a sell and <0 for a buy). Additionally, we had deduced equation $(1)$: $\tau_i = \pi_i^{-2}, \ 1 \leq i \leq N.$

In the remainder of this section, we will assume that equation $(26)$ holds for varying velocity.

The *capture of a trajectory* is defined as $\sum_{i=1}^{N} n_i \tilde{S}_i$ and the expected *total cost of trading* (buy or sell) is the difference

$$X S_0 - \sum_{i=1}^{N} n_i \tilde{S}_i \stackrel{\text{def}}{=} E(\vec{\pi}). \quad (28)$$

After a computation, the expected value of the total cost of trading (eqn. (28)) gives

$$E(\vec{\pi}) = \sum_{i=1}^{N} X_i \, n\pi_i^{-1} \ \gamma_i + \sum_{i=1}^{N} n\pi_i^{-1} \ \ h_i \ , \quad (29)$$

with $X_i = X - \sum_{j=1}^{i} \frac{n}{\pi_j}$ the expected number of shares remaining to be traded at the $i$ −segment. The first component of the sum represents the *total* permanent impact cost and the second , the *total* temporary impact cost of the completed trade. In both, we see the dependency on the variable execution speed $\pi_i$. Note that $E(\vec{\pi}) > 0$ always, either a buy or a sell.

As explained previously, we will use the expressions $(20)$ and $(24)$ for the permanent and temporary impact functions of the security price in each segment:

$$\gamma_{i+1} = \frac{(-1)}{i(i+1)} \frac{1}{p_1} \frac{\Sigma_1 \, \Sigma_2}{\Sigma_{i+1}} \ \frac{r_1^+}{n} \pi_{i+1}, \ \ i \geq 1 \qquad (20)$$

$$h_{i+1} = -\frac{1}{i \, p_1} \frac{\Sigma_1 \, \Sigma_2}{\Sigma_{i+1}} r_1^+, \qquad i \geq 1. \qquad (24)$$

Additionally, as in (Almgren, R. and Chriss, N., 2000), we will evaluate the *variance of the cost* $V(\vec{\pi}) = \langle (E(\vec{\pi}) - \langle E(\vec{\pi}) \rangle_G)^2 \rangle_G$ . For that, we will sum the term representing the volatility of the asset

$$\sigma \, \pi_{i+1}^{-1} \ \varsigma_{i+1}, \quad (25)$$

to the equation $(19)$. The $\varsigma_{i+1}$ are random variables with zero Gaussian mean and unit variance and $\sigma$ is a constant with units $[\sigma] = \frac{\$}{share \times \sqrt{transaction}}$.

Therefore, the *variance* of $E(\vec{\pi}\,)$ takes the form

$$V(\vec{\pi}\,) = \sigma^2 \sum_{i=1}^{N} \frac{1}{{\pi_i}^2} {X_i}^2. \qquad (26)$$

We next write the *risk-adjusted cost function*

$$U(\vec{\pi}\,) \stackrel{\text{def}}{=} E(\vec{\pi}) + \lambda V(\vec{\pi}\,), \qquad (27)$$

where $\lambda$ is the *risk parameter* with units $[\lambda] = \$^{-1}$.

Applying the previous expressions, we obtain:

$$U(\vec{\pi}) = n\,\pi_1^{-1}\gamma_1 X - \sum_{i=2}^{N} (i\,n\,\pi_i^{-1} + X - n \sum_{j=1}^{i} \pi_j^{-1}) \frac{1}{i(i-1)} \frac{\Sigma_1\,\Sigma_2}{p_1\,\Sigma_i} r_1^{+} + \lambda\sigma^2 \sum_{i=1}^{N} \pi_i^{-2} \left( X - \sum_{j=1}^{i} \frac{n}{\pi_j} \right)^2 \; (28),$$

with the constraints:

$$X = n \sum_{j=1}^{N} \pi_j^{-1} \; (29)$$

and

$$T \stackrel{\text{def}}{=} l_N - l_0 = \sum_{j=1}^{N} {\pi_j}^{-2} \; (30).$$

Here, $T$ is the institutional trade duration written in the market's transaction time; i.e. the market executes $T$ transactions while the institutional trade is active.

### 1.5 The Optimal Trajectory for λ=0 in the simplified model

In the past chapter, we extended the information arbitrage theory to varying execution speeds based on the assumption that market participants form expectations about future imbalances in order flow. Testing the null hypothesis of a random walk behavior of the market, we determined that an institution caused the imbalance when it was greater than the statistical standard deviation. In contrast, smaller fluctuations were assumed to be statistical noise. From that, we were able to write that the total transaction time is related to the execution schedule as $T = \sum_{j=1}^{N} {\pi_j}^{-2}$, which is the constraint (30).

The expectations about the remaining size of the hidden order were based on a distribution of the number of detectable segments. Permanent and temporary impacts of the security price were independent of the trading velocity at each segment but the *total* cost components resulted dependent through $n_i = \frac{n}{\pi_i}$.

The next step is to obtain an optimal trajectory of trading $(\pi_1^{-1}, \pi_2^{-1}, \dots, \pi_N^{-1})$ given the previous results. Each $\pi_i^{-1}$ represents the number of hidden order executions in the trading $i$ −segment.

We use the Lagrange multipliers method. The constraints (equations $(29)$ and (30)) set the value of the total number of hidden order transactions $X$, and the total transaction time available to execute, $T$. The optimization problem reduces to solving the equations

$$0 = \frac{\partial}{\partial \pi_k^{-1}}[U + \alpha\left(X - n\sum_{j=1}^{N}\pi_j^{-1}\right) + \beta(T - \sum_{j=1}^{N}\pi_j{}^{-2})], (31)$$

where $\alpha$ and $\beta$ are Lagrange multipliers.

Hereafter, we will use the assumption that the number of intervals is Pareto-distributed with exponent $\alpha$ and replace $\sum_i$ by the Hurwitz-Zeta function $\zeta(\alpha+1, i)$, with $\alpha = 1.5$.

Solving for the inverse of the institutional participation rates, we find

$$\pi_k^{-1} = -\left(\frac{n}{2\beta}\right)\left(\frac{r_1^+}{p_1}\zeta[2.5,1]\zeta[2.5,2]A_{k,N} + \alpha\right), \qquad 1 \le k \le N, \ (32)$$

where

$$A_{k,N} = \left((k-1)\zeta[2.5,k]\right)^{-1} - \sum_{i=k}^{N}(i(i-1)\zeta[2.5,i])^{-1}, \qquad 2 \le k \le N, (33)$$

$$A_{1,N} = -\gamma_1 X\left(\frac{r_1^+}{p_1}\zeta[2.5,1]\zeta[2.5,2]\right)^{-1} - \sum_{i=2}^{N}(i(i-1)\zeta[2.5,i])^{-1}, (34).$$

From the constraints, we get

$$\beta_{1,2} = \pm\frac{r_1^+ n\ \zeta[2.5,1]\zeta[2.5,2]}{2p_1\sqrt{NT - \frac{X^2}{n^2}}}\sqrt{-\gamma_1^2 X^2\left(\frac{r_1^+}{p_1}\zeta[2.5,1]\zeta[2.5,2]\right)^{-2} + N\sum_{k=1}^{N}A_{k,N}^2}, (35)$$

$$\alpha_{1,2} = \frac{\left(-2\frac{\beta_{1,2}}{n^2} + \gamma_1\right)X}{N}, \quad (36).$$

Therefore, we are able to obtain two different optimal trajectories $\left(\pi_{(1),1}^{-1}, \pi_{(1),2}^{-1}, \dots, \pi_{(1),N}^{-1}\right)$ and $\left(\pi_{(2),1}^{-1}, \pi_{(2),2}^{-1}, \dots, \pi_{(2),N}^{-1}\right)$ for each set of parameters. One of them minimizes the cost function and the other one maximizes it. To find the optimal solution, we calculate the second derivatives on $\pi_k^{-1}$ of the utility function plus the constraints $(29)$ and $(30)$. After that; we conclude that the trajectory that makes the absolute value of the cost a minimum corresponds to $\beta < 0$; i.e., the pair $(\beta_1, \alpha_1)$.
To obtain a physical solution, we have to fix the parameters of the model such that:

$$N > \frac{X^2}{n^2}\frac{1}{T}, \quad (I)$$

and

$$\pi_i^{-1} > 1, \ \ 1 \le i \le N. \ \ (II)$$

The condition $(I)$ gives us a minimal value for $N$ that we will call $N_{minimun}$:

$$N_{minimun} = \left[\frac{X^2}{n^2}\frac{1}{T}\right] + 1, \quad (37).$$

[ ] denotes the integer part of a real number.
Moreover, the condition $(II)$ gives us a maximum $N_{max}$.

The expression for the cost $(28)$ becomes

$$U(\lambda = 0, N) = n\,\pi_1^{-1}\,\gamma_1\,X - \frac{r_1^+}{p_1}\;\zeta[2.5,1]\zeta[2.5,2]\left\{(X - \pi_1^{-1})\sum_{k=2}^{N} A_{k,N} + \sum_{k=2}^{N} \pi_k^{-1}\,A_{k,N}\right\}.$$

We consider an example with the *sell* parameter values

$T = 10^4\ transactions$ , $X = 10^3\ lots,\ n = 1\ lot^{\dagger},\ (38)$

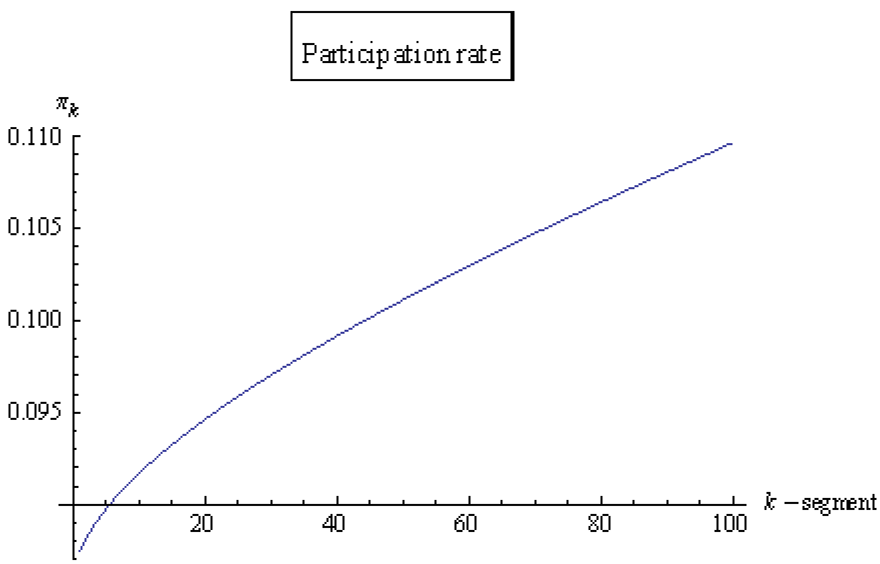
which implies for condition $(I)$ that $N_{minimun} = 101$.

The liquidity and initial returns depend on the stock traded; for the present example we use typical observed impact values for trades of 1% of the average daily volume. Executing at a rate of $\pi = 0.1,$ a detectable interval comprises $\tau = \pi^{-2} = 100$ prints of which the institution takes $\pi^{-1} = 10$. An impact of $10bps$ for the first interval in a $S_0 = 50\$$ security, translate to $10^{-3} \times 50\ \$/share = S_1 - S_0 = -\gamma_1\,n/\pi$ and $r_1^+ = -5 \times 10^{-2}\ \$/share$. We have $p_1 = \frac{1}{\zeta(2.5,1)}$ for $\alpha = 1.5$, then

$$\frac{r_1^+}{p_1} = -6.7 \times 10^{-2}\,\frac{\$}{share}, \qquad \gamma_1 n = 5 \times 10^{-3}\ \$/share, \quad (39).$$

From condition $(II)$ and $(32)$, we obtain $N_{max} = N_{minimun} = 101$.
The conclusion is that, for the set of parameters (38) and (39), we will have an optimal physical solution (i.e., $\pi_i$ a real number and $\pi_i^{-1} > 1,\ \ 1 \le i \le N)$ if $N = 101$.

Below, we show the chart representing the set $(\pi_{(1),1}, \pi_{(1),2}, \dots, \pi_{(1),N})$ of optimal participation rates $(32)$ for $N = 101$.




**Graph 1:** Optimal expected participation rate in each $k$ −segment for an information arbitrage theory without risk (λ= $0$ ). Blue line follows a trajectory that is very close to a linear strategy, around $\pi = 0.1$, as in Almgren-Chriss results for $\lambda = 0$ with exception of the first interval, with a value of $\pi_1 = 0.486$ (not shown in the graph).

Furthermore, we use equation $(28)$ to calculate the value of the minimum cost, resulting in:

† Taking $n = 1\ lot$ to represent a typical partial fill, i.e. approximately 250 shares in the U.S. markets.

$$U(\lambda = 0, N = 101) = 210437\$.$$

The optimal solution is not a realistic trading trajectory for most situations, because it leaves too much of the trade to the end. As we will see in the next section, the excessive back-loading of the execution profile results from the simplifying assumptions made in this first model. The main objective of this section was to demonstrate that the consideration of information leakage can alter the optimal execution profile by favoring slower execution rates at the beginning of a trade.

## 2. A theory for a variable speed of trading

### 2.1 Expectations of future trading, and temporary impact

In the previous section, we developed a model to illustrate how consideration of the information leakage cost changes the optimal trading trajectory $\vec{\pi}$. We found from random walk concepts that it is sensible to consider that segments of $\pi_i^{-1}$ hidden order transactions will be detectable by liquidity providers. Assuming that trade size is proportional to the number of detectable segments and that the number of segments is Pareto-distributed, we showed that it is optimal to trade more slowly at the beginning of a trade. We will call this assumption the Pareto detection hypothesis.

However, in this model the permanent and temporary impacts *of the security price* are independent of the rate of trading, which is contrary to empirical observations (Gomes, C. and Waelbroeck, H., 2008).

In the following, we will discuss the Pareto detection hypothesis and propose a more realistic model where market participants formulate expectations about the likely size of a trade measured in traded shares, $\tilde{\xi}_k \stackrel{\text{def}}{=} \sum_{j=1}^{k} n_j$ , instead of the likely size in number of detectable segments traded by the institution. Although the alternate model is more difficult to solve, we will see that numerical solutions confirm the previous result that it is suboptimal to start a trade at high speed.

Empirical data reveals a Pareto distribution of order sizes when measured in either time duration or number of shares (Bouchoud, J-P., Farmer, J. D., Lillo, F.). Nevertheless, these studies do not address the possibility that trading velocity could be significantly modified during the course of execution.

In reality, a reduction in trading velocity is typically accompanied by partial price reversion. In the context of Information Arbitrage Theory, we explain this reversion as an adjustment in expectations: liquidity providers, who had been following the high velocity signal and observe a change to a lower velocity, will lower their expectations regarding the remaining impact.

With all those ideas in mind, we propose the following set of assumptions:

1. Temporary impact represents expectations of future impact
2. Expectations of future impact are based on the assumption that total trade size [shares] is Pareto-distributed, and on the *number of shares already filled* so far in the hidden order $\tilde{\xi}_k$
3. Expectations of future impact assume that the temporary impact after $\tilde{\xi}_k$ shares is the same as it would be if the trade had been executing at the most recently observed speed $\pi_k$ since the beginning with the same number of traded shares

From these assumptions and drawing from the constant-velocity information arbitrage theory, it follows that temporary impact is

$$\widetilde{S_k} - S_{k-1} = \tilde{\mu}\pi_k^{\beta}\left|\tilde{\xi}_k\right|^{\alpha-1}. \qquad (40a)$$

Here, $\alpha = 1.5$, and $S$ is the trailing average of the market price $\tilde{S}$ so far in the execution. Empirical observations from Pipeline data suggest $\beta = 0.3$, close to theoretical predictions of 0.25 (Bouchoud, J-P., Farmer, J. D., Lillo, F.). The constant $\tilde{\mu} > 0$ for a buy and $\tilde{\mu} < 0$ for a sell.

Information Arbitrage Theory for a variable rate of trading is built on three hypotheses:

1. Hidden orders can be detected by observing aggregated market data over a segment containing $\pi_i^{-2}$ market transactions, where $\pi_i$ is the participation rate in the segment *i*.
2. Temporary impact of the security price is related to price formation as eq. $(40a)$.
3. The liquidity takers break even on average for every trade size as:

$$S_{k-1} = \frac{\sum_{i=1}^{k-1} n_i \, \widetilde{S_i}}{\sum_{i=1}^{k-1} n_i}, \qquad N \geq k \geq 2\,, \quad (41)$$

Note that we are not forcing any efficiency condition to be valid. Instead, we adopt equation $(40a)$ that, as we showed in the previous section, follows from the efficiency condition for trades that are executed at a constant rate. We will come back to this point in section 2.3. The next step is to determine the cost of trading and the optimal trading trajectories.

## 2.2 The Optimal Trajectory for λ=0

Using the random walk model to determine the size of detectable intervals (see 1.1), the equation $(40a)$ can also be written as

$$\widetilde{S_k} = S_{k-1} + \mu \pi_k^{\beta} \left( \sum_{i=1}^{k} \frac{1}{\pi_i} \right)^{\alpha-1} . \quad (40b)$$

Here, $\mu \stackrel{\text{def}}{=} \tilde{\mu} |n|^{\alpha-1}$ and units $[\mu] = \frac{\$}{share}$. Combining (40b) with the breakeven condition $(41)$, we derive the expression for the expected price at $k$, as a function of the participation rate schedule $\pi_i, 1 \leq i \leq k$:

$$\widetilde{S_k} = S_0 + \mu \left\{ \pi_1^{\beta-\alpha+1} + \pi_k^{\beta} \left( \sum_{i=1}^{k} \pi_i^{-1} \right)^{\alpha-1} + \sum_{i=2}^{k-1} \pi_i^{\beta-1} \left( \sum_{j=1}^{i} \pi_j^{-1} \right)^{\alpha-2} \right\}, 2 \leq k \leq N, \quad (42a)$$

$$\widetilde{S_1} = S_0 + \mu \, \pi_1^{\beta-\alpha+1}. \quad (42b)$$

The expected total cost of the trade is

$$U(\lambda = 0) = E \stackrel{\text{def}}{=} X \, S_0 - \sum_{i=1}^{N} n_i \, \tilde{S}_i \,, \qquad (43)$$

where $X$ is the total number of shares traded until the last detectable interval $N$ and $n_i = n\pi_i^{-1}$ is the number of shares traded in the $i$ –segment.

After a calculation, using equations $(42)$, we arrive at:

$$U(\lambda = 0) = |\mu X| \sum_{k=1}^{N} \pi_k^{\beta-1} \left( \sum_{i=1}^{k} \pi_i^{-1} \right)^{\alpha-2} , \quad (44)$$

with the constraints

$$X = n\sum_{i=1}^{N}\pi_i^{-1}, (45)$$

$$T = \sum_{i=1}^{N}\pi_i^{-2}. \quad (46)$$

Here, $X$ is the total number of shares traded and $T$ is the total duration of the market trade, in units of transaction time. Note that $U$ is positive over the relevant domain $\pi_i > 0$.

We now make a change of variables:

$$\xi_k \stackrel{\text{def}}{=} \tilde{\xi}_k\, n^{-1} = \sum_{j=1}^{k}\pi_j^{-1}, \quad N \geq k \geq 1, \qquad (47)$$

Equation $(47)$ represents the total number of transactions $\xi_k$ traded until the $k$ detectable interval and, it is always a positive real number.

The constraint $(45)$ becomes the final condition

$$\xi_N = X/n, \qquad (48)$$

and $(46)$ is

$$\sum_{k=1}^{N}(\xi_k - \xi_{k-1})^2 = T. \quad (49)$$

We add the initial condition:

$$\xi_0 = 0. \qquad (50)$$

After the change of variables, expression $(44)$ is:

$$U(\lambda = 0) = |\mu X|\sum_{k=1}^{N}(\xi_k - \xi_{k-1})^{1-\beta}(\xi_k)^{\alpha-2} \quad . \qquad (51)$$

As a result of minimizing eq.$(51)$ together with the constraint (49) represented by the Lagrange multiplier $\gamma$, we get the equation for the optimum

$$0 = |\mu X|\,(\alpha-2)\xi_j^{\alpha-3}\left(\xi_j - \xi_{j-1}\right)^{1-\beta} + |\mu X|(1-\beta)\left\{\xi_j^{\alpha-2}\left(\xi_j - \xi_{j-1}\right)^{-\beta} - \xi_{j+1}^{\alpha-2}\left(\xi_{j+1} - \xi_j\right)^{-\beta}\right\} + 2\gamma\left(2\xi_j - \xi_{j-1} - \xi_{j+1}\right). \qquad (52)$$

Next, we will present solutions of eq.$(52)$ for different values of the parameters $\alpha$ and $\beta$.

1. Solutions for $\beta = 1$

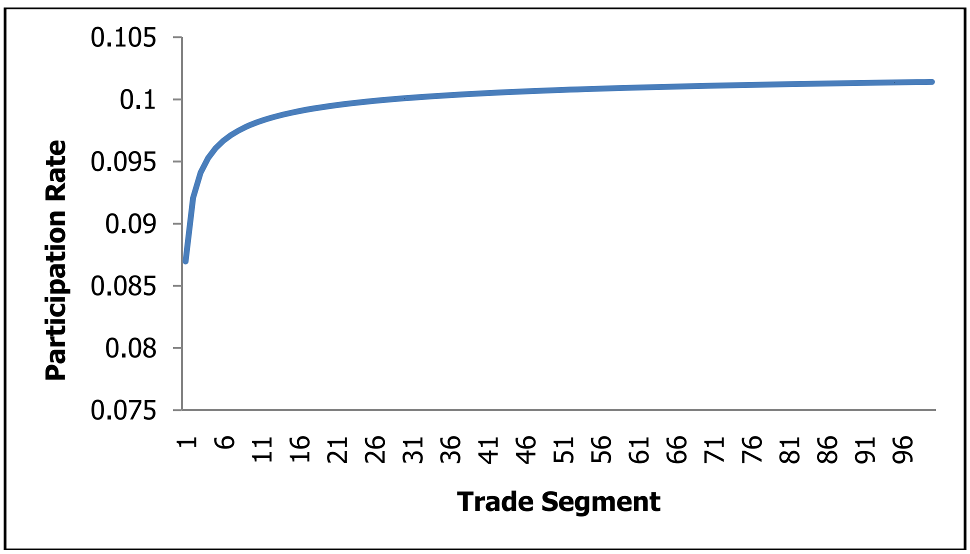
This is the case where temporary impact has a lineal dependency on the trading rate, in equation $(40a)$.

1.1. For $\alpha = 1.5$, we obtain

$$0 = -\frac{|\mu X|}{2} + 2\gamma\, \xi_j^{3/2} \left(2\xi_j - \xi_{j-1} - \xi_{j+1}\right). \qquad (53)$$

We solve equation (53) by recurrence using the initial conditions $\xi_0 = 0$ and $\xi_1 = 11.5$ . The parameters are $|\mu| = 1(\frac{\$}{share})$, $|X| = 1000\, lots^{‡}$, $N = 100$, $T = 10000$. For $\gamma/|n| = 10(\frac{\$}{share})$, the solution satisfies the constraints $(48)$ and $(49)$. Note that $|\mu|$ sets the scale for measuring absolute cost, we have chosen $|\mu| = 1$ arbitrarily here, as we are interested primarily in the shape of the execution profile. We will examine the absolute scale in subsection 2.4 where we consider an example of an actual trade with realistic values for the parameters.

Below, we show the graph of the participation rate $\pi_j = \left(\xi_j - \xi_{j-1}\right)^{-1}$ versus $j$, $1 \le j \le 100$.




**Graph 2:** An optimal participation rate profile $\pi_j = \left(\xi_j - \xi_{j-1}\right)^{-1}$ is shown as a function of the trade segment $j$, $1 \le j \le 100$, for a theory without risk ($\lambda = 0$). This solution illustrates that trading cost can be reduced by beginning an execution at a lower participation rate, in order to minimize information leakage. The cost is $\frac{U(\lambda=0)}{|\mu X|} = 5.773$; as we will see below this is a local minimum of the cost function.

It is worth noting that the iterative equation (53) is a non-linear recursion, therefore there is no guarantee that the solution we have found analytically using the Lagrange parameters method should be the global optimum. Also, by writing the constraints as in (45, 46) we are restricting the space of possible solutions to those that completely fill each detectable segment, possibly ignoring solutions that end with the final segment only partway filled (i.e., all $X$ shares are filled but the last segment of the

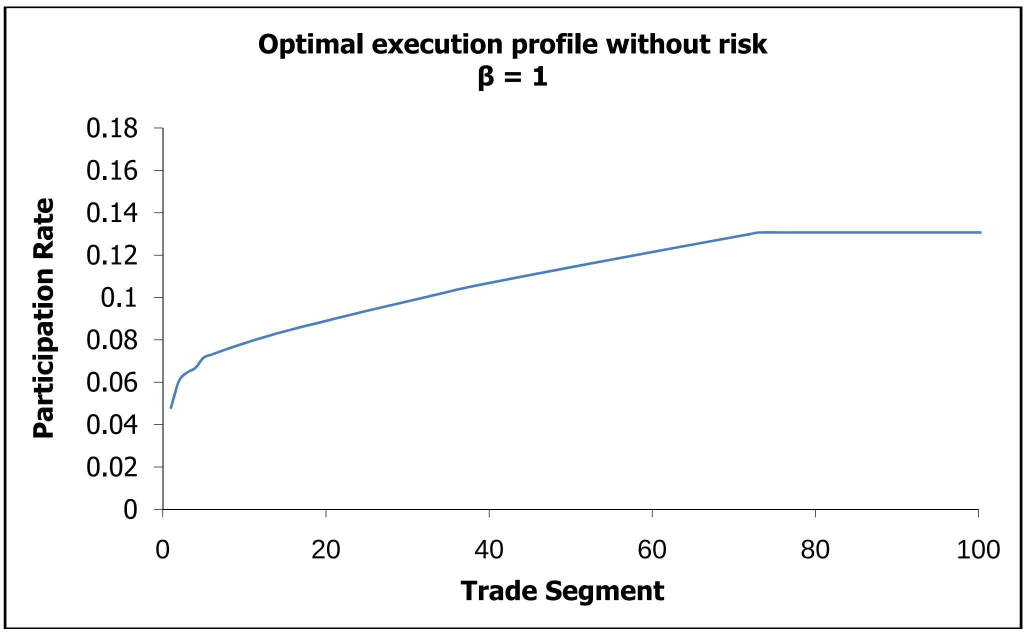
‡ $1 lot = n\ shares$

trade does not contain enough hidden order activity to enable detection of the hidden order). To find the global optimum we address both of these issues next.

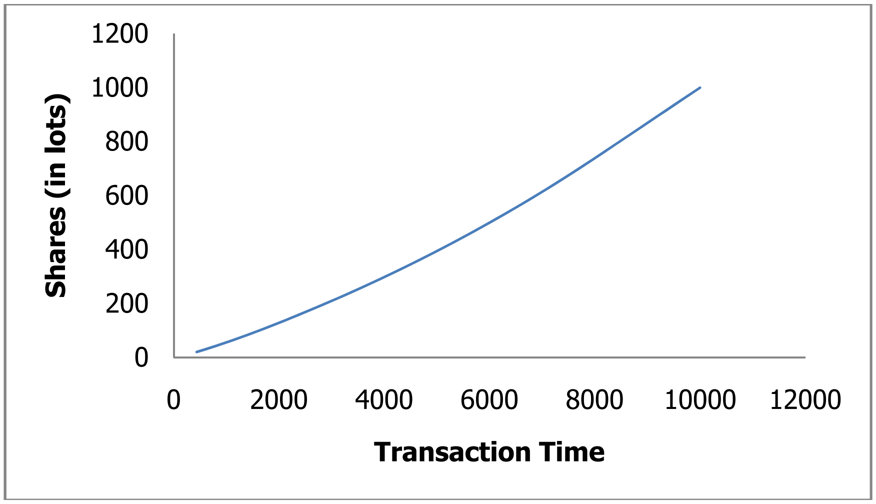
We restrict the search of optimal trading solutions to trading trajectories s of the form $\{\pi_1, \pi_2, \dots, \pi_k, \pi^*, \pi^*, \dots, \pi^*\}$ , $k$ being an integer, where the first $k$ trading velocities are subject to optimization and the remaining velocities are set to the single fixed value that satisfies the constraints, namely

$$\pi_{k+1} = \pi_{k+2} = \cdots = \pi_N = \pi^* = \frac{X - \sum_{i=1}^{k} 1/\pi_i}{T - \sum_{i=1}^{k} 1/\pi_i^2} .$$

In order to find the global optimum, we performed the optimization of $(51)$ using the simulated annealing algorithm (Kirpatrick,S.;C.D. Gelatt, M.P.Vecchi., 1983).This method produced a lower cost solution than the first one represented in Graph 4. Here, we give the graphs of the expected optimal participation rates and the number of optimal expected shares traded by the institution




**Graph 3:** Expected optimal participation rate $\pi_j$ in function of the trade segment, $1 \le j \le N = 106$ , for a theory without risk, found using the simulated annealing optimization method. As above, the lower initial speed leads to lower trading cost. The optimal cost is $\frac{U(\lambda=0)}{|\mu X|} = 5.392$ (global optimum).




**Graph 4:** Expected optimal number of transactions $\xi_j$ executed by the institution in function of the transaction time $t_j = \sum_{i=1}^{j} \pi_i^{-2}$, for a theory without risk.

*Asymptotic behavior of the solution in the continuum limit*

In the continuum limit, eq. (53) becomes

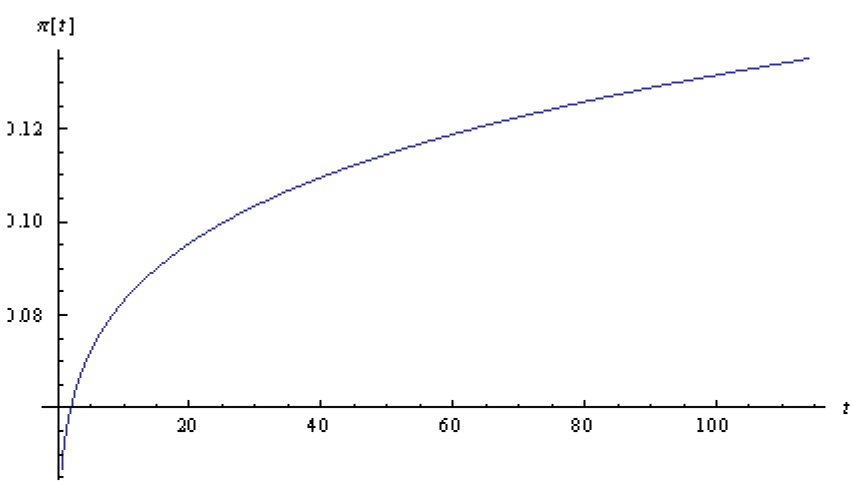
$$0 = \frac{|\mu X|}{2} + 2\gamma\, \xi(t)^{\frac{3}{2}}\, \frac{d^2\xi(t)}{d^2t}, \quad (54)$$

which has the solution

$$\xi(t) = \left(25\frac{|\mu X|}{16\gamma}\right)^{\frac{2}{5}} t^{\frac{4}{5}}\ , \qquad (55).$$

This solution indicates that the optimal total number of executed transactions $\xi(t)$ increases with the time $t$ as a power law with exponent $0.8$.

Below, we plot the participation rate $\pi(t) = \left(\frac{d\xi(t)}{dt}\right)^{-1} = \frac{5}{4}\left(\frac{16}{15\,T}\right)^{\left(\frac{4}{5}\right)} \xi_N^{\frac{3}{5}}\ t^{\frac{1}{5}}$. The value of the parameters $\gamma$ and the final time $t_N$ were obtained through the constraints $\xi_N = \int_0^{t_N} \pi(t)^{-1}\ dt$ and $T = \int_0^{t_N} \pi(t)^{-2}\ dt$. For $|X| = 1000 lots$ and $T = 10000$, is $t_N = 114.286$. This solution is similar to that found with simulated annealing.




**Graph 5:** Participation rate in function of time $t,\ 0 \le t \le t_N$, in the continuum limit for a theory without risk. We see the resemblance with Graph 3**.** The solution is for $|X| = 1000\ lots,\ T = 10000$.

1.2. For $\alpha = 1$ the recurrence equation is

$$0 = -|\mu X| + 2\gamma\, \xi_j^2\, (2\xi_j - \xi_{j-1} - \xi_{j+1}). \qquad (56)$$

In the continuum limit, equation (56) becomes

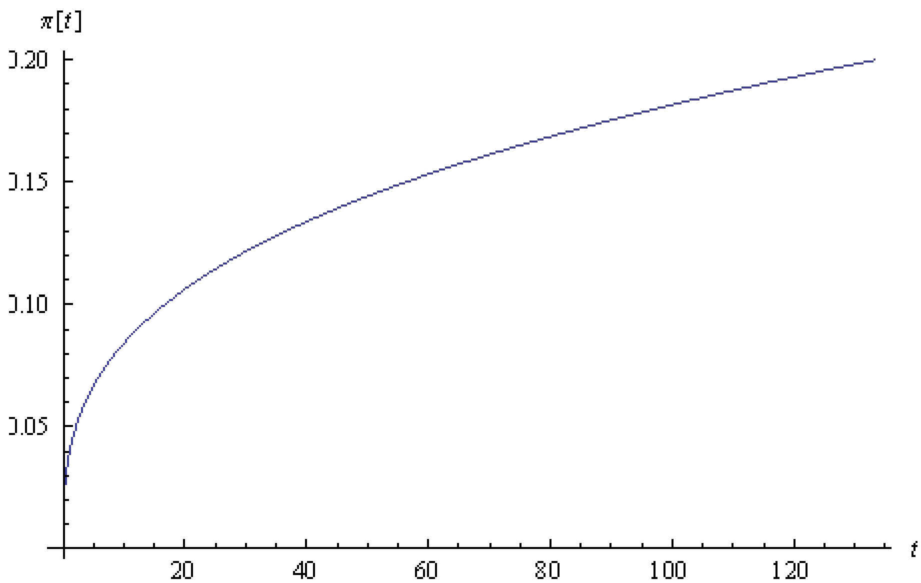
$$0 = |\mu X| + 2\gamma\xi^2\, \frac{d^2\xi(t)}{d^2t}, \qquad (57)$$

with the solution

$$\xi(t) = \left(\frac{9|\mu X|}{4\gamma}\right)^{\frac{1}{3}} t^{\frac{2}{3}}\,. \qquad (58)$$

From the constraints, we get for $|\mu| = 1$: $\gamma/|n| = 4\left(\frac{\xi_N}{T}\right)^2$ and $t_N = \frac{4}{3}\frac{\xi_N{}^2}{T}$.

Below, we represent the participation rate $\pi(t) = \frac{3}{2}\left(\frac{16\,\xi_N}{9\,T^2}\right)^{\frac{1}{3}} t^{\frac{1}{3}}$, for the parameters $|X| = 1000 lots, T = 10000$.




**Graph 6:** This is the optimal participation rate as a function of the time $t, 0 \le t \le t_N = 400/3$, for $\lambda = 0$, in the continuum limit.

2. Solutions for $\beta = 0$.

   This is the case where temporary impact is a function only of the total number of shares acquired and independent of the trading rate, in equation $(40a)$.

   2.1. For $\alpha = 1.5$, we find a polynomial equation that does not lend itself well to numerical solution but can be solved in the continuum limit

$$\frac{d^2\xi(t)}{d^2t} = 0. \quad (59)$$

The optimal linear solution is:

$$\xi(t) = C\, t, \quad (60)$$

with $C$ a constant value.

2.2. For $\alpha = 2$ and $\beta = 0\ or\ 1$, we obtain the same equation

$$\left(2\xi_j - \xi_{j-1} - \xi_{j+1}\right) = 0, \qquad (61)$$

with the optimal linear solution

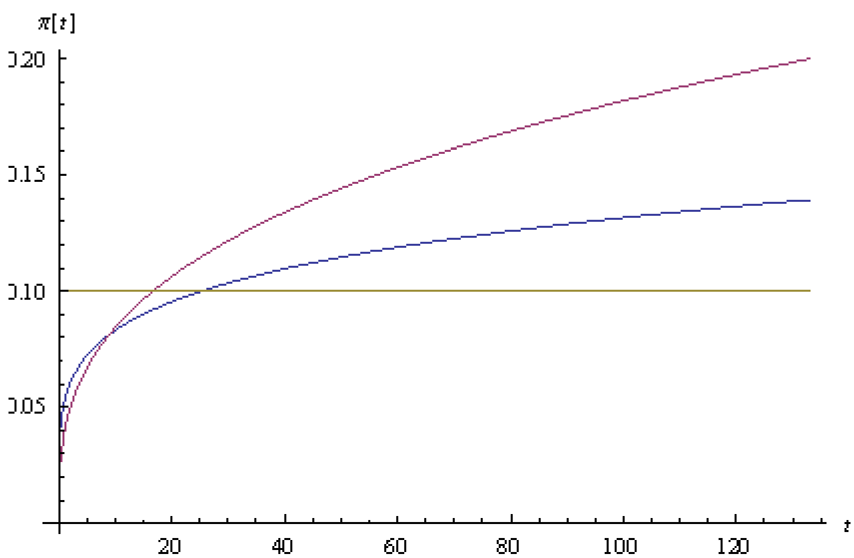
$$\xi_j = 10\,j, \qquad (62)$$

for $|X| = 1000\ lots, N = 100, T = 10000$.

2.3. For $\alpha = 1$ and $\beta = 0$ , we obtain again- in the continuum limit- an equation and a solution of the same type as $(59)$and (60).

As equations $(60)$ and (62) show, *if* we assumed that impact cost were independent of the trading rate *and* ignored risk, it would be optimal to trade at a constant participation rate for the entire trade duration – this result agrees with (Almgren, R. and Chriss, N., 2000). However, given that impact costs tend to increase with trading speed we find in Eqn. (55) and $(58)$ that it is optimal to trade with a lower participation rate at the beginning of a trade. We will consider how the consideration of risk affects this result in the following subsection 2.4.

Below, we show the participation rate plots for the different values of $\alpha$ and $\beta$.




**Graph 7:** Optimal participation rates for a theory without risk. The red line represents $\alpha = 1, \beta = 1$. The blue line is $\alpha = 1.5, \beta = 1$. The green line represents the cases $\beta = 0, \alpha = 1, 1.5, 2$ and $\beta = 1, \alpha = 2$. The constant velocity trajectory would be optimal (without risk) if $\beta = 0$, where temporary impact is a function of shares filled but not of the trading velocity - but this is not what is observed in actual trading. For real-world parameter values we find that lower costs can be achieved by trading more slowly at the beginning.

**2.3 Efficiency Condition**

In the first Section of this paper we developed a theory for constant participation rate and postulating three other assumptions: the efficiency condition, Pareto-distributed order sizes, and the breakeven condition.

In developing Section (2) for a variable participation rate, we assumed that the rationality condition (41) was valid. This is equivalent to the equation (10) or breakeven condition for constant participation rate. Instead of proposing a transitional probability, we proposed a temporary impact function that it was written in equation $(40)$.

Our question is now: if we want the efficiency condition (7) to be satisfied for a variable participation rate, what is the corresponding transitional probability $p^{+}(k \,/\, k-1)$? The answer is

$$p^{+}(k/k-1) = \frac{\tilde{S}_{k-1} - S_{k-1}}{\tilde{S}_{k} - S_{k-1}}.$$

Using equations $(40)$, $(41)$ and $(42)$, we obtain

$$p^{+}(k \,/\, k-1) = \left(\frac{\pi_{k-1}}{\pi_k}\right)^{\beta} \left(\frac{\xi_{k-1}}{\xi_k}\right)^{\alpha-2} \frac{\xi_{k-2}}{\xi_k}\,, \quad k \geq 3. \quad (63)$$

In Section 2.1., we proposed the formula $(40a)$ for the permanent impact of the security price on basis of the information arbitrage theory for constant participation rate and on the Pareto-distributed order sizes in the asymptotic limit $k \gg 1$. For $\pi = \pi_i, 1 \leq i \leq N$, equations $(40\,)$ and $(63)$ lead to:

$$\widetilde{S_k} = S_{k-1} + \mu\, \pi^{\beta-\alpha+1}\, k^{\alpha-1}, \qquad (64)$$

$$p^{+}(k/k-1) = \left(\frac{k-1}{k}\right)^{\alpha-2} \frac{k-2}{k}\,, \qquad k \geq 3. \quad (65)$$

We should verify if the above equations are consistent with the information arbitrage theory for constant participation rate in the limit $k \gg 1$.

Comparing equation $(64)$ with equation $(24)$ and $(65)$ with $(4)$ for the constant participation rate theory, we conclude that equations match if

$$k \gg 1,$$

$$\mu \propto \frac{\Sigma_1\, \Sigma_2}{p_1}, \quad (66a)$$

$$\pi^{\beta-\alpha+1} \propto |r_1^{+}|, \quad (66b).$$

*Proof*: From equations $(24)$ and $(64)$, the term $\frac{1}{(k-1)\sum_{i=0}^{\infty} p_{k+i}}$ should be proportional to $k^{\alpha-1}$, then equations match iff relations $(66)$ are valid and

$$\frac{1}{(k-1)k^{\alpha-1}} \propto p(N \geq k) = \sum_{i=0}^{\infty} p_{k+i}\,, \qquad k > 1. \quad (67)$$

Subtracting $p(N \geq k+1)$ from $p(N \geq k)$, we obtain:

$$p_k \propto \frac{1}{(k-1)k^{\alpha-1}} - \frac{1}{k\,(k+1)^{\alpha-1}}, \quad k > 1, \quad (68).$$

Equation $(68)$ turns out to be $\frac{1}{k^{\alpha}(1-k^{-1})} - \frac{1}{k^{\alpha}(1+k^{-1})^{\alpha-1}} \xrightarrow{k\gg1} k^{-\alpha}(1+k^{-1}) - k^{-\alpha}(1-(\alpha-1)k^{-1}) = \alpha\, k^{-(\alpha+1)}$, which is the Pareto distribution $(3)$.

### 2.4 Risk-adjusted Optimization

In the last section of this paper, we considered the implications of information arbitrage theory for variable speed of trading without risk aversion ($\boldsymbol{\lambda = 0}$). In what follows, we will find optimal trading trajectories for a theory with varying participation rate and arbitrary risk aversion. That requires us to minimize the total risk-adjusted cost function:

$$U = |\mu X| \sum_{k=1}^{N} \pi_k^{\beta-1} \left( \sum_{i=1}^{k} \pi_i^{-1} \right)^{\alpha-2} + \lambda\sigma^2 n^2 \sum_{k=1}^{N} \pi_k^{-2} \left( \xi_N - \sum_{j=1}^{k} \pi_j^{-1} \right)^2. \quad (69)$$

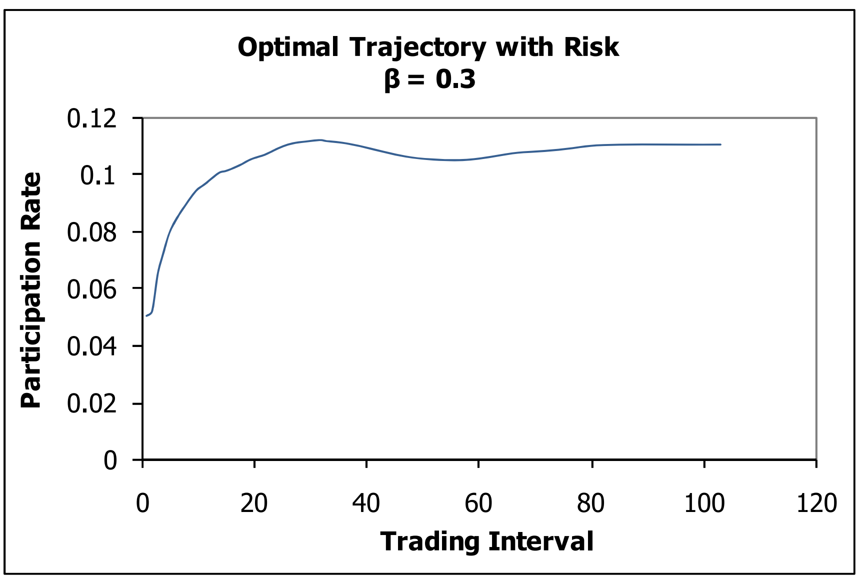
The constraints are the total trading time $T$:

$$T = \sum_{i=1}^{N} \pi_i^{-2}, \quad (30)$$

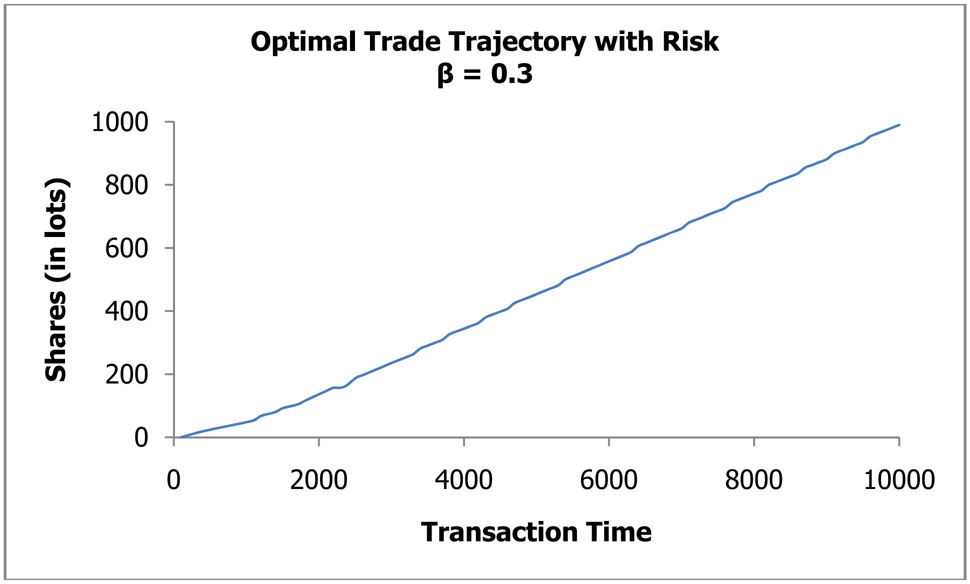
and the total number of traded shares:

$$X = n \sum_{j=1}^{N} \pi_j^{-1}. \quad (29)$$

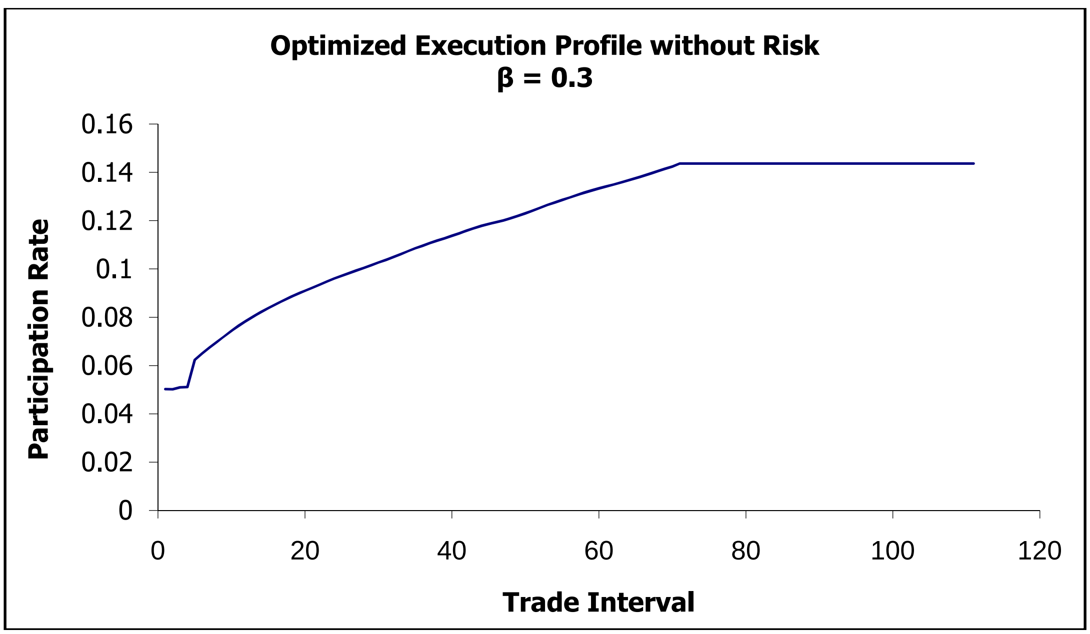
We performed the optimization of (69) using the simulated annealing algorithm (Kirpatrick,S.;C.D. Gelatt, M.P.Vecchi., 1983). The results are summarized below for $|X| \sim 1000\ lots$, $T \sim 10000$ (transaction time), $\alpha = 1.5, \beta = 0.3$. It is important to mention that there does not exist a solution for strict equalities in the constraints.



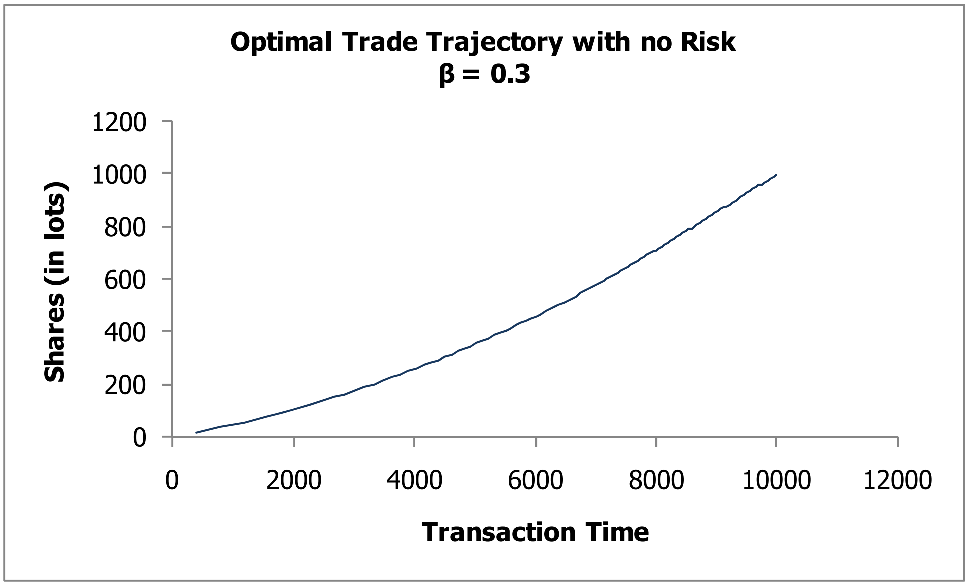

**Graph 8:** An optimal participation rate schedule $\{\pi_k\}_{k=1}^{N=101}$ is found using the simulated annealing method for a risk-averse trader, with parameter $\frac{\lambda}{\mu}\sigma^2 n = 1\ x\ 10^{-9}$. The optimal cost is $\frac{E}{|\mu X|} = 28.223$, ($\frac{E}{|\mu n|} = 28223$), risk term $\frac{\lambda V}{|\mu n|} = 3.7$.




**Graph 9:** Optimal expected number of traded transactions $\xi_k$ in function of the trading time $t_k$ for a risk-averse parameter $\frac{\lambda}{\mu}\ \sigma^2 n\ = 1\ x\ 10^{-9}$.



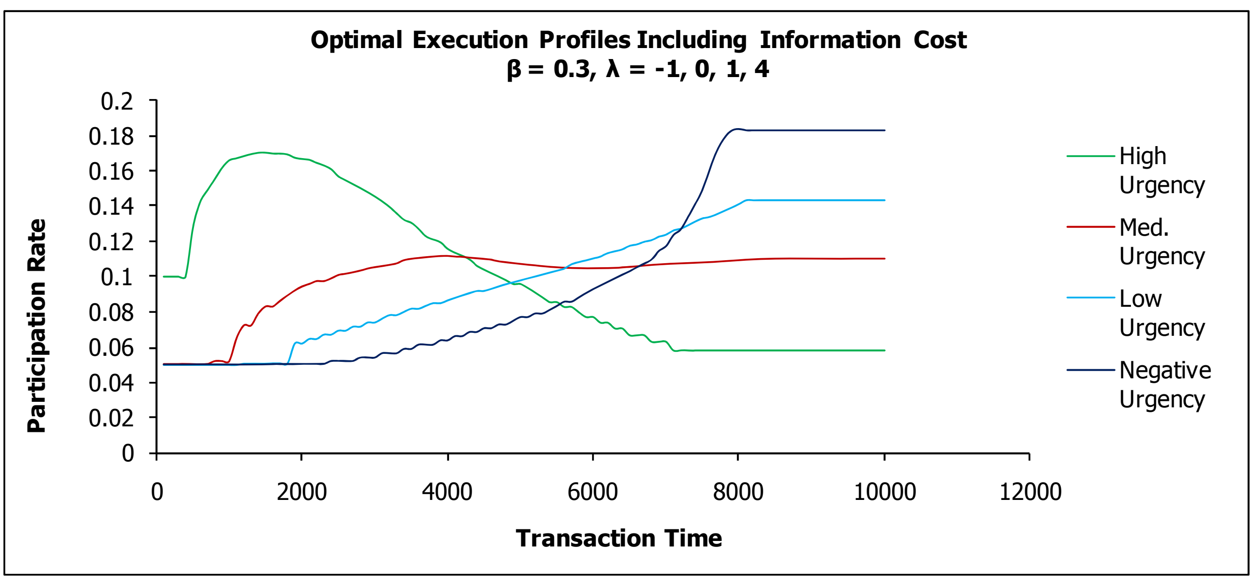

**Graph 10:** Optimal participation rates for $\lambda = 0$. This profile is similar to that shown in graph 5 for $\beta = 1$, but slightly more concave. The optimal cost is $\frac{U(\lambda=0)}{|\mu n|} = 27907, (\frac{U(\lambda=0)}{|\mu X|} = 27.907)$. Comparing to the risk-averse trade, we notice a slight reduction in the optimal trading cost (from 28.223).



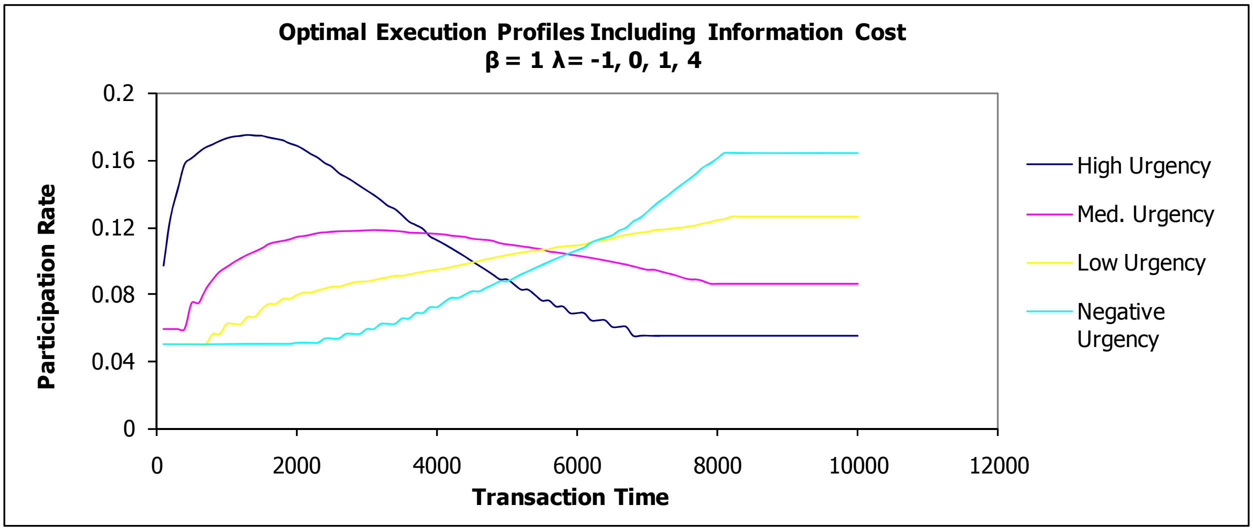

**Graph 11:** Optimal number of traded transactions $\xi_k$ as a function of trading time $t_k$, for $\lambda = 0, \beta = 0.3$. This is similar to graph 6 for $\beta = 1$.

In the following, we will use the notation $L \stackrel{\text{def}}{=} \frac{\lambda}{\mu}\, \sigma^2 n \, \times 10^9$.

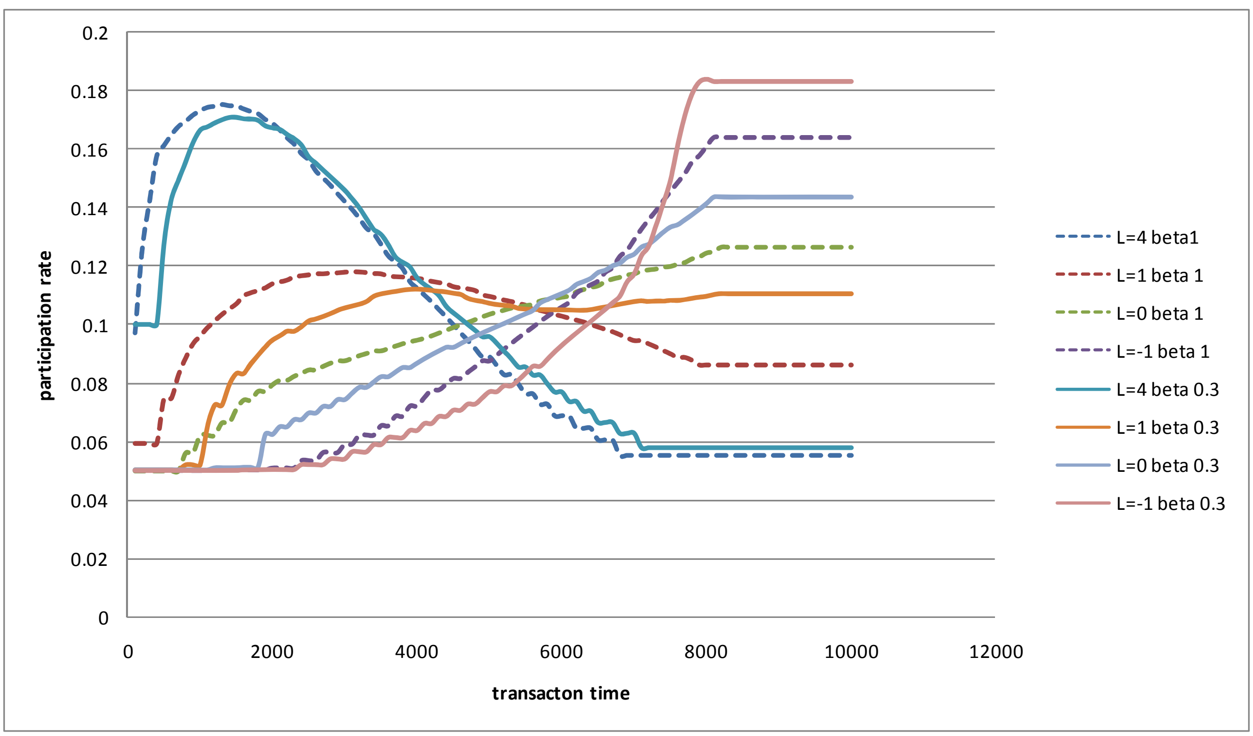
The optimal solutions for various values of risk aversion are shown in Graph 13 for the parameter values $\alpha = 1.5, \beta = 0.3$ that match actual impact data. High-risk aversion demands front-loading the execution profile, while the desire to minimize costs associated with information leakage require on the contrary avoiding high participation rates too early in the trade. The two forces together lead to solutions that begin with a moderate speed, increasing to a maximum partway through the trade in the risk-averse case. Note that this treatment ignores the effect of short-term alpha for trades that are motivated by news. When trading in presence of short-term alpha it may be optimal to accept increased impact costs by trading fast early in the trade, but this lies outside the scope of this paper.



**Graph 12:** Optimal expected trajectories of participation rates $\{\pi_k\}_{k=1}^{N=127}$ comparing with various levels of risk aversion, from a risk-taking trader with parameter $L = \frac{\lambda}{\mu}\sigma^2 n = -1 \, x \, 10^{-9}$ to a highly risk averse trader with $L = 4 \, x \, 10^{-9}$. All optimal trajectories represented here execute $|X| = 1000$ units in an allotted transaction time of $T = 10000$, with $\beta = 0.3$, for an average participation rate of $10\%$.



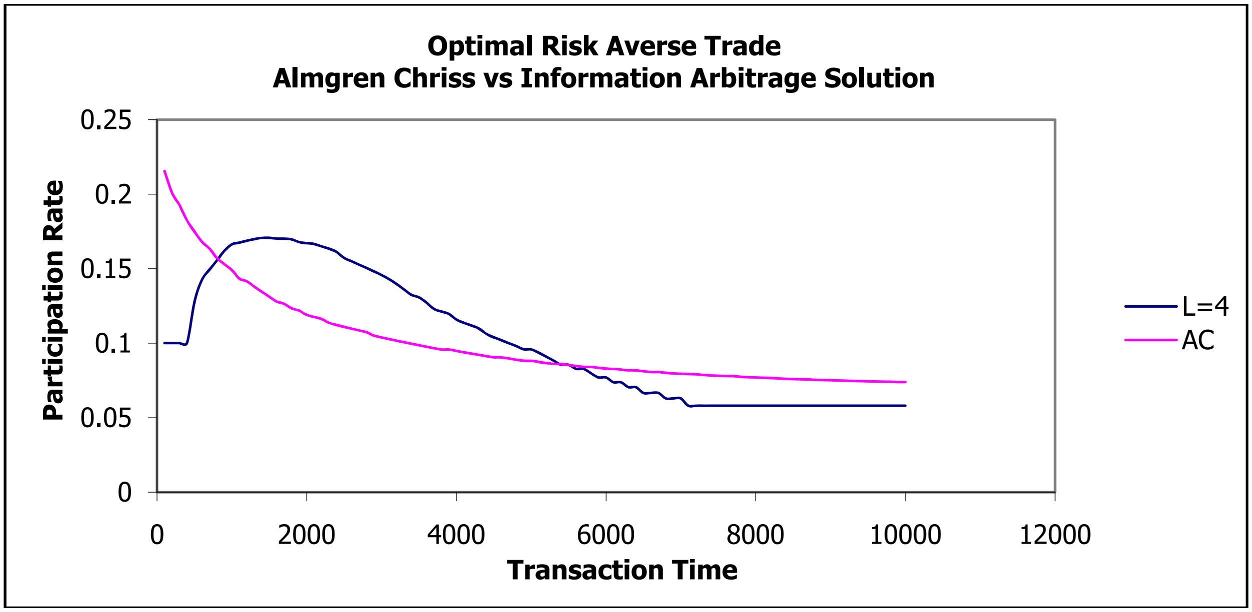

**Graph 13:** Optimal execution profiles as a function of transaction time, for various levels of risk aversion and $\beta = 1$. Higher risk aversion suggests a more front-loaded profile, but consideration of information leakage cost makes it optimal to avoid high participation rates at the very beginning of a trade.




**Graph 14:** The optimal expected participation rate is shown as a function of the transaction time for different values of $\lambda$ and $\beta$. For the same risk, the trajectories for $\beta = 1$ and $\beta = 0.3$ are not significantly different, especially for $L = 4 \times 10^{-9}$. For $\beta = 0.3$ the increase in trading velocity is postponed, by about $200$ units of transaction time for $L = 1 \times 10^{-9}$. After the intersection point, the optimal trajectory for

$\beta = 1$ becomes slower by about $3\%$. The curves for $\lambda = 0$ (green and light blue) show qualitatively the same effects but quantitatively smaller than $L = 1$.

Below, we show the comparison between the optimal solution using information arbitrage theory and the optimal solution according to the formulation of Almgren and Chriss (Almgren, R. and Chriss, N., 2000), for a high level of risk aversion.




**Graph 15:** Comparison of the optimal solution for the risk averse case, using the Almgren Chriss formulation (red line) or the solution presented here that considers the cost of information leakage (blue). The solution that accounts for the cost of information leakage maintains a moderate execution speed for the first $200$ transactions (or $30$ minutes, in this example), before increasing the execution speed. The higher speed is then maintained until approximately $60\%$ of the trade is executed and $45\%$ of the time allotted to the trade is consumed, at which point it crosses $10\%$ falling towards a $6\%$ participation rate for completing the trade.

### 2.5 Optimization respect to the VWAP

In the preceding sections, we found optimal trading trajectories that minimized a linear combination of the expected implementation shortfall and its variance.

The implementation shortfall was defined as the difference between the initial book value of the shares, $XS_0$, and the capture of the trajectory $\sum_{i=1}^{N} n_i \widetilde{S_i}$, which is the full trading revenue upon completion of all trades. Nevertheless, traders have other objectives or benchmarks rather than $XS_0$. These are:

1. Post reversion price or closing price. This is useful to measure the effect of an entry trade on assets under management marked to market after the trade is done and post-trade reversion is complete.
2. Volume-weighted average price during the transactional period or VWAP. This is the average price transacted by the market during the period chosen to execute the trade. The VWAP benchmark is useful to evaluate exit trades that are not too large relative to the ADV because it is less exposed to market effects than implementation shortfall and therefore can be measured more accurately. Note that we consider only the case where the transaction period is fixed in

terms of transaction time; this particular VWAP price is sometimes referred to as "participation-weighed price" or PWP because it represents the average price over the period of time that would be required to complete the trade at a constant participation rate[§].

We will start analyzing the difference between VWAP and the capture for a buy:

$$\Delta(buy) := -\mathrm{VWAP} + \mathrm{capture} \overset{\mathrm{def}}{=} -\frac{1}{t_N}\sum_{i=1}^{N} \tau_i\, \widetilde{S}_\iota + \frac{1}{\xi_N}\sum_{i=1}^{N} \pi_i^{-1}\, \widetilde{S}_\iota\,. \qquad (70a)$$

Here, $t_N$ is the institutional trade duration and $\xi_N$ is the total number of institutional transactions executed in that period. $\{\pi_i\}_{i=1}^{N}$ is the set of participation rates from the first to the last detectable segment and $\tau_i = \pi_i^{-2}$ is the minimum expected transactional market period for the detection of the $i^{th}$ segment. $\pi_i^{-1}$ is the expected value of the number of the institutional transactions executed in the $i-$detectable segment, and $\tilde{S}_i$ is the price per share paid by the institution in the $i-$segment. For a sell: $)(sell) := \mathrm{VWAP} - \mathrm{capture}.$

Following the equations $(42)$, we have:

$$\Delta(buy, \{\pi_i\}_{i=1}^{N}) = -\frac{\mu}{t_N}\sum_{k=1}^{N} \pi_k^{\beta-1}\, \xi_k^{\alpha-2}\,(\pi_k^{-1}\xi_k - t_k) - \sigma \sum_{k=1}^{N} \pi_k^{-1}\, \varsigma_k \left(\frac{\xi_k}{\xi_N} - \frac{t_k}{t_N}\right), \qquad (70b)$$

where

$$\xi_k \overset{\mathrm{def}}{=} \sum_{i=1}^{k} \pi_i^{-1}, \quad (71)$$

$$t_k = \sum_{j=1}^{k} \tau_j = \sum_{j=1}^{k} \pi_j^{-2}\,, \quad (72)$$

$\mu, \alpha, \beta, \sigma$ are constant parameters and $\varsigma_i$ are random variables with zero Gaussian mean and unit variance associated to the volatility of the asset.

The variance of $\Delta$ is:

$$V(\{\pi_i\}_{i=1}^{N}) := \langle(\Delta - \langle\Delta\rangle)^2\rangle = \sigma^2 \sum_{k=1}^{N} \pi_k^{-2}\left(\frac{\xi_k}{\xi_N} - \frac{t_k}{t_N}\right)^2. \quad (73)$$

We minimize the sum $\langle\Delta\rangle + \tilde{\lambda}V$ applying the Simulated Annealing method, with

$$\langle\Delta\rangle + \tilde{\lambda}\, V = -\frac{\mu}{t_N}\sum_{k=1}^{N} \pi_k^{\beta-1}\, \xi_k^{\alpha-2}\,(\pi_k^{-1}\xi_k - t_k) + \tilde{\lambda}\,\sigma^2 \sum_{k=1}^{N} \pi_k^{-2}\left(\frac{\xi_k}{\xi_N} - \frac{t_k}{t_N}\right)^2, \qquad (74)$$

and constraints to set the number of shares and time allotted for the trade.

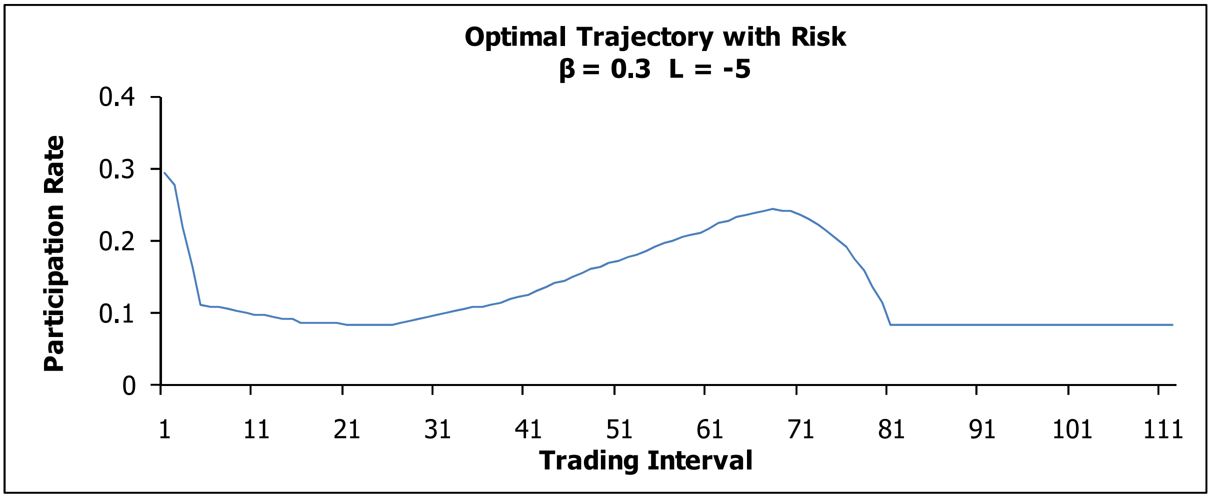
### 2.5.1 Strong constraints (non-discretionary case)

[§] Some trading desks use also the "interval VWAP" as benchmark. Interval VWAP is the average price over whichever period was actually used to execute; this is a poor choice of benchmark because the evaluation period and therefore the benchmark price itself can be easily controlled by the execution strategy.

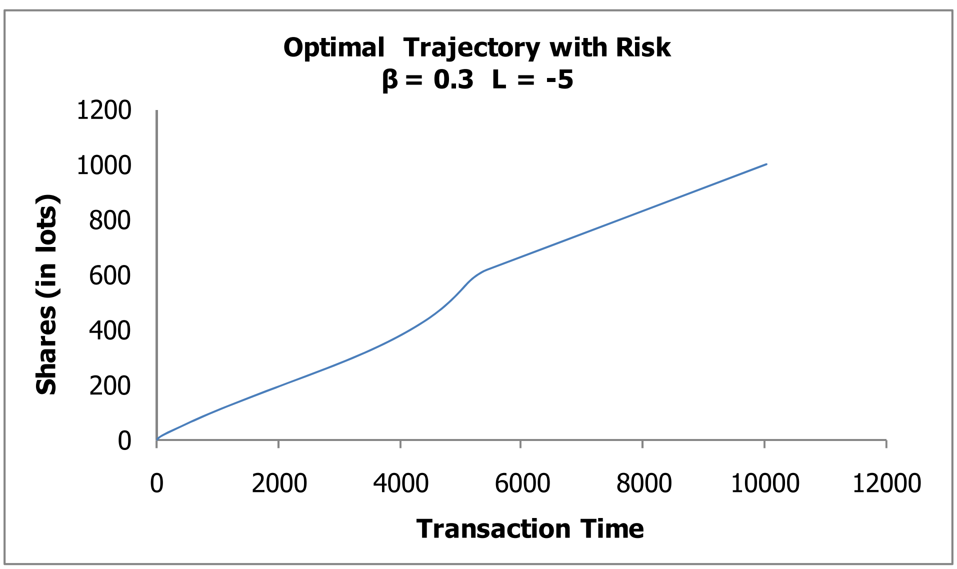
We set the constraints

$$t_N = T, \quad (75)$$
$$\xi_N = X/n. \quad (76)$$

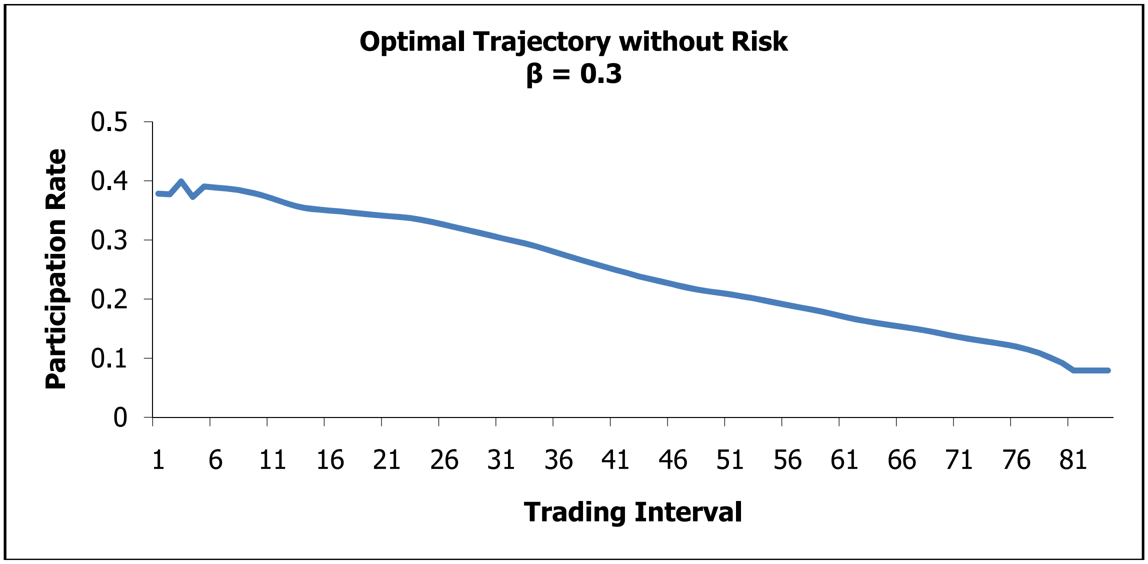
Below, we summarize the results for an order of $|X| \sim 1000\ lots$ and a transaction time of $T \sim 10000$. The exponents are $\alpha = 1.5, \beta = 0.3$. We explore various values for the risk parameter $\tilde{L} = \frac{\tilde{\lambda}}{\mu}\sigma^2 \times 10^7$.



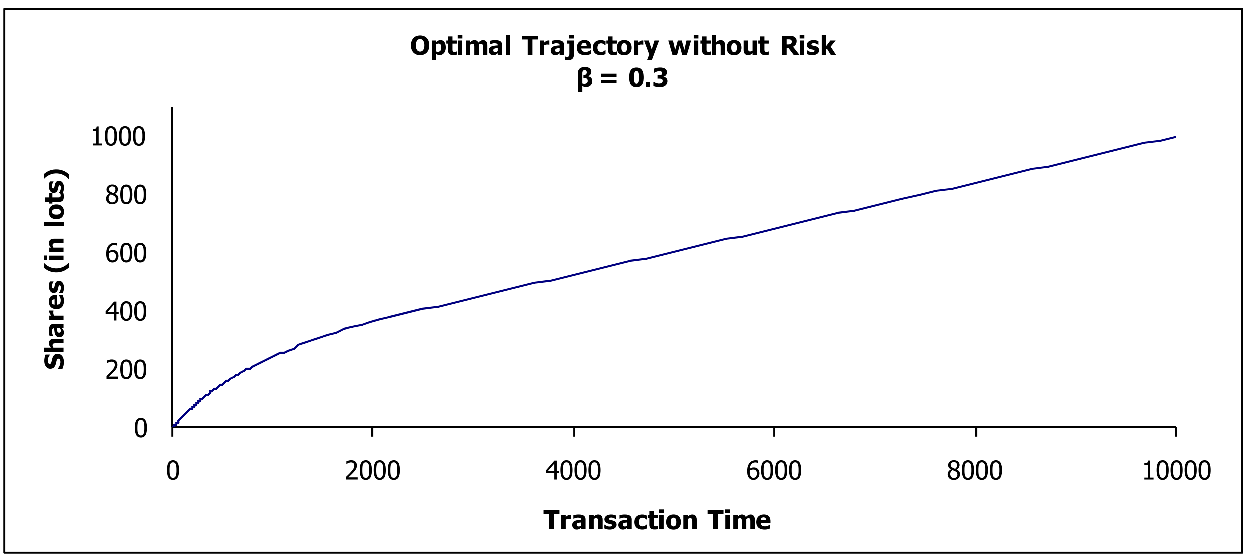

**Graph 16:** Optimal execution trajectory for a risk-adjusted VWAP cost function with negative risk aversion, shows a propensity to front-load the execution to capture better-than-VWAP prices early in the trade, but also preserve some munitions to increase the participation rate late in the trade (intervals $60 - 80$) in accord with negative risk aversion $(L = -5)$. Traders would use negative risk aversion solutions when they hope to be able to find better prices later on – as would be the case in a high volatility market without long-term momentum.



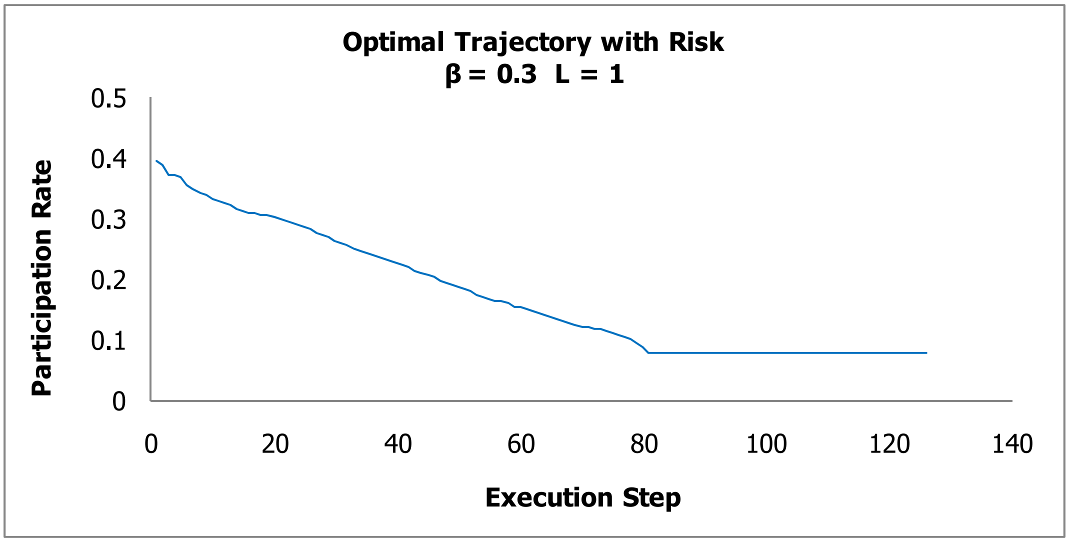

**Graph 17:** Optimal fill profile for a risk-adjusted VWAP cost function with negative risk aversion.



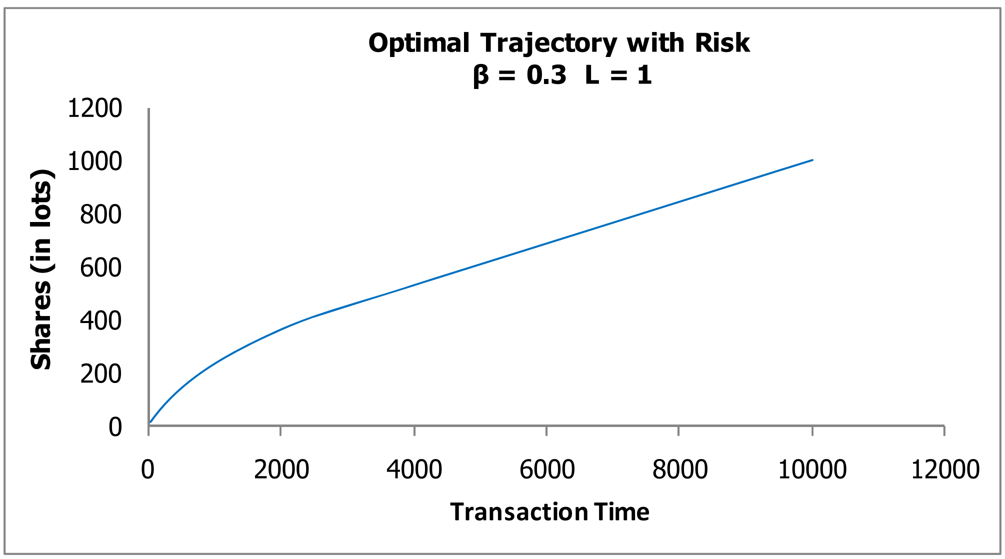

**Graph 18:** Optimal execution trajectory for a risk-neutral VWAP benchmark. Front-loading the execution captures better prices early on relative to the average price over the entire execution period.




**Graph 19:** Optimal execution trajectory for a risk-neutral VWAP benchmark. Front-loading the execution schedule captures better prices early on relative to the average price over the entire execution period. The average price may be worse than with a shortfall-optimized solution, but that is irrelevant when optimizing performance relative to VWAP.




**Graph 20:** Optimal execution trajectory for a risk-averse VWAP cost function. Front-loading is exacerbated by risk aversion.




**Graph 21:** Optimal fill profile for a risk-averse VWAP trade.

### 2.5.2 Weak constrains (discretionary case)

The constraints $(75)$ and $(76)$ are too rigid for most cases. Oftentimes, traders are asked to execute with "discretion." Looking for good price points and perhaps not executing the entire trade if market conditions are unfavorable. Alternatively, some portfolio managers expect the trader to communicate her observations and will add shares if the liquidity is available or reduce the trade size if the impact seems too great. In both cases (whether the trader has discretion or the portfolio manager adjusts percentage of the total order $|X|$ within the specified time and within a reasonable limit instructions), we represent the discretionary execution instructions as a constraint to fill at least a price:

$$p\,X/n \leq \xi_N \leq X/n, \;\; 0 < p \leq 1, \quad (77)$$
$$t_N \leq T, \qquad (78)$$
$$\tilde{S}_i \leq S_{\{limit\}}, (for\ a\ buy) \;\; 1 \leq i \leq N \;. \;\; (79)$$

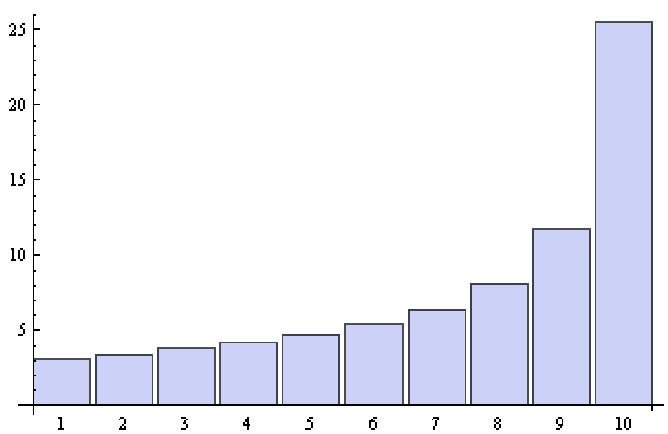
Here, $\tilde{S}_i$ is the price per share paid by the institution in the $i$ –segment.

Because of the nonlinear complexity of the relation $(79)$, in this paper, we will consider only the constraints $(77)$ and $(78)$ (no limit prices).

We performed the simulation for $N = 10\ steps$, which represents average of trading in each step. We used Mathematica7 with different optimization methods. Simulated Annealing or Differential Evolution is the best optimizer. Results are shown below for $\alpha = 1.5, \beta = 0.3, T = 1000, \mu = 1 \frac{\$}{share}\ (buy), |X| = 100\ lots, p = 0.5$ and $\tilde{\lambda} = 0$ (no risk aversion).

$$\langle \Delta(buy) \rangle = -1.39\ \$/share$$
$$\xi_N = 76.2869\ transactions$$
$$t_N \approx 1000\ transactions$$

, Number of lots vs. detectable step




**Graph 22:** Execution profile for a risk-neutral VWAP-optimized execution with a discretionary size shows a preference for front-loading the execution. The higher initial trading speed implies that the first trades are more quickly detected leading to a small number of filled lots in each detectable interval.

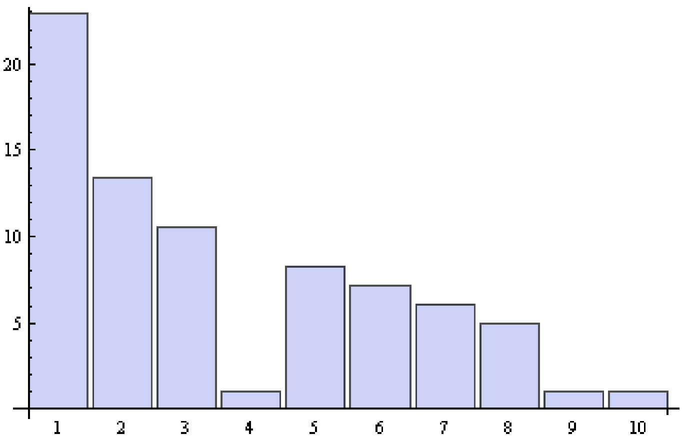
With this solution, we calculate the total cost of trade given by formula $(69)$ with $\lambda = 0$. This is $U(\lambda = 0) = 630.84\ \$ \times lot$[**]$/share$ and $\frac{U(\lambda=0)}{\xi_N lots} = 8.27 \frac{\$}{share}$.

In comparison, if we minimize the shortfall $(69)$ for $N = 10\ steps$, with the weak constraints $(77)$ and $(78)$ such that $p|X| = 76.29\ lots$, we obtain:

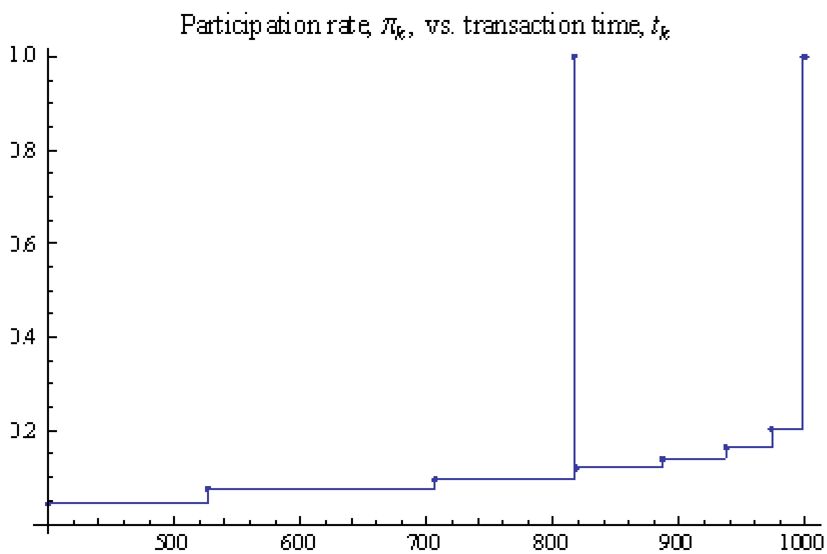
$$U(\lambda = 0) = 448.98\$ \times \frac{lot}{share},\ \xi_N = 76.29 transactions\ , \frac{U(\lambda = 0)}{\xi_N lots} = 5.89\ \$/share$$

, Number of lots vs. detectable step

[**] $1 lot = n\ shares$




**Graph 23:** Execution profile for a risk-neutral execution aiming to minimize implementation shortfall, shows a preference for back loading the execution.




**Graph 24:** The participation rate schedule is shown for the risk-neutral trade aiming to minimize implementation shortfall depicted also in graph 24. Isolated surgical strikes are found in some numerical solutions, indicating that the implementation shortfall minimization objective is consistent with this type of trading strategy. In practical applications, such strikes are used to take advantage of a favorable short-term price move or to avoid averse selection at the end of a trade.

Next, we consider optimal solutions with respect to the VWAP benchmark, with various values of the risk aversion parameter $\tilde{L} = \tilde{\lambda}\, \sigma^2$

- Negative risk aversion $\tilde{L} = -1$

  The best result was obtained by Simmulated Annealing with perturbation scale 3 and boltzmann exponent as
  $-\, df/exp(i/10)$ , where i is de iteration number. The same result was also obtained by Differential Evolution. We can accept this result as the global minimum:

$$\langle \Delta \rangle + \tilde{\lambda}\, V = -147.88 \frac{\$}{share},$$

$\{n[1] = 29.62, n[2] = 3.96, n[3] = 3.84,$
$n[4] = 3.76, n[5] = 3.69, n[6] = 3.64, n[7] = 3.60, n[8] = 3.58, n[9] = 3.56, n[10] = 3.56\}.$

The total number of traded transactions is $\xi_N$ =62.81 and the total transactional time is $t_N = 1000 = T$ . Other local minimimum solutions do not change the trajectory significantly.

- Moderate risk aversion: $\tilde{L} = 1$

$$\left\{-0.14 \frac{\$}{share}, \{n[1] \to 1.97, n[2] \to 2.34, n[3] \to 6.75, n[4] \to 5.36, n[5] \to 5.59, n[6] \to 5.52, n[7] \to 5.50, n[8] \to 5.50, n[9] \to 5.57, n[10] \to 5.90\}\right\},$$

which was obtained by Simmulated Annealing with perturbation scale = 1 and Boltzman factor $-\frac{df\, log\,(i+1)}{10}$, where $i$ is the iteration number of the method.
For $\tilde{L} = 1$, the total number of traded transactions is $\xi_N = 50$ and the total trading time is $t_N = 271.72$. In this case, the institution do not use all the market transactional time $T$.

- Strong risk aversion: $\tilde{L} = 4$

$$\{-0.01 \frac{\$}{share}, \{n[1] = 8.45, n[2] = 8.92, n[3] = 8.84, n[4] = 8.83, n[5] = 8.82, n[6] = 8.82, n[7] = 8.81, n[8] = 8.81, n[9] = 8.82, n[10] = 8.91\}$$

obtained by Simulated Annealing with perturbation scale 50 or 100 and Boltzmann exponent $-df/(Exp[i/10])$ .

For risk $\tilde{L} = 4$, $\xi_N = 88.04$ and $t_N = 775.22$.

Below, we give the values of the statistical mean $\langle \Delta(buy)\rangle/\mu$ and the risk term or weighted variance $\tilde{\lambda}V$ for the different values of $\tilde{L}$ for any value of $\mu$.

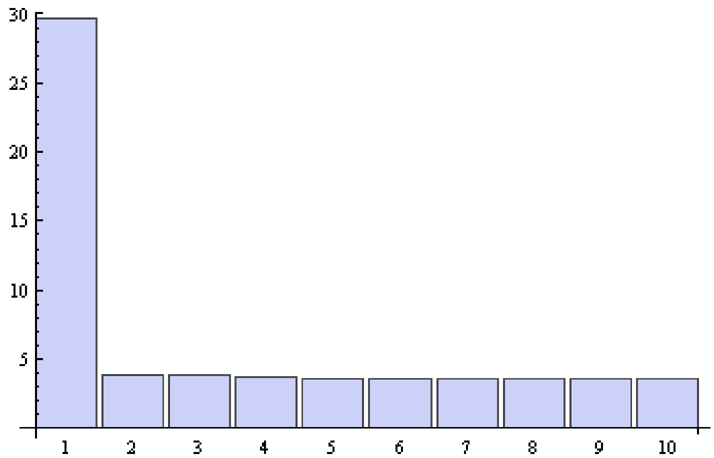
$$\frac{\langle \Delta\rangle}{\mu} = 2.53, \quad (\tilde{\lambda}/\mu)V = -150.41 \text{ for } (\tilde{\lambda}/\mu)\sigma^2 = -1$$
$$\frac{\langle \Delta\rangle}{\mu} = -0.22, \quad \left(\frac{\tilde{\lambda}}{\mu}\right)V = 0.08 \text{ for } (\tilde{\lambda}/\mu)\sigma^2 = 1$$
$$\frac{\langle \Delta\rangle}{\mu} = -0.02, \quad \left(\frac{\tilde{\lambda}}{\mu}\right)V = 0.01 \text{ for } (\tilde{\lambda}/\mu)\sigma^2 = 4$$
$$\frac{\langle \Delta\rangle}{\mu} = -1.39, \quad \text{for } \tilde{\lambda} = 0$$

The values $\tilde{\lambda}\sigma^2 \propto 10^{-1}$ are especially interesting because the value of the risk $\tilde{\lambda}V$ is of the same order or less than the statistical mean value $\langle\Delta\rangle$ and/or the optimal trajectories are non-trivial (variable speed). We show the results below:

- For $\tilde{\lambda}\sigma^2 = -10^{-1}$, the solution is

$\{-12.52 \frac{\$}{share}, \{n[1] = 29.66, n[2] = 3.92, n[3] = 3.81, n[4] = 3.72, n[5] = 3.65, n[6] = 3.60, n[7] = 3.57, n[8] = 3.55, n[9] = 3.54, n[10] = 3.54\}\}.$

, Number of lots vs. detectable step for a risk-taker



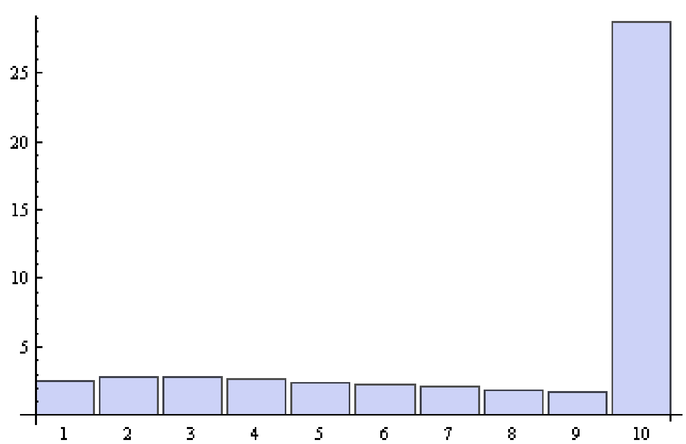

**Graph 25:** The participation rate schedule is shown for negative risk aversion; most of the shares are filled at a low participation rate just above 3% in a first very large detectable interval, the remainder is filled at higher speed at the very end.
For $\tilde{\lambda}\sigma^2 = -10^{-1}$, $\xi_N = 62.56$ and $t_N = 1000$.

Note that the optimal trajectory for this case is the same as the case $\tilde{\lambda}\sigma^2 = -1$, so $\xi_N$ and $t_N$ coincide for both cases, but the optimal function $\langle \Delta \rangle + \tilde{\lambda} V$ is about $10^{-1}$ lesser.

- For $\tilde{\lambda}\sigma^2 = 10^{-1}$, the best solution is

  $\{-0.70 \frac{\$}{share}, \{n[1] = 2.58, n[2] = 2.83, n[3] = 2.83, n[4] = 2.66, n[5] = 2.45, n[6] = 2.24, n[7] = 2.05, n[8] = 1.89, n[9] = 1.76, n[10] = 28.72\}$.

, Number of lots vs. detectable step for a risk-averse trader




**Graph 26:** The participation rate schedule is shown for an averse trader averse to risk of deviating from VWAP; the trade begins aggressively leading to short detectable intervals (and rapid information leakage); the residual is filled at a lower speed in a final interval with a much longer detection time. Front-loading is optimal because the variance on price relative to VWAP is greatest near the end of the trade.

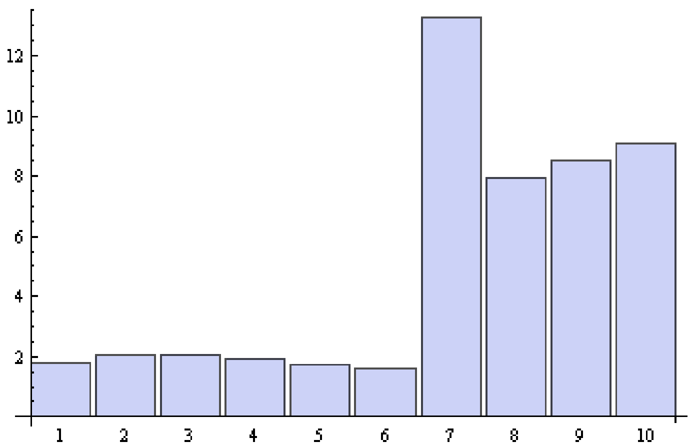
For $\tilde{\lambda}\sigma^2 = 10^{-1}$, $\xi_N = 50$ and $t_N = 876.44$. In comparison with $\tilde{L} = 1$, the optimal trajectory stops almost being constant at the end of the trade. The total number of traded transactions is the same for both cases ($\xi_N = 50$) but $\tilde{L} = 10^{-1}$ takes a longer trading time. It is also less costly.

- For $\tilde{\lambda}\sigma^2 = 4 \times 10^{-1}$, the best solution found is

  $\{-0.32 \frac{\$}{share}, \{n[1] = 1.80, n[2] = 2.05, n[3] = 2.08, n[4] = 1.95, n[5] = 1.76, n[6] = 1.59, n[7] =$ *13.26,n8=7.96,n9=8.50,n10=9.06*,

  obtained by Simulated Annealing with Perturbation Scale□$100$, Boltzmann Exponent Function $[\{i, df, f0\}, -df/(Exp[i/10])]\}]$.

, Number of lots vs. detectable step for a higher risk-averse trader

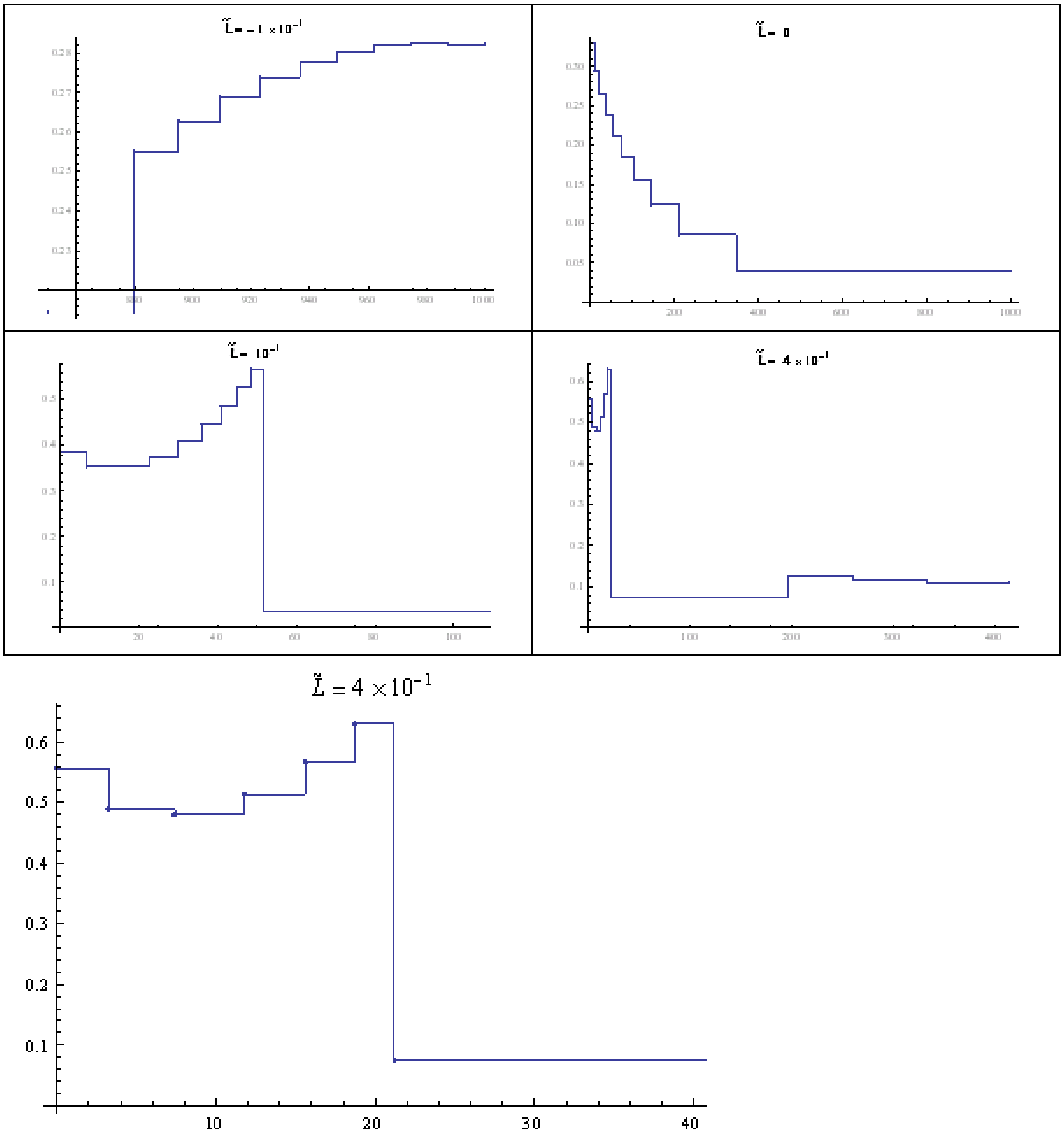



**Graph 27:** The participation rate schedule for a higher level of risk aversion L=0.4. Interestingly, the profile is flatter than for L=1; in the limit of very high-risk aversion the flat execution profile becomes optimal as this has zero variance to the VWAP benchmark.

For $\tilde{\lambda}\sigma^2 = 4 \times 10^{-1}$, $\xi_N = 50$, $t_N = 414.67$. This trajectory is of variable rate, faster and less costly than $\tilde{L} = 4$, though the number of traded transactions is in the limit lower bound 50 (lesser than for $\tilde{L} = 4$).

Additionally, we give the plots of the participation rates $\{\pi_k\}_{k=1}^{10}$ in function of the transaction times $\{t_k\}_{k=1}^{10}$, for the different values of the risk constant $\tilde{L} = (\tilde{\lambda}/\mu)\sigma^2$.

### VWAP-optimized solutions (non-discretionary case)




**Graph 28:** VWAP-optimized participation rate profiles are shown. Negative risk aversion leads to optimal solutions with back-loaded schedules; VWAP optimal solutions without risk aversion are front-loaded to get shares filled early in the trade before most of the market impact has taken place. Solutions averse to VWAP risk are front-loaded over a narrowing window, for higher risk aversion parameters, the flat profile (exactly matching VWAP) becomes optimal.

For $\tilde{L} = -1 \times 10^{-1}$, the participation rate is $\pi = 0.0337$ , for a transactional time $t_k$ running from $0$ to $880\ transactions$. For $\tilde{L} = 0, \pi = 0.33$ corresponds to $t_k \in [0,9] transactions$.

*Reminder*: For $\tilde{L} = 10^{-1}$, $t_{10} = 876.44\ transactions$, i.e., $t_k$ runs from $0$ to $876\ transactions$.

For $\tilde{L} = 4 \times 10^{-1}$, $t_{10} = 414.67$ .

Below, we show an amplified graph of the optimal solution for $\tilde{L} = 4 \times 10^{-1}$, highlighting the early part of the execution schedule where $t_k$ runs from $0$ to $40\ transactions$.

Below, we give the values of the statistical mean $\frac{\langle \Delta(buy) \rangle}{\mu}$ and the risk term or weighted variance $\frac{\tilde{\lambda} V}{\mu}$ for the different values of $\tilde{L} = \left(\frac{\tilde{\lambda}}{\mu}\right)\sigma^2$. All the quantities are non-dimensional.

| $\tilde{L}$ | $\frac{\langle \Delta \rangle}{\mu}$ | $\frac{\tilde{\lambda}}{\mu} V$ | $\xi^*_N(transactions)$ | $t_N(transactions)$ |
|---|---|---|---|---|
| $\mathbf{1 \times 10^{-1}}$ | $-0.93$ | $0.23$ | $50$ | $876.44$ |
| $\mathbf{4 \times 10^{-1}}$ | $-\ 0.46$ | $0.13$ | $50$ | $414.67$ |
| $\mathbf{-1 \times 10^{-1}}$ | $2.52$ | $-15.04$ | $62.6$ | $1000$ |
| $\mathbf{0}$ | $-1.39$ | $0$ | $76.3$ | $1000$ |

**Table I:** Cost contributions (slippage and risk) for various VWAP trading strategies. The risk-neutral strategy beats the VWAP benchmark by 1.39. Only the negative risk aversion solution misses the VWAP benchmark, but it opens the door to giant deviations from the VWAP with a risk value of 15.

For a buy $\mu > 0$, then $\langle \Delta \rangle < 0$ (i.e. the captured price is better than $VWAP$) for the risk-averse or indifferent trader. That means those traders are buying at lower prices than the same-period average price on the market. For the risk-taker, the realized price is worse than VWAP but the range of possibilities is much larger with $\frac{\langle \Delta \rangle}{\mu} \in [-12.52, 17.56]$.

With the optimal trajectories found above for the different values of $\tilde{L}$, we calculate the non-dimensional implementation shortfall of the trade $E/\mu n$ and per share $\frac{E}{\mu n \xi_N}$

| $\tilde{L}$ | $E/\mu n$ | $\frac{E}{\mu n \xi_N}$ |
|---|---|---|
| $\mathbf{1 \times 10^{-1}}$ | $346.56$ | $6.93$ |
| $\mathbf{4 \times 10^{-1}}$ | $372.81$ | $7.46$ |
| $\mathbf{-1 \times 10^{-1}}$ | $327.54$ | $5.23$ |
| $\mathbf{0}$ | $630.84$ | $8.27$ |

**Table II:** Implementation shortfall for the VWAP-optimized solutions. Normalized shortfall is greatest for the most front-loaded execution profiles – this shows that the typical trading desk instructions (to perform well relative to both arrival price and VWAP benchmarks) represents a frustrated optimization problem.

In comparison, we minimize the total cost $(69)$ for $N = 10\ steps$, with the weak constraints $(77)$ and $(78)$ such that $\frac{pX}{n} = \xi^*_N$ (values given in *Table I*). We calculate the total cost of trade, $\frac{U}{\mu n}$, given by formula $(69)$ for different values of $L = (\lambda/\mu)\sigma^2 n$, and arbitrary parameters. Here, $E/\mu n$ is the shortfall, $\frac{E}{\mu n \xi_N}$ is the shortfall per share, $\frac{U}{\mu n \xi_N}$ is the total cost per share, and $\frac{\lambda}{\mu n}V$ is the risk term. All of these quantities are non-dimensional.

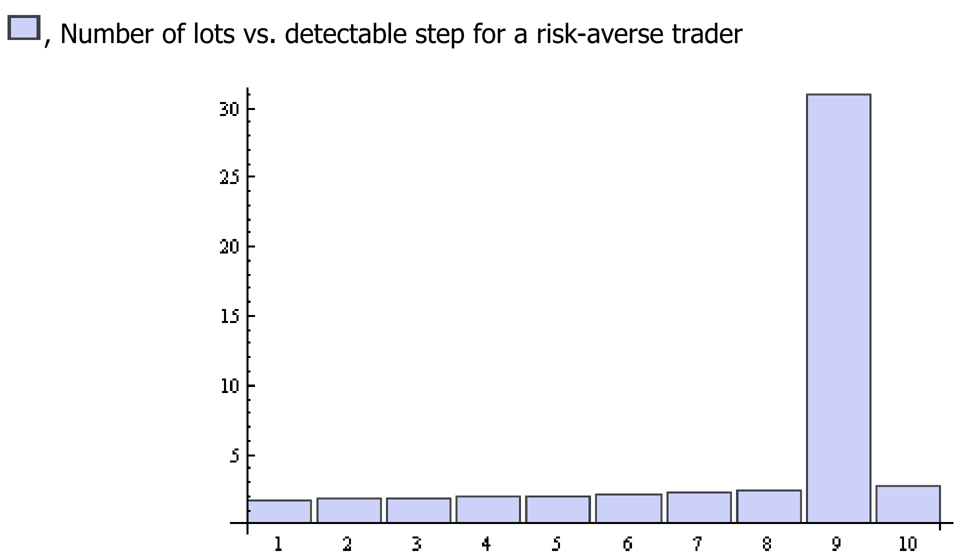
| $L$ | $E/\mu n$ | $\frac{E}{\mu n \xi_N}$ | $\frac{\lambda}{\mu n}V$ | $\frac{U}{\mu n}$ | $\frac{U}{\lambda n \xi_N}$ | $\xi_N$ $(transactions)$ | $t_N$ $(transactions)$ |
|---|---|---|---|---|---|---|---|
| $1 \times 10^{-1}$ | 352.52 | 7.05 | 6250.2 | 6602.73 | 132.05 | 50 | 1000 |
| $1 \times 10^{-3}$ | 324.23 | 6.48 | 79.3 | 403.53 | 8.07 | 50 | 1000 |
| $-1 \times 10^{-4}$ | 308.89 | 4.94 | −101.5 | 207.34 | 3.31 | 62.56 | 1000 |
| $0$ | 448.98 | 5.89 | 0 | 448.98 | 5.89 | 76.3 | 1000 |

**Table III:** Optimal solutions for risk-adjusted implementation shortfall, for the same number of transacted lots as considered in Table II. In the risk-neutral case, the shortfall is reduced from 8.27 to 5.89, a 29% reduction from the front-loaded profile that optimizes VWAP performance. This reveals a significant risk in managing institutional trading desks: since the VWAP performance is easier to measure with accuracy, it is often chosen as an optimization objective either on its own or in conjunction with implementation shortfall; as we see here, this common error can have a giant effect on a firm's trading costs.

The corresponding trajectories are:

- Risk-adjusted cost optimization for a highly high risk aversion $L = 10^{-1}$

  The optimal solution is $\{6602.73, \{n[1] \to 1.75, n[2] \to 1.80, n[3] \to 1.87, n[4] \to 1.95, n[5] \to 2.05, n[6] \to 2.16, n[7] \to 2.30, n[8] \to 2.47, n[9] \to 31, n[10] \to 2.69\}\}$.

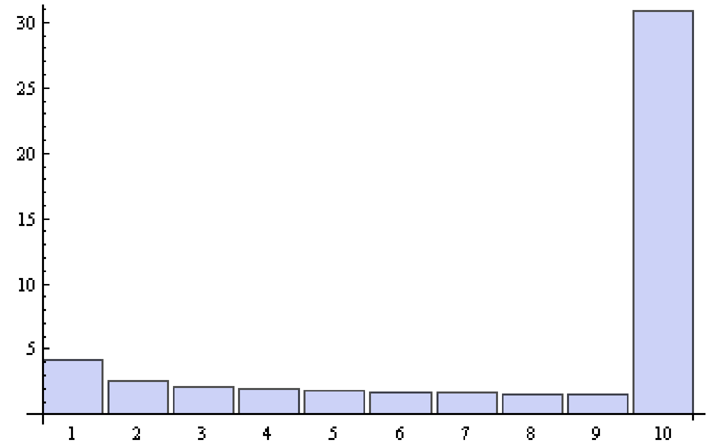



**Graph 29:** Shortfall-optimized solution for a highly risk-averse trader.

- Risk-adjusted cost optimization for a risk averse trader with $L = 10^{-3}$.

  The optimal solution is $\{403.57, \{n[1] = 4.24, n[2] = 2.57, n[3] = 2.13, n[4] = 1.91, n[5] = 1.78,$ $n[6] = 1.69, \ n[7] = 1.63, n[8] = 1.60, n[9] = 1.59, n[10] = 30.88\}\}$

, Number of lots vs. detectable step for a risk-averse trader

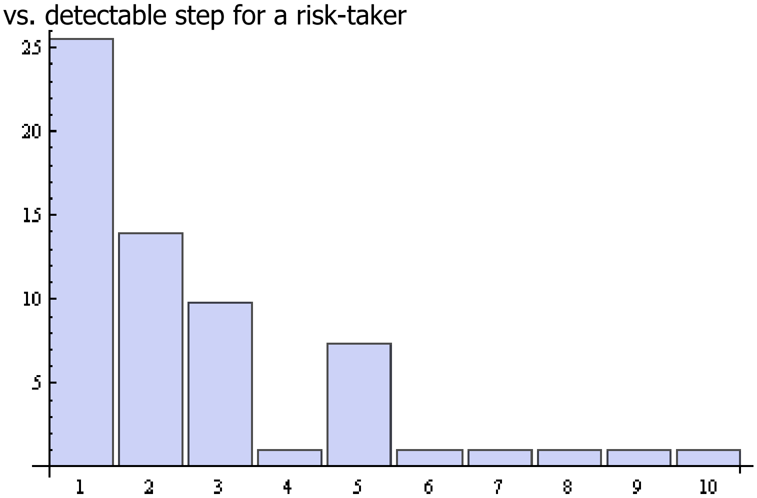



**Graph 30:** Shortfall-optimized solution for a risk averse trader.

- Risk-adjusted cost optimization with negative risk aversion $L = -10^{-4}$

  The optimal solution is
  $\{207.34, \{n[1] = 25.52, n[2] = 13.88, n[3] = 9.80, n[4] = 1.00, n[5] = 7.36, n[6-10] = 1\}\}$ .

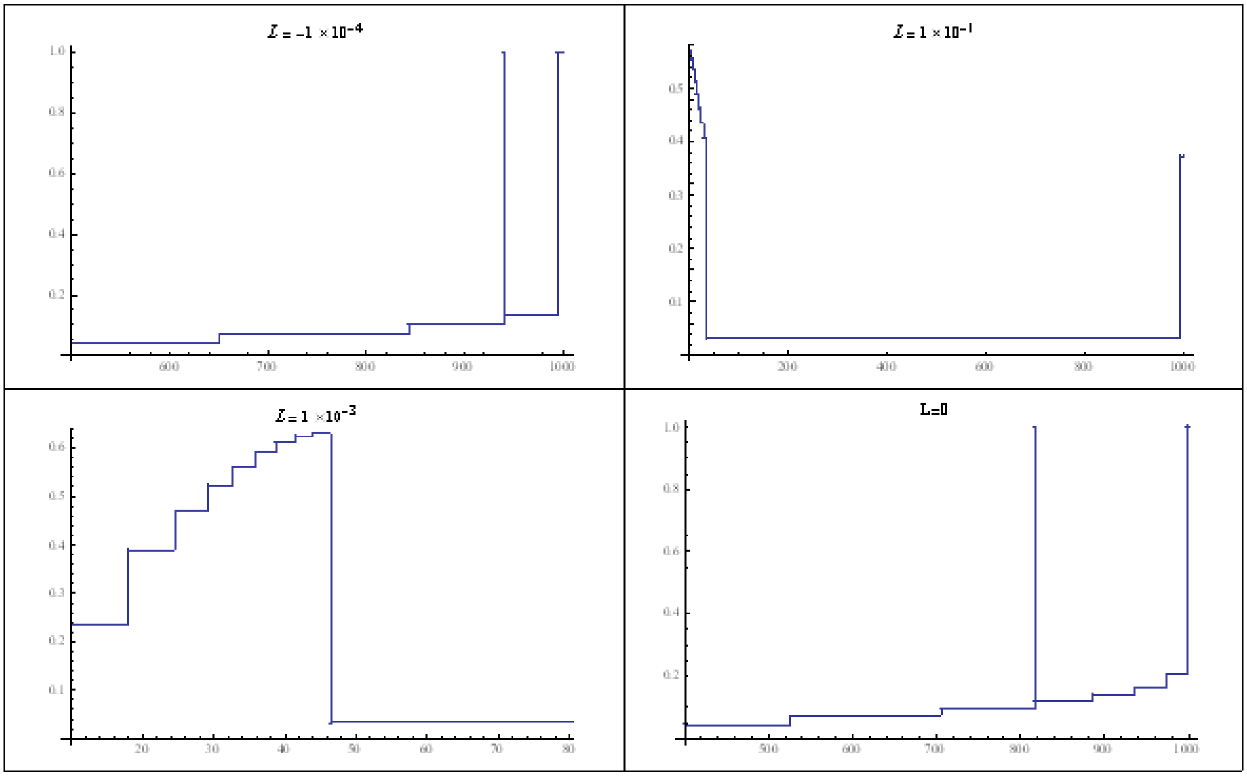
, Number of lots vs. detectable step for a risk-taker

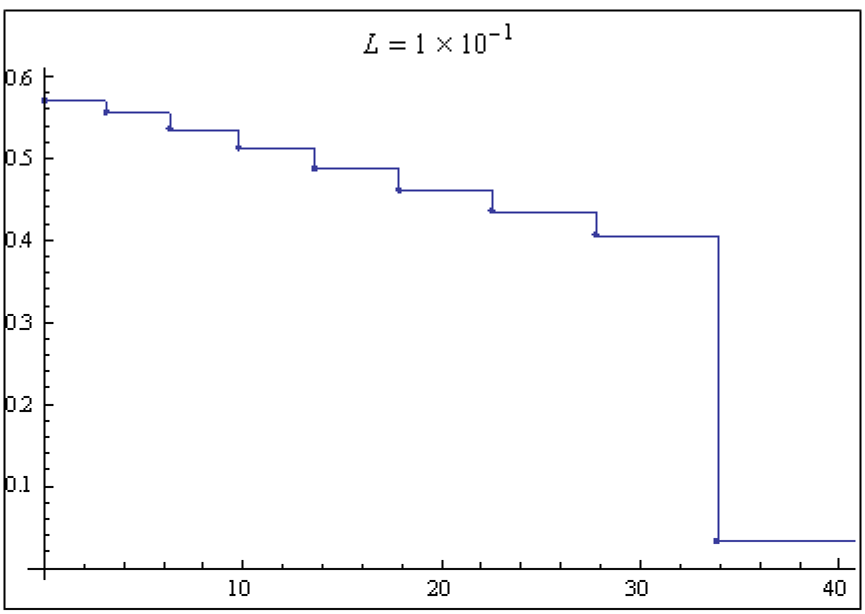



**Graph 31:** Shortfall-optimized solution for a risk taker.

Additionally, we give the plots of the participation rates $\{\pi_k\}_{k=1}^{10}$ in function of the transaction times $\{t_k\}_{k=1}^{10}$, for the different values of the risk constant $L = (\lambda/\mu)\sigma^2 n$.

**Cost-optimized solutions (non-discretionary case)**




**Graph 32:** Shortfall-optimized participation rate profiles are shown for various levels of risk aversion. Negative and neutral risk aversion leads to back-loaded optimal profiles with surgical strikes; risk averse solutions are front-loaded and incur higher shortfalls.
For all the cases, the total transactional time is $t_{10} = 1000\ transactions$.

Below, we give an amplification for the case $L = 10^{-1}$, where $t_k$runs from $0$ to $40$.

It is worthwhile to add some color to the theory by considering an example of a medium trade to put actual numbers on the potential cost savings from using different trading trajectories. Below, we consider a mid-cap trade of $|X| = 25000\ shares$, $n = 250\ shares\ per\ transaction,$ in a $S_0 = \$50$ security, executed at an average participation rate of 10% ($\pi = 0.1$). The impact for a 15-minute interval is estimated to be 10bps for this security, and the entire trade completes in $N = 10$ such intervals. From formula $(40b)$, with $\beta = 0.3, \alpha = 1.5$ we find that a 10bps impact for the first interval correspond to $\widetilde{S_1} - S_0 = 10^{-3} \times \$50 = \mu \times\ 0.1^{-0.2}$ , i.e. $\mu = 0.0315\ \$/share$. If we take an $30\%$ of annual volatility, as in Almgren-Chriss $\sigma = \frac{0.95(\frac{\$}{share})}{\sqrt{day}}$. In our case, we work with transactions units as measure of time and $\tau = \frac{1}{\pi^2} = 100$ is the average number of the market transactions in each detectable interval. If one day consists of 6 hours and 30 minutes and each detectable interval last 15 minutes then, 1 day represents 2600 market transactions. Therefore, $\sigma = \frac{0.019\frac{\$}{share}}{\sqrt{transaction.}}$.

The costs of various solutions are listed below for this example corresponding to *Tables III, $I$* and $II$ respectively.

| $L$ | $\boldsymbol{Risk\ parameter,}$ $\boldsymbol{\lambda}$ $(\$^{-1})$ | $\boldsymbol{Traded\ shares}$ | $\boldsymbol{Shortfall\ per\ share}$ $(\frac{\$}{\boldsymbol{share}})$ | $\boldsymbol{Shortfall,}$ $\boldsymbol{E}$ $(\$)$ | $\boldsymbol{Variance,}$ $\sqrt{\boldsymbol{V}}$ $(\$)$ |
|---|---|---|---|---|---|
| $-1 \times 10^{-4}$ | $-3.49\ \times 10^{-5}$ | 15650 | 0.16 | 2432.49 | 4786.66 |
| 0 | 0 | 19075 | 0.19 | 3535.74 | 6476.54 |
| $1 \times 10^{-3}$ | $3.49\ \times 10^{-4}$ | 12500 | 0.20 | 2553.34 | 1337.96 |

**Table IV:** Comparable shortfall costs of optimal solutions to the risk-adjusted cost optimization (corresponding to Table III)

| $\tilde{L}$ | $\boldsymbol{Risk\ parameter,}$ $\tilde{\boldsymbol{\lambda}}$ $(\boldsymbol{share}/\$)$ | $\boldsymbol{Traded\ shares}$ | $\langle\Delta\rangle$ $(\frac{\$}{\boldsymbol{share}})$ | $\boldsymbol{Variance,}$ $\sqrt{\boldsymbol{V}}$ $(\$/\boldsymbol{share})$ |
|---|---|---|---|---|
| $-1 \times 10^{-1}$ | $-8.73$ | 15650 | 0.08 | 0.23 |
| 0 | 0 | 19075 | $-0.04$ | 0.09 |
| $1 \times 10^{-1}$ | 8.73 | 12500 | $-0.03$ | 0.03 |
| $4 \times 10^{-1}$ | 34.90 | 12500 | $-0.02$ | 0.01 |

**Table V:** Comparable VWAP performance of optimal solutions to the risk-adjusted VWAP optimization problem (corresponding to Table I)

| $\tilde{L}$ | Traded shares | Shortfall per share $(\frac{\$}{share})$ | Shortfall, E (\$) | Variance, $\sqrt{V}$ (\$) |
|---|---|---|---|---|
| $-1 \times 10^{-1}$ | 15650 | 0.17 | 2579.36 | 4725.47 |
| 0 | 19075 | 0.26 | 4967.87 | 3834.24 |
| $1 \times 10^{-1}$ | 12500 | 0.22 | 2729.20 | 1342.57 |
| $4 \times 10^{-1}$ | 12500 | 0.24 | 2935.89 | 2017.43 |

**Table VI:** Comparable shortfall costs of optimal solutions to the risk-adjusted VWAP optimization problem (corresponding to Table II)

**3. Conclusions**

Information arbitrage theory explains the shape of the market impact function from the assumptions that portfolio managers choose order sizes to break even on average and the market sets efficient prices given the information revealed in the order flow about the size distribution of hidden orders.

In this paper, we extended the theory to describe trades that are executed with a varying speed. This required an assumption about the relationship between trading speed and information leakage, which we based on observations of aggregate order flow imbalances according to a biased random walk model. It also requires an assumption about how the market forms expectations of future order flows when observing a trade that was executed with a varying velocity profile. We assumed that expectations depend only on the current speed and total shares traded. Finally, we did not consider underlying alpha; instead we assumed that prices were driven by an unbiased random walk. These simplifying assumptions enabled us to write down the basic equations for market impact of a varying-speed execution and find numerical solutions to optimal execution problems.

It is worth taking a moment to look at these assumptions critically with an eye on the real world. Some of the hypothesis validation analysis lies outside the scope of this paper, so we will simply state some results.

The breakeven condition can be verified using institutional trade data and is quite accurate for trade sizes that dominate the total dollars traded by the funds. The exceptions are very small trades, which are numerous but not economically all that significant, and very large trades, which are quite rare. In both cases the trades are unprofitable if marked-to-market at T+1 average prices. There are logical reasons why a fund engages in such trade sizes: small trades are often required to meet liquidity demands as money flows in and out of a mutual fund; giant trades by oversized funds are associated with longer-term bets where alpha is expected to be realized over a longer period of time, but where the risk that alpha might materialize sooner argues against delaying the execution over a longer time period.

The efficiency condition has been the subject of numerous analyses and is generally very close to being satisfied, we will not dwell on it further here.

The relationship between the transactional time required to detect a hidden order and the speed of trading has been studied extensively in the context of Pipeline's algorithm switching engine (Stephens &

Waelbroeck, 2009) ; a more detailed analysis of aggregate order flow metrics was provided in (Stephens & Waelbroeck, Relating Market impact to Aggregate Order Flow - Why Concavity?, 2009) . As shown there, the market reacts to the presence of persistent demand for a stock with a characteristic timescale that is similar to that which results from the simple random walk argument made in this paper. While these observations lend some support to our hypothesis both in terms of the typical timescale for hidden order detection and its dependency on trading speed, the assumption of a biased random walk should not be viewed as much more than a simplifying assumption intended to define the right average scaling properties: the actual hidden order detection problem is complex and subject of intense investigation by high frequency trading operations; any edge that can be acquired by better predicting future supply and demand in a stock translates directly into profits for these firms.

Our assumption that the market forms expectations of future order flow looking only at current speed and shares filled is intended as a simplifying assumption and not expected to be exact: Given the preponderance of execution profiles based on the Almgren-Chriss algorithm one would expect market participants to recognize a steadily diminishing execution speed as a signature of order flow coming from a risk-averse, front-loaded strategy, and formulate expectations of future order flow accordingly. However, in order to take this into account systematically one would need to understand the different execution profiles that are used by practitioners and their relative probabilities, which would clearly make the theory intractable.

In this paper, we presented two information arbitrage models. The first model did not include a change in the trading velocity from one segment to the next as a cause for the change in the price. If one simply replaces a constant velocity by a changing trading velocity in the temporary impact formula (24), one obtains a non-linear recurrent optimization equation on the trading velocities. Because a rather complicated optimization problem could discourage the best-predisposed users, we cut the temporary impact function to the first order in $1/k$ in the Hurwitz Zeta function – of course large values of $k$ are economically most significant to practitioners so not much is lost in making this approximation. This gave rise to a more workable theory that we called the Information Arbitrage Theory with variable rate of trading, which was described in section 2 and encompasses all of the assumptions discussed in the previous paragraphs.

Armed with the Information Arbitrage Theory with variable rate of trading, we showed that the execution strategy that minimizes implementation shortfall is back-loaded: to minimize overall costs one must trade more slowly at the beginning to avoid information leakage and increase the trading rate progressively to complete on schedule. To minimize the risk-adjusted shortfall requires adopting a profile that reaches a maximum speed sooner and eventually slows down again as the trade nears completion.

Next we considered optimization with respect to the much-used VWAP benchmark. Not surprisingly, the optimal execution strategy to beat the VWAP benchmark is to trade most aggressively at the beginning, to capture more shares before the price moves away. What is remarkable however is that this is exactly the opposite approach from that which minimizes implementation shortfall. It is also interesting to observe that substantial savings to the VWAP price are achievable, which explains why some brokers are willing to provide VWAP executions on a principal basis free of commissions.

The contrast between these two optimal solutions, optimizing for implementation shortfall or VWAP performance, shows that a trading desk operating with instructions to optimize performance to both objectives is faced with a frustrated optimization problem. Improving VWAP performance increases implementation shortfalls, and vice-versa. Since the effect of trades on AUM dollars is at the end of the day measured by the implementation shortfall, we would argue that it is counter-productive for a trading desk to adopt a VWAP objective even as part of their optimization solution – it can only drive the desk in the wrong direction. It is ironic in this regard that the often-used Almgren Chriss solution is often regarded as a risk-adjusted implementation shortfall algorithm, since the resulting front-loaded execution schedule is in fact appropriate to optimize relative to the VWAP but significantly increases implementation shortfall. The confusion at many trading desks about the choice of optimization objective may partly explain the popularity of this type of algorithm.

Of course these conclusions and this paper are all based on the consideration of average prices in an environment with no underlying alpha. This brings into question the last of the assumptions listed at the beginning of this section, that prices follow a random walk. Again assumption this is difficult to defend, but necessary in order to construct a solvable theory. There are several important deviations from the random walk assumption. First, the Pareto tail of price returns clearly shows that extreme price changes are more common than would be expected from a random walk model. Second, portfolio managers at times act on information that is both timely and shared with other managers, and therefore the stock price would move regardless of their trading. Even in the absence of information, price is not strictly an unbiased Brownian motion: traders rely on various technical indicators and assign significance to particular price levels, as demonstrated for example in (Gomes, C. and Waelbroeck, H., 2009) this has significant implications including a degree of predictability of market liquidity and occasionally also short-term alpha. These deviations from the random walk assumption bring important consequences to the optimal execution problem and are clearly deserving of further study.

A surprising implication of information arbitrage theory is that it predicts that permanent impact is not a linear function of the number of shares traded. Since this conflicts with widespread beliefs, the point deserves further examination. A commonly voiced argument for the linearity of permanent impact is that, if impact were not linear, it would be possible to manipulate prices by buying at one speed and selling at another. We do not think this argument is valid, because market impact depends on the specific conditions of each trade and the trades involved in the manipulation argument are not representative of the average institutional trades. We cannot expect that permanent impact will be the same for uninformed price manipulation trades as for the average institutional trade. Institutional trades are generally not random: if permanent impact represents the average *information content* of institutional trades, a theory of permanent impact cannot be expected correctly to describe the results of uninformed price manipulation trades.

The authors are not aware of any compelling empirical facts that would confirm the hypothesis that permanent impact is linear. Almgren et al. (Almgren, R., Thum, C., Hauptmann E., Li, .H, 2005) analyzed *total permanent* impact in Citibank trade data by looking at the stock's price 30 minutes after trade completion. They found that *if* one postulates that market impact is a stationary process and therefore permanent impact is linear in execution time, it follows that total permanent impact costs depend on the constant average trading velocity (or trade rate) with an exponent of 0.89±0.10, which is not significantly different from one. Combined with the assumed linear dependency on time this would indicate that permanent impact is linear in the trade size. The results of this study however do not appear to be robust to using institutional trade data from a different broker [Almgren, personal communication]. More importantly, the study relies on the *assumption* that trading is a stationary process, so to take it as a proof that permanent impact is linear would be a circular argument. Indeed, it is not difficult to devise a non-stationary process where permanent impact is a square root of trade size, and yet, when the data is fit to a model that is linear in time and a power of trading velocity, the optimal exponent for trading velocity comes out to 0.89.

As we have seen, information arbitrage theory makes several falsifiable predictions: the shape of the impact function as a function of trade size; the non-linearity of permanent impact; and the higher expected cost of executing in a front-loaded strategy in absence of alpha. While an extensive test of these results lies beyond the scope of this paper the evidence at the authors' disposal supports all of these claims. The relevance of our claims at least is not in doubt - the estimated cost differences of various execution profiles are as high as 30-50% for some of the examples considered in this paper.

**Acknowledgements**

The authors are in debt to F. Lillo for conversations on the application of information arbitrage theory to variable speed trading. One of us (HW) wishes to thank J. D. Farmer for making us aware of the core ideas behind arbitrage and hidden orders, and the Santa Fe Institute for their hospitality in numerous productive visits.

**Works Cited**

Abramowitz, M., and Stegun, I. A. (1972). *Handbook of Mathematical Functions.* New York: Dover Publications, Inc.

Almgren, R. and Chriss, N. (2000). Optimal Execution of Portafolio Transactions. *Journal of Risk* , 3(2):5-39.

Almgren, R. (2003). Optimal execution with nonlinear impact functions and trading enhanced risk. *Applied Mathematical Finance.* , 10,1-18.

Almgren, R., Thum, C., Hauptmann E., Li, .H. (2005). Direct Estimation of Equity Market Impact. *Risk* .

Bertismas, D., and Lo, A. (1998). Optimal control of execution costs. *Journal of Financial Markets* , 1,1-50.

Bouchoud, J-P., Farmer, J. D., Lillo, F. How markets slowly digest changes in supply and demand. Academic Press.

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

Chan,L.K.C., and Lakonishok, J. (1993). Institutional trades and intraday stock price behavior. *Journal of Financial Economics,* , 33,173-199.

Farmer, J. D., Gerig, A., Lillo F., Waelbroeck H. (2009). Fair pricing and the market impact of large institutional trades.

Gabaix, X., Gopikrishnan, P.,Plerou, V. and Stanley, H.E. (2006). Institutional investors and stock market volatility. *Quarterly Journal of Economics,* , 121, 461-504.

Gomes, C. and Waelbroeck, H. (2009). An Empirical Study of Liquidity Dynamics and Resistance and Support Levels. *J. Quantitative Finance* , (to be published).

Gomes, C. and Waelbroeck, H. (2008). *Effect of Trading Velocity and Limit Prices on Implementation Shortfall.* Pipeline Financial Group Preprint PIPE-2008-09-003.

Gopikrishnan, P.,Plerou, V., Gabaix, X., Stanley, H. E. (2000). Statistical Properties of Share Volume Traded in Financial Markets. *Physical Review E* , 62(4):R4493-R4496.

Kirpatrick,S.;C.D. Gelatt, M.P.Vecchi. (1983). Optimization by Simulated Annealing. . *Science.* , New Series 220 (4598):671-680.ISSN 00368075.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica* , 53,1315-1335.

Lax, M., Cai, W., Xu M. (2006). *Random Processes in Physics and Finance.* New York: Oxford University Press.

Lillo, F., Farmer, J. D., and Mantegna, R.N. (2003). Master Curve for Price-Impact Function. *Nature* , 421, 129-130.

Loeb, T. F. (1983). Trading Cost: The critical link between investment information and results. *Financial AnalystsJournal* , 39(3):39-44.

Moro, E.,Moyano, L.G.,Vicente, J., Gerig, A., Farmer, J.D.,Vaglica, G.,Lillo,F.,and Mantegna, R.N. (2009). Market impact and trading protocols of hidden orders in stock markets. . *Technical Report* .

Obizhaeva,A., and Wang, J. (2006). Optimal trading strategy and supply/demand dynamics. *Technical Report.AFA 2006 Boston Meetings Paper.*

Stephens, C. R., & Waelbroeck, H. (2009). Algorithm Switching: Co-adaptation in the Market Ecology. *J. of Trading* , 59-73.

Stephens, C. R., & Waelbroeck, H. (2009). *Relating Market impact to Aggregate Order Flow - Why Concavity?* New York: Pipeline Preprint 2009-10-003.

Stephens, C. R., Waelbroeck, H. (2009(in press)). Institutional Investor Guides: Algorithmic Trading IV.

Torre, N. (1997). *Market Impact Model Handbook.* Berkeley: Barra Inc.