



## **OPTIMAL LIQUIDITY TRADING**

**Gur Huberman**

*Columbia Business School*

**Werner Stanzl**

*Yale School of Management*

**OCTOBER 7, 2000**

This paper can be downloaded without charge from the  
Social Science Research Network Electronic Paper Collection at:

[http://papers.ssrn.com/paper.taf?abstract\\_id=249973](http://papers.ssrn.com/paper.taf?abstract_id=249973)

# Optimal Liquidity Trading\*

Gur Huberman

Werner Stanzl

Columbia Business School

Yale School of Management

October 7, 2000

\*We are grateful to Prajit Dutta and Larry Glosten for numerous conversations and comments. Address for correspondence: Werner Stanzl. Yale School of Management, International Center for Finance, 135 Prospect Street, P.O. Box 208200, New Haven, CT 06520-8200, USA. Phone: (203) 436-0666, fax: (203) 436-0630, email: [werner.stanzl@yale.edu](mailto:werner.stanzl@yale.edu).

## **Abstract**

We study optimal liquidity trading in a framework where trade size has a price impact. A liquidity trader wishes to trade a fixed number of shares within a certain time horizon and to minimize the mean and variance of the costs of trading. Explicit formulas for the optimal trading strategies show that risk-averse liquidity traders reduce their order sizes over time and execute a higher fraction of their total trading volume in early periods when price volatility increases or price sensitivity decreases. In the presence of transaction fees, numerical simulations suggest that traders want to trade more frequently when price volatility or price sensitivity goes up. In the multi-asset case, price effects across assets have a substantial impact on trading behavior, as does continuous-time trading.

What is the optimal trading sequence of a person who wishes to buy (or sell) a certain portfolio within a certain time, and knows how trades affect prices? Bertsimas and Lo (1998) show that to minimize the expected costs of trading a fixed number of shares, a trader should split his orders evenly over time. They introduce an autocorrelated news process and consider different price processes to study how the even-split result changes in a risk-neutral world. We extend Bertsimas and Lo in several other directions by allowing risk aversion, nonstationary price-impact functions, transaction fees, and continuous trading. We focus exclusively on linear price-impact functions following Huberman and Stanzl (2000), who argue that in the absence of arbitrage opportunities, price-impact functions are linear.

The trader wishes to minimize the mean and variance of total trading costs. A time-consistent solution of this optimization problem exists and is unique if arbitrage opportunities are ruled out. The most important feature is that the optimal trades are independent of past random shocks such as the arrival of new information. If the price-impact function is stationary, trade sizes decline over time. The comparative statics show that lower aversion to risk, lower price volatility, and higher sensitivity of price changes to trade size all lead to less aggressive initial trading, i.e., to more evenly distributed trade sizes. If price-impact functions are not stationary, more aggressive trading is desired when price sensitivity is lower.

In practice, multiple assets are traded simultaneously. In this case, the traded volume of one asset presumably affects not only its own price but also the prices of other assets. To account for cross-price impacts, we extend the analysis to allow for trading a portfolio of securities and derive dynamic trading rules that describe how to optimally rebalance a portfolio.

If the time between trades shrinks to zero, the discrete-time solution of the liquidity trader's problem with risk aversion does not exhibit a well-defined continuous-time limit: the trader would like to trade with infinite speed in the beginning, for he only cares about the variance of his trading

costs when the time between trades is small. However, by incorporating an additional temporary price impact we obtain a well-behaved convergence of the discrete-time solution to its continuous-time limit.

Another modification that allows continuous-time modeling is the incorporation of a fixed per-transaction fee. In the absence of transaction costs, a trader who could select the frequency of his transactions would trade continuously. In the presence of fixed per-transaction costs, a higher price volatility or higher price sensitivity with respect to trade size increases the optimal trading frequency. When the trader transacts more often, he can submit smaller orders in each period, which lessens the total impact of his trades on the prices. In addition, the more frequent the transactions, the smaller is the price volatility between trades.

An alternative to the mean-variance minimization is the objective to minimize the mean and the total (quadratic) variation of the trading costs over time. In this case, the trader cares more about the instantaneous volatilities of the costs accumulated over the whole trading horizon than only about the volatility of the total trading costs. Traders who face cash constraints during trading may prefer this objective to the mean-variance minimization. We prove, however, that the optimal trading strategy has similar qualitative properties as the mean-variance optimal trading strategy.

The remainder of this paper is structured as follows. The liquidity trader's minimization problem is introduced in Section 1, together with examples of speculative and insider trading that boil down to liquidity trading when the number of shares in the portfolio is fixed. Section 2 establishes the existence and uniqueness of the liquidity trader's optimization problem and discusses the properties of the optimal trading behavior. Section 3 examines portfolio trading. Section 4 studies the optimal trading frequency in the presence of fixed transaction costs. Section 5 investigates the convergence behavior of the discrete-time solutions and Section 6 generalizes the discrete-time framework to continuous time. Section 7 contains concluding remarks. All proofs are in the Ap-

pendix.

## 1 The Optimization Problem

Consider a market in which a *single asset* is traded over  $N$  periods. At each period, traders submit their orders simultaneously, and the price change from one period to the next depends on the aggregate excess demand. (Presumably, there is a market maker outside the model who absorbs this excess demand.) Orders are placed before the price change is known. Only market orders are considered.

From the perspective of an individual trader, the total trading volume at time  $n$  is given by  $q_n + \eta_n$ , where  $q_n$  denotes the trader's order size and  $\eta_n$  is a random variable representing the unknown volume of the others. (Negative quantities are sales.) We assume that  $\{\eta_n\}_{n=1}^{\infty}$  is an *i.i.d.* stochastic process with zero mean and finite variance  $\sigma_{\eta}^2$ , defined on the probability space  $(\Omega, \mathcal{F}, \varphi)$ .

The initial price of the asset at time  $n$ ,  $\hat{p}_n$ , which is observed by each trader before choosing his quantity  $q_n$ , is the last price update computed after the trades in the previous period  $n - 1$ . Given the initial price, an individual trader faces the transaction price  $p_n = \hat{p}_n + \lambda_n(q_n + \eta_n)$ , where the real number  $\lambda_n > 0$  measures the price sensitivity with respect to trading volume. Hence, a trader expects to pay  $(\hat{p}_n + \lambda_n q_n)q_n$  if he wants to buy the quantity  $q_n$ . After all trades have been executed at time  $n$ , the new price update for the next period is calculated according to  $\hat{p}_{n+1} = \alpha \hat{p}_n + (1 - \alpha)p_n + \varepsilon_{n+1}$ , where  $0 \leq \alpha \leq 1$  and  $\varepsilon_{n+1}$  incorporates news into the price.

The *updating weight*  $\alpha$  determines the size of the price updates. The lower the  $\alpha$ , the stronger is the impact. If  $\alpha = 0$ , expected price updates and transaction prices coincide, and the price dynamics reduce to

$$p_n = p_{n-1} + \lambda_n(q_n + \eta_n) + \varepsilon_n. \quad (1)$$

Trade size has only a temporary price impact if  $\alpha = 1$ . The stochastic process  $\{\varepsilon_n\}_{n=1}^\infty$ , defined on  $(\Omega, \mathcal{F}, \varphi)$ , is *i.i.d.* with zero mean and variance  $\sigma_\varepsilon^2$ , and it is independent of  $\{\eta_n\}_{n=1}^\infty$ . Note that the zero-mean assumptions are not made for convenience; if one of the two stochastic processes exhibited a nonzero mean, then arbitrage opportunities as discussed in Huberman and Stanzl (2000) would arise.

To define the information set of an individual trader we introduce the vector  $H_n \triangleq [\{q_j\}_{j=1}^{n-1}, \{\varepsilon_j\}_{j=1}^n, \{\eta_j\}_{j=1}^{n-1}]^T$  containing the variables known to the trader before he submits his order in period  $n$ , and the sigma-algebra  $\sigma(H_n)$  that it generates. Then, the set  $M(H_n)$  of all  $\sigma(H_n)$ -measurable functions thus comprises all information available to the trader before his trade at time  $n$ . Note that unlike  $\eta_n$ , the trader does know the news  $\varepsilon_n$ . Furthermore, the trader can only choose a trading strategy  $q_n$  that is an element of  $M(H_n)$ . This setup should best capture real trading activity where the latest news is known before submitting an order, while others' trades are not.

To make later references easier, the price dynamics are summarized by

$$\hat{p}_n = \alpha \hat{p}_{n-1} + (1 - \alpha)p_{n-1} + \varepsilon_n \quad (2)$$

$$p_n = \hat{p}_n + \lambda_n(q_n + \eta_n),$$

for  $n \geq 1$  and given an initial price  $p_0 = \hat{p}_0 > 0$ , with the special case (1) when  $\alpha = 0$ .

The liquidity trader's optimization problem, then, can be formulated as

$$L(Q) \triangleq \inf_{\{q_n \in \mathbf{M}(H_n)\}_{n=1}^N} E\left[\sum_{n=1}^N p_n q_n\right] + \frac{R}{2} \text{Var}\left[\sum_{n=1}^N p_n q_n\right] \quad (3)$$

$$\text{subject to } \sum_{n=1}^N q_n = Q \text{ and (2),}$$

where  $Q > 0$  ( $Q < 0$ ) denotes the number of shares he wants to buy (sell) and  $R \geq 0$  is the risk-aversion coefficient. Expectation and variance are evaluated at time 0 before any of the random price elements are realized. The liquidity trader, aware of the price impact of his trades summarized in (2), minimizes the mean and variance of the total trading costs that must be incurred to enlarge (reduce) his portfolio by  $Q$  shares. Note that the first line in (3) reads  $\sup_{\{q_n \in \mathbf{M}(H_n)\}_{n=1}^N} -E[\sum_{n=1}^N p_n q_n] - \frac{R}{2} \text{Var}[\sum_{n=1}^N p_n q_n]$  for the seller, i.e., he maximizes revenues from trading minus its variance. We will refer to  $E[\sum_{n=1}^N p_n q_n] + \frac{R}{2} \text{Var}[\sum_{n=1}^N p_n q_n]$  as the *loss function*, with  $(q_1, \dots, q_N)$  being the argument of this function.

Problem (3) can also be employed to study speculative or insider trading as in the Kyle (1985) framework provided the total trading volume is fixed or announced before trading (e.g., a new S.E.C. rule on insider trading stipulates that potential insiders have to announce their trades before they actually trade and that they are obliged to commit to their announced trades). A speculator (insider) believes (knows) that eventually the security will trade at a price  $v$ , and tries to profit from the discrepancy between that eventual price and the prices at which he can trade over the next  $N$  periods. Assume that in each period the final price,  $v$ , becomes public information with probability  $1 - \rho$  and that prices stay constant at  $v$  thereafter (in Kyle (1985),  $\rho = 1$  and the equilibrium price process is given by (1)). Further, no updates on  $v$  are made during trading. Buying  $q_n$  at price  $p_n$  thus yields a profit of  $E[\rho^n(v - p_n)q_n]$ . The objective is therefore

$$\pi(Q) \triangleq \sup_{\{q_n \in \mathbf{M}(H_n)\}_{n=1}^N} E[\sum_{n=1}^N \rho^n(v - p_n)q_n] \quad (4)$$

$$\text{subject to } \sum_{n=1}^N q_n = Q \text{ and (2).}$$

Since the revelation of the value  $v$  is assumed to be independent of all random price elements in (2),  $E[\sum_{n=1}^N \rho^n(v - p_n)q_n] = E_\rho[\sum_{n=1}^N (v - p_n)q_n]$  holds after an appropriate change of the probability

measure ( $E_\rho$  is the expectation with respect to this measure). Hence,  $\pi(Q) = \sup E[\sum_{n=1}^N \rho^n (v - p_n)q_n]$  subject to  $\sum_{n=1}^N q_n = Q$  is equivalent to  $L(Q) = \inf E_\rho[\sum_{n=1}^N p_n q_n]$  subject to  $\sum_{n=1}^N q_n = Q$ , and  $\pi(Q) = vQ - L(Q)$ . But this is, apart from the underlying probability measure, the liquidity trader's minimization problem in (3) if  $R = 0$ . Hence, speculators or insiders who fix the number of shares they trade are risk-neutral liquidity traders.

## 2 Minimizing the Mean and Variance of the Trading Costs

This section formulates a recursive version of the problem in (3), provides proofs for the existence and uniqueness of the solution, and presents explicit formulas for the optimal trading policy. The optimal trading path is independent of the resolution of uncertainty, and the traded amounts decline with time. To provide the basic intuition, we begin with a stripped-down version of (3), with  $\alpha = 0$  and two periods. The focus is only on the buyer's problem, because the seller's problem is similar.

### 2.1 The Two-Period Problem

With the price process  $p_n = p_{n-1} + \lambda_n(q_n + \eta_n) + \varepsilon_n$ ,  $n = 1, 2$ , the costs of trading  $q_1$  in period 1 and  $q_2 = Q - q_1$  in period 2 amount to

$$\begin{aligned} C_2 = & [p_0 + \lambda_1(q_1 + \eta_1) + \varepsilon_1]q_1 \\ & + [p_0 + \lambda_1(q_1 + \eta_1) + \lambda_2(Q - q_1 + \eta_2) + \varepsilon_1 + \varepsilon_2](Q - q_1). \end{aligned} \tag{5}$$

The pair  $(q_1, q_2)$  that minimizes  $E[C_2] + \frac{R}{2}Var[C_2]$  is

$$q_1 = \frac{2\lambda_2 - \lambda_1 + R(\lambda_2^2\sigma_\eta^2 + \sigma_\varepsilon^2)}{2\lambda_2 + R(\lambda_2^2\sigma_\eta^2 + \sigma_\varepsilon^2)}Q, \quad (6)$$

$$q_2 = \frac{\lambda_1}{2\lambda_2 + R(\lambda_2^2\sigma_\eta^2 + \sigma_\varepsilon^2)}Q.$$

The stochastic term  $\varepsilon_1$  is a sunk cost by the time the first (and only) decision is made, and therefore does not affect the optimal trades in (6).

To understand how optimal trading is affected by risk aversion, set  $\lambda_1 = \lambda_2 = \lambda$  for the moment. Then,  $q_1 = q_2 = Q/2$  if  $R = 0$ , i.e., the risk-neutral trader splits his total quantity evenly across the two periods. These amounts are not optimal for a risk-averse trader, because the marginal loss at these quantities is  $-1/2R(\lambda^2\sigma_\eta^2 + \sigma_\varepsilon^2)Q < 0$ . In other words, the risk-averse trader is willing to incur higher expected costs in return for a lower variance. In fact, he always wants to equate the marginal change in the expected value of the trading costs,  $\lambda(2q_1 - Q)$ , to the marginal change in its variance weighted with the coefficient  $R$ ,  $R(\lambda^2\sigma_\eta^2 + \sigma_\varepsilon^2)(Q - q_1)$ . Thus, optimality requires that  $q_1 > q_2$  and that a buyer purchases shares in each period. (Note that the trader chooses  $q_1 = Q$  if he aims at minimizing the variance only.)

The ratio  $q_1/q_2 = 2\lambda_2/\lambda_1 - 1 + R(\lambda_2^2\sigma_\eta^2 + \sigma_\varepsilon^2)/\lambda_1$  enables us to perform comparative statics. The optimal trade size at time 1 increases when  $\lambda_1$  decreases, or when  $\lambda_2$ ,  $R$ ,  $\sigma_\eta^2$ , or  $\sigma_\varepsilon^2$  rises. Hence the trader purchases less in price-sensitive periods and shifts his trading volume in the first period when price volatility or the level of risk aversion goes up.

## 2.2 Existence and Uniqueness of Optimal Trading

To solve problem (3) we define a recursive version of (3) and apply dynamic programming arguments to find a time-consistent solution. Again, we consider here only  $\alpha = 0$ ; the Appendix tackles the

more general case  $0 \leq \alpha \leq 1$ .

The state at time  $n$  consists of the price  $\tilde{p}_{n-1} \triangleq p_{n-1} + \varepsilon_n$  which is to be paid for zero quantity, and  $Q_n$ , the number of shares that remain to be bought. The control variable at time  $n$  is  $q_n$ , the number of shares purchased in period  $n$ . Randomness is represented by the  $\varepsilon_k$ 's ( $n+1 \leq k \leq N$ ) and the  $\eta_k$ 's ( $n \leq k \leq N$ ). The objective is the weighted sum of the expectation and variance of the trading costs, and the law of motion is governed by (1) and the following state equations which describe the dynamics of the remaining number of shares to be purchased:

$$Q_1 = Q, Q_{n+1} = Q_n - q_n, \text{ and } Q_{N+1} = 0 \quad (7)$$

for  $1 \leq n \leq N$ . Note that  $Q_1 = Q$  and  $Q_{N+1} = 0$  represent the restriction that  $Q$  shares must be traded within the next  $N$  periods.

Since the objective function in (3) is not additive-separable, it is not obvious at first sight whether there exists an equivalent dynamic program for (3). (A dynamic program is equivalent to (3) if it produces the same solutions as (3).) In what follows, we show that an equivalent dynamic program indeed exists. Consider

$$\begin{aligned} L_n(\tilde{p}_{n-1}, Q_n) = & \inf_{q_n \in \mathbf{M}(H_n)} E_n[p_n q_n + L_{n+1}(\tilde{p}_n, Q_{n+1})] \\ & + \frac{R}{2} Var_n[p_n q_n + L_{n+1}(\tilde{p}_n, Q_{n+1})] \end{aligned} \quad (8)$$

subject to (1) and (7).  $E_n$  and  $Var_n$  denote the conditional expectation and variance in period  $n$ . Beginning at the end in period  $N$  and applying the recursive equation above and the price-impact curve (1) and (7), the functional equation can be solved backwards as a function of the state variables. The procedure ends when we reach the first period in which we know the whole

optimal trading sequence and the value of its loss.

The equivalence of (3) and (8) can be demonstrated by looking at the case  $N = 3$ , where the main ideas of the proof can be illustrated. (For the general case consult the Appendix.) Let us begin by solving the recursive problem in (8). In the last period the trader has no choice but trading the amount  $q_3 = Q_3 = Q - q_1 - q_2$ . Further, his loss is given by  $L_3(\tilde{p}_2, Q_3) = (\tilde{p}_2 + \lambda_3 Q_3)Q_3 + \frac{R\sigma_\eta^2}{2}\lambda_3^2 Q_3^2$  because uncertainty, represented by  $\varepsilon_3$ , has been resolved before the order  $Q_3$  is submitted. In the second period, the trader faces the loss function

$$L_2(\tilde{p}_1, Q_2) = \min_{q_2} E_2[p_2 q_2 + L_3(\tilde{p}_2, Q_3)] + \frac{R}{2} \text{Var}_2[p_2 q_2 + L_3(\tilde{p}_2, Q_3)] \quad (9)$$

which he minimizes with respect to  $q_2$ . It is not hard to verify that the optimal  $q_2^{DP}$  satisfies

$$[2\lambda_3 + R(\lambda_3^2 \sigma_\eta^2 + \sigma_\varepsilon^2)]q_2^{DP} - [2\lambda_3 - \lambda_2 + R(\lambda_3^2 \sigma_\eta^2 + \sigma_\varepsilon^2)]Q_2 = 0. \quad (10)$$

Note that  $q_2^{DP}$  does not explicitly depend on the random shock  $\varepsilon_2$ . History only enters through  $Q_2$ , the remaining shares to be traded. This fact can be interpreted as follows. The total costs of trading in periods 2 and 3 can be written as the difference between  $p_3 Q_2$ , the costs of buying all the remaining shares in the last period, and  $(p_3 - p_2)q_2$ , the “cost savings” of trading at time 2 ( $p_3 \geq p_2$  always holds). From (1), however, it follows that the conditional expectation and variance of the price differential,  $p_3 - p_2$ , are not a function of  $\varepsilon_2$ , and that  $Q_2 \varepsilon_2$  is the only term contained in  $E_2[p_3 Q_2]$  and  $\text{Var}_2[p_3 Q_2]$  that includes  $\varepsilon_2$ . Therefore, the shock  $\varepsilon_2$  has no impact on the optimal  $q_2^{DP}$ .

Finally, at time 1, the liquidity trader computes  $L_1(\tilde{p}_0, Q) = \min_{q_1} E_1[p_1 q_1 + L_2(\tilde{p}_1, Q - q_1)] + \frac{R}{2} \text{Var}_1[p_1 q_1 + L_2(\tilde{p}_1, Q - q_1)]$  where all uncertainty is captured by the terms  $\eta_1$  and  $\varepsilon_2$ , which are

contained in  $p_1 q_1$  and  $L_2(\tilde{p}_1, Q - q_1)$ . The optimal  $q_1^{DP}$  is obtained from

$$(2D + R\sigma_\varepsilon^2)q_1 - (2D - \lambda_1 + R\sigma_\varepsilon^2)Q = 0, \quad (11)$$

where

$$D \triangleq \lambda_2 \left[ \frac{4\lambda_3 - \lambda_2 + 2R(\lambda_3^2 \sigma_\eta^2 + \sigma_\varepsilon^2)}{2[2\lambda_3 + R(\lambda_3^2 \sigma_\eta^2 + \sigma_\varepsilon^2)]} + \frac{R}{2} \lambda_2 \sigma_\eta^2 \right].$$

The same decomposition argument used to explain that  $q_2^{DP}$  is independent of  $\varepsilon_2$  can be applied to verify that  $q_1^{DP}$  is not affected by  $\varepsilon_1$ . Furthermore, the residual trades enter the optimal solution only through their variance, because they are unknown before orders are submitted and have a zero conditional mean. Thus, the whole optimal trading sequence,  $\{q_n^{DP}\}_{n=1}^3$ , is deterministic. Observe that this is also true for any arbitrary  $N > 3$  (see Propositions 2, 3, and 4 below). Consequently, the variance formula (50) in the Appendix can be used to derive

$$L_n(\tilde{p}_{n-1}, Q_n) = \min_{\{q_j\}_{j=n}^3} E_n \left[ \sum_{j=n}^N p_j q_j \right] + \frac{R}{2} \text{Var}_n \left[ \sum_{j=n}^N p_j q_j \right] \quad \text{for } n = 1, 2, \quad (12)$$

proving that  $\{q_n^{DP}\}_{n=1}^3$  constitutes a time-consistent solution to (3).

On the other hand, a time-consistent solution to (3),  $\{q_n^*\}_{n=1}^3$ , satisfies

$q_2^* = \arg \min_{q_2} E_2[\sum_{n=2}^3 p_n q_n] + \frac{R}{2} \text{Var}_2[\sum_{n=2}^3 p_n q_n]$ , which implies  $q_2^* = q_2^{DP}$  and  $\min_{q_2} E[\sum_{n=2}^3 p_n q_n] + \frac{R}{2} \text{Var}[\sum_{n=2}^3 p_n q_n] = L_2(\tilde{p}_1, Q_2)$ . Thus, from  $q_1^* = \arg \min_{q_1} E_1[\sum_{n=1}^3 p_n q_n] + \frac{R}{2} \text{Var}_1[\sum_{n=1}^3 p_n q_n]$  it follows that  $q_1^* = q_1^{DP}$  and that the value of the minimal loss induced by  $q_1^*$  coincides with  $L_1(\tilde{p}_0, Q)$ . Hence, we conclude the equivalence between (3) and (8) for the case  $N = 3$ , and that, due to (10) and (11), the solution must be unique.

The question of the existence of a solution to (3) and (8) is closely related to whether arbitrage as studied in Huberman and Stanzl (2000) is possible. In an environment like here, where prices are

unknown when trades are submitted, Huberman and Stanzl require expected costs,  $E[\sum_{n=1}^N p_n q_n]$ , to be nonnegative for any  $N$  when  $\sum_{n=1}^N q_n = 0$  to rule out arbitrage. They show that the price process (1) is arbitrage-free if and only if the matrices

$$\Lambda_N \triangleq \begin{bmatrix} 2\lambda_2 & \lambda_2 & \lambda_2 & \dots & \lambda_2 \\ \lambda_2 & 2\lambda_3 & \lambda_3 & \dots & \lambda_3 \\ \lambda_2 & \lambda_3 & 2\lambda_4 & \dots & \lambda_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_2 & \lambda_3 & \lambda_4 & \dots & 2\lambda_N \end{bmatrix} \quad (13)$$

are positive semidefinite for all  $N \in \mathbf{N}$ .

Now, the existence of a solution to (3) and (8) is guaranteed by (1) being arbitrage-free whenever traders are risk-averse. In the case of risk neutrality, a slightly stronger condition is needed to ensure the existence of a solution. The proposition below, which is a special case of Corollary 2 in the Appendix, documents these facts.

**Proposition 1** *Suppose one of the following conditions is met:*

- i.  $R > 0$  (trader is risk-averse) and the price process (1) is arbitrage-free or*
- ii.  $R = 0$  (trader is risk-neutral) and the matrices  $\{\Lambda_N\}_{N=1}^\infty$  are all positive definite.*

*Then, the liquidity trader's problem (3) has a unique, time-consistent solution for all  $N \in \mathbf{N}$  that can be derived by solving the dynamic programming problem in (8).*

To give an example of arbitrage-free  $\Lambda_n$ 's where (3) is inextricable in the case  $R = 0$ , consider  $N = 3$ ,  $Q > 0$ ,  $\lambda_1 = \lambda_2$ , and  $\lambda_n = \lambda_2/4$  for  $n \geq 3$ . Expected costs become

$$C_3(q_2, q_3) = \lambda_1[q_2^2 + \frac{q_3^2}{4} + q_2 q_3 - Q(q_2 + q_3)]$$

due to the constraint  $\sum_{n=1}^3 q_n = Q$ . Clearly, all  $\{\Lambda_n\}_{n \geq 2}$  are all positive semidefinite, but  $C_3$  attains no minimum on  $\mathbf{R}^2$ , even though it is bounded from below.

Obviously, if the price-impact slopes are constant and positive, then the  $\Lambda_N$ 's are positive definite. Thus, according to Proposition 1, a solution to the liquidity trader's problem exists, regardless of the trader's type. If the  $\lambda_n$ 's change over time, the  $\Lambda_N$ 's have to be evaluated numerically to apply Proposition 1. Huberman and Stanzl (2000) derive a recursive formula for the determinant of  $\Lambda_N$  that can be used for this computational evaluation. Note that empirical studies imply varying price-impact slopes (for example, see Chordia et al.).

### 2.3 Optimal Trading Behavior

Having established the equivalence between (3) and the corresponding dynamic program, we turn now to the optimal solution itself. If  $\alpha = 0$  and the price-impact function is time-stationary, i.e., the  $\lambda_n$ 's are constant, then the solution to (3) is summarized as follows:

**Proposition 2** *If the price-impact sequence is constant, then the optimal trading quantities of a risk-averse trader are given by*

$$q_n = D \left( A_+ r_+^{N-2-n} - A_- r_-^{N-2-n} \right) Q \quad (14)$$

for  $1 \leq n \leq N-1$ , and

$$q_N = D \lambda^2 (r_+ - r_-) Q$$

where

$$r_{\pm} \triangleq 1 + \frac{\sigma}{2\lambda} \left[ R\sigma \pm \sqrt{R(4\lambda + R\sigma^2)} \right], \quad (15)$$

$$A_{\pm} \triangleq [\lambda^2 + 3\lambda R\sigma^2 + R^2\sigma^4] r_{\pm} - \lambda(\lambda + R\sigma^2),$$

$$D^{-1} \triangleq \frac{r_+^{N-3} - 1}{1 - r_-} A_+ - \frac{r_-^{N-3} - 1}{1 - r_+} A_- + (3\lambda + R\sigma^2) (\lambda + R\sigma^2) (r_+ - r_-) > 0,$$

and  $\sigma^2 \triangleq \lambda^2 \sigma_\eta^2 + \sigma_\varepsilon^2$ . All quantities are strictly positive and the sequence of trades is strictly decreasing. If  $R = 0$ , then  $q_n = \frac{Q}{N}$  for  $1 \leq n \leq N$ .

Therefore, it is optimal for a buyer to purchase shares in each period. The intuition behind declining trade size is as in the two-period example. Due to the price dynamics (1), the variance of the trading costs at time  $n$  depends only on the remaining shares to be traded,  $Q_n - q_n$ , and is increasing in  $Q_n - q_n$ . Given the risk-averse utility, the variance that is produced by distributing trades evenly across time is too high. Thus, the risk-averse trader wants to reduce the variance by submitting a larger order in period  $n$  than in  $n + 1$ . The formulas in (14) and (15) also show that the optimal trades are not directly a function of the  $\eta_n$ 's and the  $\varepsilon_n$ 's, a property explained in the three-period example above.

To illustrate the shape of the solution to (3) given in Proposition 2, we conduct some numerical analysis. Figures 1-3 show the form and the basic comparative-static properties of the formulas. In the simulations we divide a trading day into 30-minute intervals to get 13 trading periods (the NYSE is open from 9:30am to 4:00pm). The amount to be traded is 100,000 shares and the initial price of the financial asset is \$40. A reasonable value for  $\lambda$  is 0.00001 (see Hausman et al. (1992) or Kempf and Korn (1999)); if 1,000 shares are traded, then the price moves by one cent (provided that the price is measured in dollars). The range for the risk-aversion coefficient,  $R$ , is assumed to be between zero and five percent. Randomness is quantified by the magnitudes of the variances of the news revelation and the residual trades, which are set at one percent and 1,000 shares, respectively. The graphs below show that the trajectory of the optimal trades is typically a geometrically decreasing function of  $n$ .

Figure 1a computes the sequence of optimal trades for different values of the risk-aversion factor. The horizontal line (at  $7692 = 100,000/13$ ) shows the strategy of a risk-neutral trader. As can be seen from this figure, higher risk aversion causes the trades to be shifted to early periods. For example, if  $R = 0.01$ , then almost all of the 100,000 shares are bought in the first two periods. The smaller  $R$  becomes, the more closely it approaches the risk-neutral horizontal. The sensitivity of trades to  $R$  is also reflected in the price changes. Figure 1b illustrates that volatility varies more in the early periods when trading strategies differ significantly for different levels of  $R$ .

Figures 2a and 2b look at the reaction of optimal trades to various levels of the variance  $\sigma_\varepsilon^2$ . Like larger values of  $R$ , a higher  $\sigma_\varepsilon^2$  causes traders to redistribute their trades from later to earlier periods (Figure 2a). The change in price is very sensitive to the level of  $\sigma_\varepsilon^2$  (Figure 2b). Besides the direct effect of variance on the price dynamics, it also alters the optimal trades. This further increases the price fluctuation. Different levels of  $\sigma_\eta^2$  show the same effect; a numerical illustration is therefore omitted here.

Figures 3a and 3b consider different values of  $\lambda$  (price sensitivity to trading volume). The higher the  $\lambda$ , the smaller are the orders in the first periods (Figure 3a). When  $\lambda$  is big, large trades in the beginning of trading would drive up prices too much, so that the succeeding purchases would take place at too high a price. Figure 3b demonstrates the sensitivity of price changes to different levels of  $\lambda$ .

If the price slopes,  $\lambda_n$ , are time-dependent, then one cannot derive a closed-form solution to (3), but at least a recursive solution can be obtained by solving the dynamic program in (8).

**Proposition 3** *For an integer  $N > 0$ , let  $\Lambda_N$  in (13) be positive semidefinite if  $R > 0$  and positive definite if  $R = 0$ . The optimal trading sequence of (3) is given by*

$$q_n = \left[ 1 - \frac{\lambda_n}{2\mu_{n+1} + R\sigma_\varepsilon^2} \right] Q_n \quad (16)$$

for  $1 \leq n \leq N - 1$ , and

$$q_N = Q_N$$

where

$$\mu_n = \lambda_n \left[ 1 + \frac{R}{2} \lambda_n \sigma_\eta^2 - \frac{\lambda_n}{2(2\mu_{n+1} + R\sigma_\varepsilon^2)} \right] \quad (17)$$

with initial condition

$$\mu_N = \lambda_N \left( 1 + \frac{R}{2} \lambda_N \sigma_\eta^2 \right),$$

and the  $Q_n$ 's satisfy (7). The minimal loss evolves according to

$$L_n(\tilde{p}_{n-1}, Q_n) = \tilde{p}_{n-1} Q_n + \mu_n Q_n^2 \quad (18)$$

for  $1 \leq n \leq N$ , and  $L(Q) = E[L_1(\tilde{p}_0, Q)]$ .

Figure 4 demonstrates that nonincreasing optimal trades like in Proposition 2 need no longer occur. Again the simulations are for a whole trading day and with  $\lambda_n$  constant but at different levels for even and odd periods. In odd periods where the slope is higher, fewer shares are traded than in even periods. Note that if the price slopes in odd periods are increased sufficiently, a buyer will even sell in these periods (not shown in Figure 4). Other than these two differences, the qualitative properties of the solution given in Proposition 3 is the same as that in Proposition 2; a further discussion of the recursive solution is thus unnecessary.

## 2.4 Trading under a Price Rule with Convex Updating

The focus of this subsection is on (3) with a positive updating weight  $\alpha$  in (2). In this case, the transaction price is higher than the price update after a trade, i.e., trades also have a temporary price impact.

Consider the two-period problem with  $\lambda_1 = \lambda_2 = \lambda$ . Trading costs are given by

$$C_2 = [p_0 + \lambda(q_1 + \eta_1) + (1 - \alpha)\varepsilon_1]q_1 \\ + [p_1 + \lambda(Q - q_1 + \eta_2) - \alpha\lambda(q_1 + \eta_1) + (1 - \alpha)\varepsilon_2](Q - q_1).$$

Minimizing the loss function  $L(C_2) = E[C_2] + \frac{R}{2}Var[C_2]$  yields the ratio of the unique optimal solution

$$\frac{q_1}{q_2} = 1 + \frac{R(1 - \alpha)^2(\lambda^2\sigma_\eta^2 + \sigma_\varepsilon^2)}{(1 + \alpha)\lambda + R\alpha\lambda^2\sigma_\eta^2}. \quad (19)$$

The effect of  $\alpha$  on the optimal trades is now easily seen from (19). More trading is postponed to the second period when  $\alpha$  increases, with the right-hand side of (19) decreasing in  $\alpha$ . This is because the variance of the trading costs is declining in  $\alpha$ . However,  $q_1 \geq q_2$  always.

The previous analysis follows through. In particular, analogues of Propositions 1, 2, and 3 can be derived, although with more complicated expressions (see Theorem 1, Proposition 9, and Corollary 2 in the Appendix). Since the results are qualitatively the same, only the analogue of Proposition 3 is presented here. For convenience, we will use the more general definition  $\tilde{p}_n \triangleq p_n + (1 - \alpha)\varepsilon_{n+1}$ .

**Proposition 4** *Under an arbitrage-free price process (2) (case  $R > 0$ ), or for positive definite  $\{\Lambda_N^\alpha\}_{N=1}^\infty$  as defined in (47) (case  $R = 0$ ), the optimal trades are*

$$q_n = \left[ 1 - \frac{\lambda_n(1 + \alpha + R\alpha\lambda_n\sigma_\eta^2)}{\xi_n} \right] Q_n \quad (20)$$

for  $1 \leq n \leq N - 1$ , and

$$q_N = Q_N$$

where the  $Q_n$ 's obey (7) and  $Q_0 \triangleq 0$ ,

$$\begin{aligned} \mu_n = & \frac{2\mu_{n+1} - 2\alpha(1+\alpha)\lambda_n + R(1-\alpha)^2\sigma_\varepsilon^2}{\xi_n} \frac{R}{2} \lambda_n^2 \sigma_\eta^2 \\ & + \lambda_n \left[ 1 + \alpha - \frac{(1+\alpha)^2 \lambda_n}{2\xi_n} \right] \end{aligned} \quad (21)$$

for  $1 \leq n \leq N-1$ , with

$$\mu_N = \alpha\lambda_{N-1} + \lambda_N \left( 1 + \frac{R}{2} \alpha \lambda_N \sigma_\eta^2 \right),$$

and

$$\xi_n = 2\mu_{n+1} + R \left[ \alpha^2 \lambda_n^2 \sigma_\eta^2 + (1-\alpha)^2 \sigma_\varepsilon^2 \right]$$

for  $1 \leq n \leq N-1$ . The minimal loss is given by

$$L_n(\tilde{p}_{n-1}, Q_{n-1}, \eta_{n-1}) = [\tilde{p}_{n-1} - \alpha\lambda_{n-1}(Q_{n-1} + \eta_{n-1})]Q_n + \mu_n Q_n^2, \quad (22)$$

$1 \leq n \leq N$ ,  $\eta_0 \triangleq 0$ . If  $R = 0$  and all  $\lambda_n$ 's are constant, then  $q_n = \frac{Q}{N}$  for all  $n$ .

Note that Proposition 4 boils down to Proposition 3 when  $\alpha = 0$ . In addition, as in the case  $\alpha = 0$ , the optimal trades depend on the history only through the deterministic state variables  $Q_n$ . Past random shocks do not enter the formulas.

Figures 5a and 5b depict numerical evaluations for these equations. With the numbers from the original example in the section above:  $Q = 100,000$ ,  $p_0 = 40$ ,  $\lambda = 10^{-5}$ , and  $R = 0.01$ . Figure 5a shows the optimal trading sequence for different levels of the updating weight  $\alpha$ , including the case  $\alpha = 0$ . As the figure illustrates, higher  $\alpha$ 's postpone trading to later periods. The result for the two-period example is thus confirmed by the general case. Trading volume appears to be sensitive to the updating weight. The bigger  $\alpha$  gets, the flatter becomes the trading curve. In the limiting

case where  $\alpha = 1$  and volume has a temporary but not a permanent effect on the price, trades are split evenly across all periods. The reason is clear: since today's quantity has no impact on future prices, the optimal trading policy must require that the same be done in each period. Figure 5b shows the corresponding price fluctuations.

### 3 Portfolio Trading

In most applications, portfolio managers trade whole portfolios. In fact, in many cases they merely rebalance a portfolio and the aggregate value of their purchases is approximately equal to the aggregate value of their sales. To examine optimal portfolio trading we extend the setup in a straightforward manner. Individuals are now allowed to trade a portfolio of at most  $M \geq 1$  securities. Prices, trades, and the stochastic variables in (2) and (3) then become  $M$ -dimensional vectors. Note that although the  $\eta_n$ 's are *i.i.d.*, the components of  $\eta_n$  can be intratemporally correlated; the same applies to  $\varepsilon_n$ . We consider here only  $\alpha = 0$ , where trades have only a permanent price impact. The price-impact slopes  $\lambda_n$  are replaced by  $M \times M$ , positive definite matrices  $\Phi_n$  that incorporate all prices and *cross-price impacts*. The price dynamics are thus described by

$$p_n = p_{n-1} + \Phi_n(q_n + \eta_n) + \varepsilon_n, \quad (23)$$

and the liquidity trader's problem (3) reads

$$L(Q) \triangleq \inf_{\{q_n \in \mathbf{M}(H_n)\}_{n=1}^N} E\left[\sum_{n=1}^N p_n^T q_n\right] + \frac{R}{2} \text{Var}\left[\sum_{n=1}^N p_n^T q_n\right] \quad (24)$$

$$\text{subject to } \sum_{n=1}^N q_n = Q \in \mathbf{R}^M \text{ and (23).}$$

The vector  $Q$  summarizes the number of shares to be traded for each asset. It can include both purchases and sales. The covariance matrices of  $\varepsilon_n$  and  $\eta_n$  are  $\Sigma_\varepsilon$  and  $\Sigma_\eta$ , respectively, and  $I_{M \times M}$  is the  $M \times M$  identity matrix.

Similar to the single-asset case, sufficient conditions for the existence of a time-consistent solution to (24) can be found; the solution is unique when it exists and can be obtained by solving the dynamic program in (8). For the risk-averse case, the absence of arbitrage guarantees the existence of a solution, while a more technical condition is required for the risk-neutral case, whose details are not presented here.

**Proposition 5** *The optimal trading sequence of (24) (if a solution exists) is*

$$q_n = [I_{M \times M} - (2\Psi_{n+1} + R\Sigma_\varepsilon)^{-1}\Phi_n] Q_n \quad (25)$$

for  $1 \leq n \leq N-1$ , and

$$q_N = Q_N$$

where

$$\Psi_n = \Phi_n \left[ I_{M \times M} + \frac{R}{2}\Sigma_\eta\Phi_n - \frac{1}{2}(2\Psi_{n+1} + R\Sigma_\varepsilon)^{-1}\Phi_n \right] \quad (26)$$

with initial condition

$$\Psi_N = \Phi_N(I_{M \times M} + \frac{R}{2}\Phi_N\Sigma_\eta),$$

and the  $Q_n$ 's satisfy (7). The minimal loss evolves according to

$$L_n(\tilde{p}_{n-1}, Q_n) = \tilde{p}_{n-1}^T Q_n + Q_n^T \Psi_n Q_n \quad (27)$$

for  $1 \leq n \leq N$ , and  $L(Q) = E[L_1(\tilde{p}_0, Q)]$ .

Again, as for the single-asset case, the optimal trades are a deterministic function of the history. Evidently, the optimal trading strategy for one security depends on the parameters and state variables of all the other securities, unless the  $\Phi_n$ 's and the covariance matrices are all diagonal. Diagonal  $\Phi_n$ 's mean that trading one asset has no impact on the prices of the other assets, and diagonal covariance matrices imply that the stochastic terms are uncorrelated.

If the price-impact matrix in (23) is time-stationary, any sequence of  $N$  portfolio trades can be replicated by splitting the individual amounts of the assets in the portfolio over  $MN$  periods with no change in the expected trading costs. For example, if the price-impact matrix is also diagonal, then trading the portfolio  $(\frac{Q_1}{N}, \frac{Q_2}{N}, \dots, \frac{Q_M}{N})$  in each period minimizes expected costs. The same costs would be incurred by trading  $\frac{Q_1}{N}$  in each of the first  $N$  periods,  $\frac{Q_2}{N}$  in each of the next  $N$  periods, ..., and  $\frac{Q_M}{N}$  in each of the last  $N$  periods. Hence, a risk-neutral trader will be indifferent between portfolio trading and trading the assets individually if he has  $MN$  periods. It is only the time constraint that induces him to prefer portfolios. A risk-averse trader, on the contrary, has in addition a time preference (more trades imply higher volatility of the costs) which makes portfolio trading even more favorable.

The formulas (25)-(27) can be conveniently used to do comparative statics to assess how the sign and the magnitude of the cross-price impacts affect the optimal portfolios and what optimal portfolio rebalancing looks like if the number of shares to be bought and sold changes. We do not present numerical simulations here, but only note that risk aversion causes trades to diminish over time for each individual security, as in the single-asset case.

## 4 Optimal Frequency of Trades

With fixed, positive transaction costs, the frequency of trades emerges endogenously. For simplicity, assume that the price-impact sequence,  $\lambda_n$ , is constant and set  $\alpha = 0$ , and consider only a single

asset. Calendar time is a fixed number  $\tau$  and trades are equally spaced in time. Hence, if the trader chooses to transact  $N$  times, the time between trades is  $\tau/N$ .

The arrival of news (represented by  $\{\varepsilon_n\}_{n=1}^\infty$ ) and other people's trades (represented by  $\{\eta_n\}_{n=1}^\infty$ ) are both *i.i.d.* processes. Their per-unit-of-time variances over the whole trading intervals are  $\sigma_{\eta\tau}^2$  and  $\sigma_{\varepsilon\tau}^2$ , respectively. Therefore, the per-interval variances,  $\sigma_\eta^2(N)$  and  $\sigma_\varepsilon^2(N)$ , satisfy  $\sigma_\eta^2(N) = \tau\sigma_{\eta\tau}^2/N$  and  $\sigma_\varepsilon^2(N) = \tau\sigma_{\varepsilon\tau}^2/N$ .

The risk-aversion coefficient, however, is not assumed to change with the trading frequency. The trader's dislike of price volatility is independent of how short the time between trades is. This is a reasonable premise, since the declining variance per period already decreases the risk part of the trader's utility function, taking into account the desired effect that volatility matters less to the trader if the time between trading becomes smaller.

When the trader has to pay a fixed fee for each of his orders, not only does he decide on the quantity but also on how often he trades within a certain time interval. Without fixed costs the trader would always want to trade at any instant in time. This is because total costs are decreasing in the number of trades, as verified below.

The loss function  $L_1(N)$ , originally given by (18), is now an explicit function of  $N$ , the number of transactions. Moreover, the per-transaction fee is  $k > 0$ ; thus, the total loss is the sum of the losses induced by the variable trading costs and the total fixed costs,  $kN$ .

For the risk-neutral trader the loss function takes a simple expression, namely  $L_1(N) = \sum_{n=1}^N (\tilde{p}_0 + \lambda n \frac{Q}{N}) \frac{Q}{N} + kN = \tilde{p}_0 Q + \lambda \frac{N+1}{2N} Q^2 + kN$ . The unique optimal solution for this problem is  $N^* = \sqrt{\frac{\lambda Q^2}{2k}}$ . Hence, the optimal frequency of trades increases in  $Q$  and  $\lambda$ , and decreases in  $k$ . The more sensitive the price reaction to trades (the higher the  $\lambda$ ), the more often the trader chooses to trade.

For the risk-averse case, the loss function is considerably more complicated. The difference equation given in (17) is a Riccati equation when all  $\lambda_n$ 's are constant. By solving (17) (see (55))

in the Appendix), we obtain a closed-form expression for the total loss  $L_1(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})$ , namely,

$$kN + \tilde{p}_0 Q + \frac{\lambda}{2} \left[ \left[ 2 + R\lambda \frac{\sigma_{\eta\tau}^2}{N/\tau} \right] - \frac{r_+(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})^{2N-1} - r_+(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})}{r_+(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})^{2N} - 1} \right] Q^2 \quad (28)$$

where  $r_+$  is defined in Proposition 2. Both  $r_+$  and  $L_1$  are functions of  $N$  as well as of the total variance,  $\frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau}$ .

It is easily checked that the third term in (28) is decreasing in  $N$  and that it tends to  $\frac{\lambda}{2}$  as  $N \rightarrow \infty$ . It can then be seen that (28) cannot be minimized with respect to  $N$  unless  $k > 0$  (fixed costs are introduced).

The first-order condition of minimizing the loss function with respect to the number of trades renders

$$\frac{d}{dN} \left[ R\lambda \frac{\sigma_{\eta\tau}^2}{N/\tau} - \frac{r_+(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})^{2N-1} - r_+(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})}{r_+(N, \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau})^{2N} - 1} \right] \geq -\frac{2k}{\lambda Q^2} \quad (29)$$

This inequality is met by an appropriate  $N$  unless  $k$  is too small. If this inequality holds for all  $N$ , then the loss function is increasing in  $N$ , implying that trading once is optimal. On the other hand, if there exists a number  $N$  such that the left-hand side of (29) is strictly smaller than the right-hand side, then there exists a number  $N^*$  that equates the two sides. This follows from the smoothness of all functions involved in (29) and the mean-value theorem. It can therefore be concluded that a solution to the loss-minimization problem above always exists when the transaction fee  $k$  is large enough.

However, the uniqueness of the solution cannot be determined. Depending on the parameters there can be multiple solutions. Only if  $R$  is small enough is the loss function strictly convex, providing a global minimum. In general, we have to take a stance on numerical values for the

parameters to evaluate the first- and second-order conditions. The Appendix contains the details of the numerical analysis. We summarize here only the main findings.

First, for all the computations find local convexity of the loss function for the periods under consideration, thereby ensuring a local unique solution for the risk-averse trader. This suggests numerical analysis to be a powerful practical tool to determine the optimal number of trades. Second, the optimal frequency of trades is increasing in the level of risk aversion. The more the trader dislikes volatility, the more often he wants to split his orders: smaller trades move the price less and a higher frequency of trades reduces the price volatility between trades. Third, the optimal frequency also goes up when the price sensitivity rises. If the price-volume elasticity is large, the trader wishes to submit smaller orders in each period, resulting in a higher trading frequency.

## 5 Convergence Behavior

This section asks where the discrete-time solution of (3) converges when the time between trades tends to zero. We prove that the discrete solution does not have a well-defined continuous-time limit when trades have only a permanent and time-stationary price impact. Only when a temporary price impact of trades is introduced, does the discrete-time solution have a proper limit in continuous time. However, the main challenge of going to continuous time is to find a reasonable definition of the trading costs. We address this issue in the next section and propose how continuous-time liquidity trading can be approached in a more general context.

At the start, some new terminology must be introduced. To make continuous-time variables clearly distinguishable from discrete-time ones, we denote the former with a tilde, while the latter keep their notation from the previous sections. In this sense,  $\tilde{Q}_t$  represents the remaining shares to be traded at continuous time  $t$ , and  $-\frac{d\tilde{Q}_t}{dt}$  is the rate of trading. More precisely,  $\tilde{Q}_t$  is assumed to be a continuously differentiable function from  $[0, \tau] \times \Omega$  to  $\mathbf{R}$  with boundary values  $\tilde{Q}_0 = Q$  and

$\tilde{Q}_\tau = 0$ . The set of all these functions shall be denoted by  $\Theta_{[0,\tau]}(Q)$ . For convenience, only the single-asset case is tackled here; the extension to the multi-asset case is straightforward. Moreover,  $\alpha$  in (2) is set to zero throughout.

We now study the limiting behavior of the solution to (3). If  $\lfloor x \rfloor$  denotes the smallest integer greater or equal to  $x$ , then the asymptotics of the optimal trades given in Proposition 2 is as follows:

**Corollary 1** *Suppose that the price-impact sequence is constant. Letting  $N \rightarrow \infty$  gives*

$$q_{\lfloor \frac{t}{\tau} N \rfloor} \rightarrow 0 \text{ for } R \geq 0, \text{ and}$$

$$\frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N} \rightarrow \infty \text{ for } R > 0, \text{ for all } t \in [0, \tau].$$

In the risk-neutral case,  $R = 0$ , we trivially have  $\frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N} = \frac{Q}{\tau}$  for all  $N \in \mathbf{N}$  and  $t \in [0, \tau]$ .

Hence, optimal behavior requires the trading sequence to converge to zero if the number of trading opportunities increases to infinity, what ever the risk type of the trader. In contrast, the difference between the trade pattern of a risk-neutral trader and that of a risk-averse one becomes more pronounced when  $N$  goes up. While the trade size per unit of time stays constant throughout for the risk-neutral case, it rises without bound in the case of risk aversion.

This result suggests that in the limit the risk-averse trader only cares about the volatility of trading costs. The mean of trading costs does not matter, because continuous trading allows the trader to split his trades in such tiny pieces that the total price impact is a fixed number. To see that this is so, consider the loss function  $l(C_N) \triangleq E[C_N] + \frac{R}{2} \text{Var}[C_N]$  with the trading costs  $C_N = \sum_{n=1}^N p_n q_n$ . Due to the price dynamics in (1) this loss function is

$$l(C_N) = p_0 Q + \lambda \sum_{n=1}^N Q_n (Q_n - Q_{n+1}) + \frac{R}{2} \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau} \sum_{n=1}^N Q_n^2. \quad (30)$$

For sufficiently large  $N$ ,  $l(C_N)$  can be approximated by

$$l_{[0,\tau]}(Q) \triangleq p_0 Q + \frac{\lambda}{2} Q^2 + \frac{R}{2} (\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2) \int_0^\tau \tilde{Q}_t^2 dt, \quad (31)$$

$\varphi - a.e.$  ( $\varphi$  almost everywhere), where  $\tilde{Q}_t \in \Theta_{[0,\tau]}(Q)$  approximates the  $Q_n$ 's (the last part of the proof of Proposition 6 in the Appendix shows that such a function can be found; in particular see equations (64)-(66) there). Then, as is evident from (31), the expected trading costs play no role in the limit; only the variance does. The approximation  $l_{[0,\tau]}(Q)$  also explains why risk-averse traders want to trade an infinite amount per unit of time, as documented in Corollary 1. To minimize  $l_{[0,\tau]}(Q)$  on the set  $\Theta_{[0,\tau]}(Q)$ , one would like to choose L-shaped functions with steeper and steeper negative derivatives at zero. Hence, no minimum exists and the continuous-time limit of problem (3) is not well defined.

In the following we sketch how to change the price dynamics to obtain a well-defined relation between the discrete-time and the continuous-time solution of the liquidity trader's problem. For this purpose, we introduce the price process

$$\hat{p}_n = \hat{p}_{n-1} + \lambda q_n + \varepsilon_n \quad (32)$$

$$p_n = \hat{p}_n + \theta \frac{q_n}{\tau/N},$$

where  $p_n$  denotes the transaction price and  $\hat{p}_n$  the *quote*. The first equation in (32) describes the quote dynamics reflecting the permanent price impact of a trade. More formally, this equation resembles (1) with the difference that, for convenience, the residual trades and news are subsumed in one variable,  $\varepsilon_n$ . This can be done if news occurs at the same time that trades take place. The second equation in (32) shows the temporary price impact of a trade. It penalizes high trading

volume per unit of time. The temporary price impact is the driving force for Proposition 6 stated below.

**Proposition 6** *Consider the problem*

$$\begin{aligned} \tilde{l}_{[0,\tau]}(Q) \triangleq \\ \inf_{\tilde{Q}_t \in \Theta_{[0,\tau]}(Q)} p_0 Q + \frac{\lambda}{2} Q^2 + \int_0^\tau \left[ \theta \left( \frac{d\tilde{Q}_t}{dt} \right)^2 + \frac{R}{2} (\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2) \tilde{Q}_t^2 \right] dt. \end{aligned} \quad (33)$$

The solution to the (discrete) liquidity trader's problem (3) with the price dynamics in (32),  $\{q_n\}_{n=1}^N$ , converges to the solution of (33),  $\tilde{Q}_t$ , in the sense that the optimal rate of trading,  $\tilde{q}_t \triangleq \lim_{N \rightarrow \infty} \frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N}$ , satisfies  $\tilde{q}_t = -\frac{d\tilde{Q}_t}{dt}$ , for  $t \in [0, \tau]$ . The explicit formulas for the risk-averse trader are

$$q_n = \frac{\sinh \left[ \psi_N \tau \frac{N+1-n}{N} \right] - \sinh \left[ \psi_N \tau \frac{N-n}{N} \right]}{\sinh(\psi_N \tau)} Q \quad \text{for } 1 \leq n \leq N, \quad (34)$$

where

$$\psi_N = \frac{N}{\tau} \cosh^{-1} \left[ \frac{\tau^2}{2N^2} \frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta + \lambda \frac{\tau}{N}} \right], \quad (35)$$

and

$$\tilde{q}_t = \sqrt{\frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta}} \frac{\cosh \left[ \sqrt{\frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta}} (\tau - t) \right]}{\sinh \left( \sqrt{\frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta}} \tau \right)} Q \quad \text{for } t \in [0, \tau]. \quad (36)$$

The risk-neutral trader again splits trades evenly:  $q_n = \frac{Q}{N}$  for all  $n$ , and  $\tilde{q}_t = \frac{Q}{\tau}$  for all  $t$ . Furthermore, we have  $\tilde{l}_{[0,\tau]}(Q) = \lim_{N \rightarrow \infty} L(Q, N)$ .

This proposition proves that the liquidity trader's problem always has a well-defined continuous-time limit when trades have permanent as well as temporary price impacts. To see why (33) is the right problem to look at for finding the continuous-time limit, just add the costs caused by the

temporary price impacts

$$\sum_{n=1}^N \theta \frac{q_n}{\tau/N} q_n = \theta \sum_{n=1}^N \frac{Q_{n+1} - Q_n}{\tau/N} (Q_{n+1} - Q_n) \approx \theta \int_0^\tau \frac{d\tilde{Q}_t}{dt} d\tilde{Q}_t$$

to  $l_{[0,\tau]}(Q)$  in (31).

The solution to the liquidity trader's problem (3) using the price process (32) has basically the same shape and properties as the solution given in Proposition 2. In particular, the optimal trading strategy is deterministic and trade size is reduced over time. (Note that (34) and (35) give the same optimal trades as Proposition 2 when  $\theta = 0$ .) Therefore, a further discussion of the formulas in Proposition 6 is unnecessary.

## 6 Liquidity Trading in Continuous Time

In this section we introduce a general continuous-time version of the liquidity trader's problem in (3). The first step is to define the price dynamics and the cost function. The multi-asset, continuous-time analogue to (1) is

$$d\tilde{p}_t = \Lambda \tilde{q}_t + \Gamma d\tilde{B}_t, \quad p_0 > 0, \quad (37)$$

where  $\tilde{q} : [0, \tau] \times \Omega \rightarrow \mathbf{R}^M$  is the instantaneous rate of trading,  $\Lambda$  is a  $M \times M$ -dimensional positive definite matrix,  $\Gamma$  is a  $M \times M$ -dimensional, instantaneous variance matrix, and  $\tilde{B} : [0, \tau] \times \Omega \rightarrow \mathbf{R}^M$  is  $M$ -dimensional Brownian motion that represents news and residual trades.

The costs are more subtle to define. As mentioned above, the limit of the discrete-time trading costs is not a good candidate for the costs in continuous time because the mean of the trading costs is a fixed number under the constraint  $\int_0^\tau \tilde{q}_s ds = Q$ . Following Back and Pedersen (1998) and

defining costs here as  $p_0^T Q + \int_0^t d\tilde{Q}_s^T d\tilde{p}_s$  does not work either, because the missing volatility term in  $d\tilde{Q}_t = -\tilde{q}_t dt$  lets  $\int_0^t d\tilde{Q}_s^T d\tilde{p}_s$  vanish.

In contrast, the cost function we propose is

$$\tilde{C}_t \triangleq p_0^T (Q - \tilde{Q}_t) + \frac{1}{2} [Q^T \Lambda Q - \tilde{Q}_t^T \Lambda \tilde{Q}_t] + \int_0^t \tilde{q}_s^T \Lambda \tilde{q}_s ds + \int_0^t \tilde{Q}_s^T \Gamma d\tilde{B}_s. \quad (38)$$

The definition in (38) can be justified as follows. Straightforward computations reveal that the trading costs in discrete time are given by

$$C_n = p_0^T \sum_{j=1}^n q_j + \sum_{j=1}^{n-1} q_j^T \Lambda \sum_{i=j+1}^n q_i + \sum_{j=1}^n q_j^T \Lambda q_j + \sum_{j=1}^n \varepsilon_j^T \sum_{i=j}^n q_i$$

up to period  $n$  when the price obeys (1). This motivates to define the trading costs by

$$\tilde{C}_t = p_0^T \int_0^t \tilde{q}_s ds + \int_0^t \tilde{q}_s^T \Lambda \tilde{Q}_s ds + \int_0^t \tilde{q}_s^T \Lambda \tilde{q}_s ds + \int_0^t \tilde{Q}_s^T \Gamma d\tilde{B}_s$$

in continuous time. The last expression, however, reduces to (38) because of  $\tilde{Q}_t = Q - \int_0^t \tilde{q}_s ds$  and  $\int_0^t \tilde{q}_s^T \Lambda \tilde{Q}_s ds = \frac{1}{2} [Q^T \Lambda Q - \tilde{Q}_t^T \Lambda \tilde{Q}_t]$ . In view of (38), the total costs of trading,  $\tilde{C}_\tau$ , are decomposed of  $p_0^T Q$ , the costs of trading the portfolio  $Q$  if price impacts were absent, and the costs of the accumulated price impacts of the trades.

Denote by  $\tilde{M}[\Theta_{[0,\tau]}(Q)]$  the set of all  $\varphi$ -a.e. continuous  $M(H_t)$ -Markov controls  $\tilde{q}_t$  that satisfy  $\tilde{Q}_t = Q - \int_0^t \tilde{q}_s ds \in \Theta_{[0,\tau]}(Q)$ . In this case  $\{\tilde{C}_t\}_{t \in [0,\tau]}$  is a semimartingale. The continuous-time version of (3) for an arbitrary loss function  $V : \mathbf{R}^{[0,\tau]} \times (\mathbf{R}^{[0,\tau]})^M \rightarrow \mathbf{R}$  ( $\mathbf{R}^{[0,\tau]}$  is the set of all functions mapping from  $[0, \tau]$  into  $\mathbf{R}$ ) then becomes

$$\tilde{L}(Q) \triangleq \inf_{\tilde{q}_t \in \tilde{M}[\Theta_{[0,\tau]}(Q)]} E \left[ V \left( \{\tilde{C}_t, \tilde{Q}_t\}_{t \in [0,\tau]} \right) \right] \quad (39)$$

subject to (37).

The state variables here are  $\tilde{Q}_t$  and  $\tilde{C}_t$  (the price is impounded in the costs and therefore need not be considered here explicitly), whereas the control variable at time  $t$  is  $\tilde{q}_t$ .

The analysis of (39) for general loss functions is quite difficult. Note that (39) falls into the category of finite-fuel stochastic control problems (see Beneš et al. (1980), Karatzas and Shreve (1986), or Karatzas et al. (2000)). Unfortunately, none of the techniques in the stochastic control literature can be applied to (39) for general  $V$ . Perhaps, extensions of the convex duality methods as described in Cvitanic and Karatzas (1992) will lead to a general solution of (39). However, (39) can be solved for specific loss functions  $V$ . We present two examples in the following.

The first example is  $V(f, g) = f(\tau)$  for all  $f \in \mathbf{R}^{[0, \tau]}$  and  $g \in (\mathbf{R}^{[0, \tau]})^M$ . In this case, (39) becomes  $\tilde{L}(Q) = \inf_{\tilde{q}_t \in \tilde{M}[\Theta_{[0, \tau]}(Q)]} E[\tilde{C}_\tau]$  subject to (37). This is the objective of a risk-neutral liquidity trader.

**Proposition 7** *Let  $V(f, g) = f(\tau)$  for all  $f \in \mathbf{R}^{[0, \tau]}$  and  $g \in (\mathbf{R}^{[0, \tau]})^M$  (risk-neutral preferences).*

*Then the solution to (39) is given by*

$$\tilde{q}_t = \frac{1}{\tau} Q \quad \text{for } t \in [0, \tau] \quad (40)$$

and

$$\tilde{L}(Q) = p_0^T Q + \left[\frac{1}{2} + \frac{1}{\tau}\right] Q^T \Lambda Q \quad (41)$$

We have thus obtained the same solution as for the discrete-time liquidity trading problem (3). Since  $\tilde{Q}_t = (1 - \frac{t}{\tau})Q$ , optimality implies that the remaining shares to be traded decline linearly with time for all assets, or equivalently, that the aggregate trading volume for each asset rises linearly with time.

As second example we could study the (mean-variance) problem  $\tilde{L}(Q) = \inf_{\tilde{q}_t \in \tilde{M}[\Theta_{[0,\tau]}(Q)]} E[\tilde{C}_\tau + \frac{R}{2}\tilde{C}_\tau^2]$  subject to (37), which is implied by the loss function  $V(f, g) = f(\tau) + \frac{R}{2}f(\tau)^2$ . Since the analytical solution of it's first-order conditions is unknown, we propose a different risk-averse objective. The main motivation for our approach is the fact that the quadratic variation process  $\langle \tilde{C} \rangle_{t \in [0, \tau]}$  of the semimartingale  $\{\tilde{C}_t\}_{t \in [0, \tau]}$  satisfies

$$\langle \tilde{C} \rangle_t = \lim_{\substack{\max(|t_n - t_{n-1}|: n=1, \dots, N, \\ 0=t_0 \leq t_1 \leq \dots \leq t_N=t) \rightarrow 0}} \sum_{n=1}^N |\tilde{C}_{t_n} - \tilde{C}_{t_{n-1}}|^2$$

for all  $t \in [0, \tau]$ , in probability (for example, see Protter (1990)). Hence, the quadratic variation is a measure of the volatility of the trading costs. Moreover,  $\langle \tilde{C} \rangle_t$  has a simple expression, namely  $\int_0^t \tilde{Q}_s^T \Gamma^2 \tilde{Q}_s ds$ .

Unlike hitherto we allow now the trader to have an infinite horizon. After having defined  $\tilde{C}_\infty \triangleq \lim_{t \rightarrow \infty} \tilde{C}_t$  and  $\langle \tilde{C} \rangle_\infty \triangleq \lim_{t \rightarrow \infty} \langle \tilde{C} \rangle_t$ , let us investigate the problem

$$\tilde{L}(Q) \triangleq \inf_{\tilde{q}_t \in \hat{M}[\Theta_{[0, \infty)}(Q)]} E \left[ \tilde{C}_\infty + \frac{R}{2} \langle \tilde{C} \rangle_\infty \right] \quad (42)$$

subject to (37), where  $\hat{M}[\Theta_{[0, \infty)}(Q)]$  denotes the set of all  $\varphi - a.e.$  continuous  $M(H_t)$ -Markov controls  $\tilde{q}_t$  that satisfy  $\int_0^\infty \tilde{q}_t dt = Q$ , and  $R > 0$  is a risk-aversion coefficient. Since  $\tilde{C}_\infty + \frac{R}{2} \langle \tilde{C} \rangle_\infty = \lim_{t \rightarrow \infty} \tilde{C}_t + \frac{R}{2} \int_0^\infty \tilde{Q}_t^T \Gamma^2 \tilde{Q}_t dt$  the loss function  $V$  representing the objective in (42) is  $V(f, g) = \lim_{t \rightarrow \infty} f(t) + \frac{R}{2} \int_0^\infty g(t)^T \Gamma^2 g(t) dt$  for suitable  $f \in \mathbf{R}^{[0, \infty)}$  and  $g \in (\mathbf{R}^{[0, \infty)})^M$ . The liquidity trading problem in (42) implies that the trader cares about the instantaneous volatilities of the costs accumulated over the whole trading horizon. This contrasts the mean-variance objective where the risk-averse trader is concerned about the volatility of the total trading costs. Traders who face cash constraints during the trading period may compute (42). They appreciate a smooth dynamics of the trading expenditures, because they prefer to have smooth cash outflows.

**Proposition 8** *The optimal trading behavior of a risk-averse trader who solves (42) is*

$$\tilde{q}_t = \sqrt{\frac{R}{2}} \Lambda^{-1} \Gamma \Lambda^{1/2} \exp \left[ -\sqrt{\frac{R}{2}} \Lambda^{-1} \Gamma \Lambda^{1/2} t \right] Q, \quad (43)$$

$$\tilde{Q}_t = \exp \left[ -\sqrt{\frac{R}{2}} \Lambda^{-1} \Gamma \Lambda^{1/2} t \right] Q, \quad (44)$$

for  $t \in [0, \infty)$ , inducing the loss function

$$\tilde{L}(Q) = p_0^T Q + \frac{1}{2} Q^T [\Lambda + \sqrt{2R} \Gamma \Lambda^{1/2}] Q. \quad (45)$$

Hence, a risk-averse liquidity trader diminishes his trade sizes over time to minimize the mean and the total volatility of the trading costs. This perfectly reproduces the results given in Propositions 2 and 6. We thus arrive at the conclusion that optimal liquidity trading requires the same behavior in discrete and in continuous time.

## 7 Concluding Remarks

This paper studies the optimal behavior of a trader who wishes to buy (or sell) a given quantity of a security within a certain number of trading rounds. He is constrained to submit only market orders and his trades affect current and future prices of the security. He therefore breaks up his trades into a sequence of smaller orders. Risk neutrality implies that these smaller orders are equal. If the trader is risk averse, though, the magnitude of his trades declines over time.

On the theoretical level, this paper does not study the optimal policy of a trader who can submit limit orders in addition to market orders, nor does it study the circumstances under which the intertrading intervals are chosen in a continuous-time setup. It would be desirable to empirically estimate the price-impact function.

In addition, liquidity, which is measured by the price-impact slopes, may be random. One way to take into account the uncertainty of the market's liquidity is to model the price-impact slopes as stochastic process and study the optimal trading behavior in the same spirit as in this paper.

# Appendix

## A.1 Existence of a time-consistent Solution

We first establish a theorem that states a sufficient and a necessary condition for the existence of a solution to problem (3). We then characterize the absence of arbitrage as defined in Huberman and Stanzl (2000) for the price process (2). This characterization constitutes a generalization of Huberman and Stanzl's no-arbitrage condition for linear, time-dependent price-impact functions. The absence of arbitrage implies the existence of a time-consistent solution to (3) if  $R > 0$ . In the risk-neutral case, a slightly stronger condition than the absence of arbitrage can be imposed to guarantee the solvability of (3).

**Theorem 1** *A sequence of trades  $\{q_n\}_{n=1}^N$  is a time-consistent solution to the liquidity trader's problem (3) if and only if it solves the dynamic program in (8). If a solution exists, then it is unique and the  $N - 1$ -square matrix,*

$$\Lambda_N^\alpha + R [\Upsilon_N + (\alpha\lambda_1\sigma_\eta)^2 I_{N-1 \times N-1}] , \quad (46)$$

*is positive semidefinite, where*

$$[\Lambda_N^\alpha]_{n,m} \triangleq \begin{cases} \alpha\lambda_1 + 2\lambda_{n+1} & \text{if } n = m \\ (1 - \alpha)\lambda_{\min(n,m)+1} & \text{if } n \neq m \end{cases} , \quad (47)$$

$$[\Upsilon_N]_{n,m} \triangleq \begin{cases} \lambda_{n+1}^2\sigma_\eta^2 + (1 - \alpha)^2 \left[ \sum_{j=2}^n (\lambda_j^2\sigma_\eta^2 + \sigma_\varepsilon^2) + \sigma_\varepsilon^2 \right] & \text{if } n = m \\ (1 - \alpha) \left[ \lambda_{\min(n,m)+1}^2\sigma_\eta^2 + (1 - \alpha)\sigma_\varepsilon^2 + \sum_{j=2}^{\min(n,m)} (\lambda_j^2\sigma_\eta^2 + \sigma_\varepsilon^2) \right] & \text{if } n \neq m \end{cases} , \quad (48)$$

*for  $n, m = 1, \dots, N - 1$ , and  $I_{N-1 \times N-1}$  is the  $N - 1$  identity matrix. On the other hand, if the*

matrix given in (46) is positive definite, then a time-consistent solution to (3) always exists.

**Proof.** We start by proving that a solution to the dynamic program in (8),  $\{q_n, L_n\}_{n=1}^N$ , also represents a time-consistent solution to (3). Thus we need to show that  $\{q_n, L_n\}_{n=1}^N$ , satisfies the equations

$$L_n = \inf_{\{q_j \in \mathbf{M}(H_j)\}_{j=n}^N} E_n \left[ \sum_{j=n}^N p_j q_j \right] + \frac{R}{2} \text{Var}_n \left[ \sum_{j=n}^N p_j q_j \right] \quad \text{for } 1 \leq n \leq N-1. \quad (49)$$

This can be easily accomplished through backward induction. In the last period  $N$ , the trade  $q_N$  has to be chosen such that  $q_N = Q - \sum_{n=1}^{N-1} q_n$ . Hence, substitute  $L_N = E_N[p_N q_N] + \frac{R}{2} \text{Var}_N[p_N q_N]$  into

$$L_{N-1} = \inf_{q_{N-1} \in \mathbf{M}(H_{N-1})} E_{N-1} [p_{N-1} q_{N-1} + L_N] + \frac{R}{2} \text{Var}_{N-1} [p_{N-1} q_{N-1} + L_N]$$

to obtain

$$\begin{aligned} L_{N-1} &= \inf_{q_{N-1} \in \mathbf{M}(H_{N-1})} \left\{ E_{N-1} \left[ \sum_{n=N-1}^N p_n q_n \right] \right. \\ &\quad \left. + \frac{R}{2} \left[ \text{Var}_{N-1} \left[ E_N \left[ \sum_{n=N-1}^N p_n q_n \right] \right] + \text{Var}_N [p_N q_N] \right] \right\} \\ &= \inf_{q_{N-1} \in \mathbf{M}(H_{N-1})} E_{N-1} \left[ \sum_{n=N-1}^N p_n q_n \right] + \frac{R}{2} \text{Var}_{N-1} \left[ \sum_{n=N-1}^N p_n q_n \right] \end{aligned}$$

by taking into account the underlying price process (2). The solution to the dynamic program,  $\{q_n, L_n\}_{n=1}^N$ , therefore satisfies equation (49) for  $n = N-1$ .

Now, suppose that  $\{q_n, L_n\}_{n=1}^N$  meets the equations in (49) for  $n+1 \leq j \leq N$ . In this case  $\{q_n, L_n\}_{n=1}^N$  satisfies equation (49) for  $j = n$  as well, completing the induction argument. This can

be seen as follows. Note that

$$Var_n[\sum_{j=n}^N p_j q_j] = Var_n \left[ E_{n+1}[\sum_{j=n}^N p_j q_j] \right] + Var_{n+1}[\sum_{j=n+1}^N p_j q_j], \quad (50)$$

because  $q_j$  is a linear function of  $Q_j$  for all  $n+1 \leq j \leq N$ , as can be seen from the formulas (20)-(21). But this fact, together with the induction hypothesis, already implies

$$\begin{aligned} L_n &= \inf_{q_n \in \mathbf{M}(H_n)} E_n[p_n q_n + L_{n+1}] + \frac{R}{2} Var_n[p_n q_n + L_{n+1}] \\ &= \inf_{\{q_j \in \mathbf{M}(H_j)\}_{j=n}^N} \left\{ E_n[\sum_{j=n}^N p_j q_j] \right. \\ &\quad \left. + \frac{R}{2} \left[ Var_n \left[ E_{n+1}[\sum_{j=n}^N p_j q_j] \right] + Var_{n+1}[\sum_{j=n+1}^N p_j q_j] \right] \right\} \\ &= \inf_{\{q_j \in \mathbf{M}(H_j)\}_{j=n}^N} E_n[\sum_{j=n}^N p_j q_j] + \frac{R}{2} Var_n[\sum_{j=n}^N p_j q_j]. \end{aligned}$$

Proceeding in a similar manner as above, one readily derives that any time-consistent solution to (3) also constitutes a solution to the dynamic program in (8), proving the first assertion of this theorem.

Since the recursive solution (20)-(22) can only be iterated in a unique way, a time-consistent solution to (3) must be unique, too. Furthermore, given that a solution to (3) must be deterministic ( $q_n$  = linear function of  $Q_n$ ,  $1 \leq n \leq N$ ), (3) can be rewritten as

$$\begin{aligned} L(Q) &= \inf_{q_{N,-1} \in \mathbf{R}^{N-1}} \left\{ \left[ p_0 + \lambda_1 \left( 1 + \frac{R}{2} \lambda_1 \sigma_\eta^2 \right) Q \right] Q \right. \\ &\quad \left. - \lambda_1 \left[ (1 + \alpha) + R \alpha \lambda_1 \sigma_\eta^2 \right] Q 1_{N-1}^T q_{N,-1} \right\} \end{aligned}$$

$$+\frac{1}{2}q_{N,-1}^T [\Lambda_N^\alpha + R[\Upsilon_N + (\alpha\lambda_1\sigma_\eta)^2 I_{N-1 \times N-1}]] q_{N,-1} \Big\}, \quad (51)$$

where  $1_{N-1}$  is the  $N-1$ -dimensional vector containing only ones and  $q_{N,-1}^T \triangleq [q_2, \dots, q_N]$ . From (51) and static optimization theory it follows that the matrix given in (46) must be positive semidefinite, establishing the second claim of this theorem.

To show the last statement requires only to take the first- and second-order conditions in (51).

□

**Proposition 9** *The price process (2) is arbitrage-free if and only if the matrix  $\Lambda_N^\alpha$  is positive semidefinite for all  $N \in \mathbf{N}$ .*

**Proof.** If (2) is arbitrage-free, then  $\sum_{n=1}^N q_n = 0$  implies expected costs

$$E\left[\sum_{n=1}^N p_n q_n\right] = \frac{1}{2} q_{N,-1}^T \Lambda_N^\alpha q_{N,-1}$$

to be nonnegative for all  $q_{N,-1} \in \mathbf{R}^{N-1}$  and  $N \in \mathbf{N}$ , i.e.,  $\Lambda_N^\alpha$  has to be positive semidefinite for all  $N \in \mathbf{N}$ .

To prove the reverse, note that (3) and (51) for  $Q = R = 0$  imply

$$\inf_{\{q_n \in \mathbf{M}(H_n)\}_{n=1}^N} E\left[\sum_{n=1}^N p_n q_n\right] = \inf_{q_{N,-1} \in \mathbf{R}^{N-1}} \frac{1}{2} q_{N,-1}^T \Lambda_N^\alpha q_{N,-1}.$$

If all  $\{\Lambda_N^\alpha\}_{N=1}^\infty$  are positive semidefinite, then expected costs are globally minimized at  $q_n = 0$  for all  $1 \leq n \leq N$ , because the function  $q_{N,-1} \mapsto q_{N,-1}^T \Lambda_N^\alpha q_{N,-1}$  is convex on  $\mathbf{R}^{N-1}$  and  $\frac{d}{dq_{N,-1}} \left[ q_{N,-1}^T \Lambda_N^\alpha q_{N,-1} \right] (0) = 0$ . Hence, expected costs are always nonnegative if  $\sum_{n=1}^N q_n = 0$ , implying an arbitrage-free market. □

Since the matrix  $\Upsilon_N + (\alpha\lambda_1\sigma_\eta)^2 I_{N-1 \times N-1}$  is positive definite, due to the definition of the

variance, Theorem 1 and Proposition 9 have the following corollary as consequence.

**Corollary 2** *Suppose one of the following conditions holds:*

- i.  $R > 0$  (trader is risk-averse) and the price process (2) is arbitrage-free or*
- ii.  $R = 0$  (trader is risk-neutral) and the matrices  $\{\Lambda_N^\alpha\}_{N=1}^\infty$  are all positive definite.*

*The liquidity trader's problem (3) then has a unique, time-consistent solution for all  $N \in \mathbf{N}$  that can be obtained by solving the dynamic program in (8).*

## A.2 Proofs of Section 2

**Proof of Proposition 2.** Writing out the first-order conditions of (3) with price process (1) yields

$$\lambda q_{n+2} - (2\lambda + R\sigma^2)q_{n+1} + \lambda q_n = 0 \quad \text{for } 1 \leq n \leq N-3 \quad (52)$$

$$(\lambda^2 + 3\lambda R\sigma^2 + R^2\sigma^4)q_{N-1} - \lambda(\lambda + R\sigma^2)q_{N-2} = 0 \quad (53)$$

$$(3\lambda + R\sigma^2)q_{N-1} + \lambda \sum_{n=1}^{N-3} q_n = \lambda Q, \quad (54)$$

where  $\sigma^2 \triangleq \lambda^2 \sigma_\eta^2 + \sigma_\varepsilon^2$ .

Solving the difference equation (52) subject to the boundary conditions (53) and (54) gives (14) and (15).

The proof that the optimal trades are positive and strictly decreasing can be easily verified by looking directly at the formulas in (14) and (15).  $\square$

**Proof of Proposition 3.** The proof of this proposition uses only the principle of induction.

In period  $N$ , the optimal loss function, as a function of the two state variables  $\tilde{p}_{N-1}$  and  $Q_N$ , is given by

$$\begin{aligned} L_N(\tilde{p}_{N-1}, Q_N) &= E_N[p_N q_N] + \frac{R}{2} \text{Var}_N[p_N q_N] \\ &= E_N[(\tilde{p}_{N-1} + \lambda_N(Q_N + \eta_N))Q_N] + \frac{R}{2} \text{Var}_N[\lambda_N \eta_N Q_N] \\ &= \tilde{p}_{N-1} Q_N + \lambda_N \left(1 + \frac{R}{2} \lambda_N \sigma_\eta^2\right) Q_N^2, \end{aligned}$$

since in the last period  $q_N = Q_N$  must be traded. Let us define  $\mu_N \triangleq \lambda_N(1 + \frac{R}{2} \lambda_N \sigma_\eta^2)$ .

Next, suppose that (16)-(18) hold for  $n+1$ . We will show that this is also true for  $n$ , completing

the induction argument. Keeping the price process (1) in mind, one can write

$$L_n(\tilde{p}_{n-1}, Q_n) = \inf_{q_n} E_n[(\tilde{p}_{n-1} + \lambda_n(q_n + \eta_n))q_n + L_{n+1}(\tilde{p}_n, Q_{n+1})] \\ + \frac{R}{2} \text{Var}_n[(\tilde{p}_{n-1} + \lambda_n(q_n + \eta_n))q_n + L_{n+1}(\tilde{p}_n, Q_{n+1})]$$

Using the induction hypothesis, this can be rewritten as

$$L_n(\tilde{p}_{n-1}, Q_n) = \inf_{q_n} (\tilde{p}_{n-1} + \lambda_n q_n)Q_n + (\mu_{n+1} + \frac{R}{2}\sigma_\varepsilon^2)(Q_n - q_n)^2 + \frac{R}{2}\sigma_\eta^2\lambda_n^2 Q_n^2.$$

Taking the first-order condition for this expression yields (16). With (16), one can compute  $L_n(\tilde{p}_{n-1}, Q_n)$  to be

$$\tilde{p}_{n-1}Q_n + \lambda_n \left[ 1 + \frac{R}{2}\lambda_n\sigma_\eta^2 - \frac{\lambda_n}{2(2\mu_{n+1} + R\sigma_\varepsilon^2)} \right] Q_n^2,$$

which completes the proof.  $\square$

### A.3 Numerical Analysis: Optimal Trading Frequency

To begin, the model parameters are sliced into equidistant plausible intervals specified below, and the loss function is then evaluated at all endpoints. We proceed by looking at first differences of the loss function to verify its convexity. Specifically, for  $R$  the interval  $[0.001, 0.05]$  is sliced into ten subintervals of the same length, and similarly for  $\lambda$  the interval  $[10^{-6}, 10^{-5}]$  is divided into 100 sub-intervals. The variance functions take the forms  $\sigma_\eta^2(N) = \frac{\sigma_{\eta\tau}^2}{N}$  and  $\sigma_\varepsilon^2(N) = \frac{\sigma_{\varepsilon\tau}^2}{N}$ , where for  $\sigma_{\eta\tau}^2$  the interval  $[500, 5000]$  is chosen and for  $\sigma_{\varepsilon\tau}^2$  the interval is  $[0.01, 0.05]$ , both split into 20 sub-intervals of same length. In all cases, we look at 130 trading rounds, equivalent to trading every 15 minutes in a week at the NYSE (the NYSE is open six and a half hours a day, five days a week; hence trading every 15 minutes produces 130 transactions). One week is a reasonable time horizon for a liquidity trader (see Bertsimas and Lo (1998)).

For all computations, we find local convexity of the loss function for the periods under consideration, thereby ensuring a local unique solution for the risk-averse trader. Therefore, the numerical analysis is a powerful practical tool to determine the optimal number of trades.

To illustrate the sensitivity of the loss function and the optimal frequency of trades to the underlying parameters, we do some comparative statics, shown in Figures 6a-6c. The trading time is again sliced into 130 periods. The block to be traded consists of 10,000 shares of the asset and the initial price is \$40.

Figure 6a shows the loss function for different levels of risk aversion. The optimal frequency of trades is decreasing in the level of risk aversion,  $R$ . In the example depicted,  $N^* = 26$  for  $R = 0.001$ ,  $N^* = 70$  for  $R = 0.01$ , and  $N^* = 128$  for  $R = 0.05$ . Risk aversion has thus a substantial impact on the frequency with which traders submit their orders. Also note here how the risk coefficient affects the curvature of the loss function: it becomes more convex as the level of risk aversion declines.

The parameter  $\lambda$  is varied in Figure 6b. As in the risk-neutral case,  $N^*$  is increasing in  $\lambda$ . In particular, for  $\lambda = 10^{-5}$ ,  $N^* = 27$ ; for  $\lambda = 5 * 10^{-5}$ ,  $N^* = 60$ ; and for  $\lambda = 10^{-4}$ ,  $N^* = 80$ . The loss function gets more convex the lower is the slope  $\lambda$ .

Finally, Figure 6c shows the impact of the fixed costs on the optimal number of trades. Not surprisingly,  $N^*$  is decreasing in the transaction costs. In numbers, for a commission fee  $k = 5$ ,  $N^* = 89$ ; for  $k = 15$ ,  $N^* = 48$ ; and  $N^* = 36$  is chosen if  $k = 25$ .

For  $\alpha > 0$ , the same qualitative results are obtained, and so this case is not analyzed.

## A.4 Proofs of Section 5

**Proof of Corollary 1.** This proof makes use of the following facts:  $r_+ = r_-^{-1}$ ,  $r_+ \rightarrow 1$ ,  $r_+^N \rightarrow \infty$ , and  $A_{\pm} \rightarrow 0$  as  $N \rightarrow \infty$ . They are applied below without explicit reference.

To show that  $q_{\lfloor \frac{t}{\tau} N \rfloor} \rightarrow 0$ , for all  $t \in [0, \tau]$ , due to Proposition 2, it suffices to prove that  $q_1 \rightarrow 0$  as  $N \rightarrow \infty$ .

For this purpose write

$$q_1 = \frac{[A_+ - A_- r_-^{2(N-3)}]}{W_N + X_N} Q,$$

where

$$W_N \triangleq \frac{1 - r_-^{N-3}}{1 - r_-} A_+ - \frac{r_-^{2(N-3)} - r_-^{N-3}}{1 - r_+} A_-$$

and

$$X_N \triangleq r_-^{N-3} \left( 3\lambda + R \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau} \right) \left( \lambda + R \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau} \right) (r_+ - r_-).$$

Obviously, the numerator of  $q_1$ ,  $A_+ - A_- r_-^{2(N-3)}$ , and  $X_N$  both converge to zero. On the other hand,  $W_N$  converges to  $\lambda^2$ , because

$$\begin{aligned} \frac{1 - r_-^{N-3}}{1 - r_-} A_+ &= (1 - r_-^{N-3}) \\ &* \left[ \lambda^2 r_+ + \lambda R (3r_+ - 1) \frac{\frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau}}{1 - r_-} + R^2 r_+ \frac{\left( \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N^2/\tau^2} \right)^2}{1 - r_-} \right] \rightarrow \lambda^2 \end{aligned}$$

and

$$\frac{r_-^{2(N-3)} - r_-^{N-3}}{1 - r_+} = \frac{r_-^{N-2} (1 - r_-^{N-3})}{1 - r_-} \leq \frac{r_-^{N-2}}{1 - r_-} \rightarrow 0,$$

as  $N \rightarrow \infty$ . This proves the first fact in Corollary 1. Another way to see this is to consider the formulas (16) and (17). Since the solution to the Riccati equation in (17)

$$\mu_n = \frac{1 - \left[ 2 + R\lambda \frac{\sigma_{\eta\tau}^2}{N/\tau} \right] r_- + r_+^{2(N-n)} [(2 + R\lambda \frac{\sigma_{\eta\tau}^2}{N/\tau}) r_+ - 1]}{\frac{2}{\lambda} r_- [r_+^{2(N-n+1)} - 1]} \quad (55)$$

converges to  $\frac{\lambda}{2}$  as  $N \rightarrow \infty$ ,  $q_1 \rightarrow 0$  in (16).

The second assertion is proved in a similar way and can be left to the reader.  $\square$

**Proof of Proposition 6.** Let us begin by solving problem (33). The necessary conditions for a minimum require  $\tilde{Q}_t$  to meet

$$2\theta \frac{d^2 \tilde{Q}_t}{dt^2} = R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2) \tilde{Q}_t \quad \text{with } \tilde{Q}_0 = Q \text{ and } \tilde{Q}_\tau = 0. \quad (56)$$

It is easily checked that this boundary-value problem has the solution

$$\tilde{Q}_t = \frac{\sinh \left[ \sqrt{\frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta}} (\tau - t) \right]}{\sinh \left( \sqrt{\frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta}} \tau \right)} Q \quad \text{for } t \in [0, \tau]. \quad (57)$$

Clearly, the function  $\tilde{Q}_t$  given in (57) is an element of  $\Theta_{[0,\tau]}(Q)$ .

Next, we tackle the discrete-time problem (3) for the price dynamics (32). The loss function,  $l(C_N)$ , for this case becomes

$$p_0 Q + \lambda \sum_{n=1}^N Q_n (Q_n - Q_{n+1}) + \frac{R}{2} \frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N} \sum_{n=1}^N Q_n^2 + \frac{N\theta}{\tau} \sum_{n=1}^N (Q_n - Q_{n+1})^2. \quad (58)$$

Minimizing  $l(C_N)$  with respect to the sequence  $\{Q_n\}_{n=2}^N$  yields the first-order conditions

$$Q_{n+1} - 2Q_n + Q_{n-1} = \frac{\tau^2}{N^2} \frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta + \lambda \frac{\tau}{N}} Q_n \quad \text{for } 1 \leq n \leq N+1, \quad (59)$$

$Q_1 = Q$ , and  $Q_{N+1} = 0$ . Some algebra reveals that the solution to the difference equation (59) has the form

$$Q_n = \frac{\sinh \left[ \psi_N \tau \left( 1 + \frac{1-n}{N} \right) \right]}{\sinh(\psi_N \tau)} Q \quad \text{for } 1 \leq n \leq N, \quad (60)$$

where

$$\cosh \left( \frac{\tau}{N} \psi_N \right) = 1 + \frac{\tau^2}{2N^2} \frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta + \lambda \frac{\tau}{N}}. \quad (61)$$

Hence, as  $q_n = Q_n - Q_{n+1}$  by definition, the formulas in (34) and (35) constitute the solution to problem (3) under the price rule (32).

We are now prepared to show that both  $\lim_{N \rightarrow \infty} Q_{\lfloor \frac{t}{\tau} N \rfloor}$  and  $\lim_{N \rightarrow \infty} \frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N}$  exist, and that they equal  $\tilde{Q}_t$  and  $-\frac{d\tilde{Q}_t}{dt}$ , respectively, where  $\tilde{Q}_t$  is given as in (57). As a first step, we study the convergence behavior of  $\psi_N$ . The Taylor expansion

$$\cosh\left(\frac{\tau}{N}\psi_N\right) = 1 + \sum_{n=1}^{\infty} \frac{1}{(2n)!} \left(\frac{\tau}{N}\right)^n \psi_N^n \quad (62)$$

indicates that for large  $N$  (neglecting terms with degree four or larger) equation (61) becomes

$$\frac{1}{2} \left(\frac{\tau}{N}\right)^2 \psi_N^2 = \frac{\tau^2}{2N^2} \frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta + \lambda \frac{\tau}{N}}.$$

This, in turn, has

$$\psi_N \rightarrow \sqrt{\frac{R(\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2)}{2\theta}} \quad (63)$$

for  $N \rightarrow \infty$  as a consequence. By virtue of (57) and (60),  $\tilde{Q}_t = \lim_{N \rightarrow \infty} Q_{\lfloor \frac{t}{\tau} N \rfloor}$ .

To find  $\lim_{N \rightarrow \infty} \frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N}$ , use (34) to calculate

$$\frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N} = \frac{\psi_N \cosh(\xi_N)}{\sinh(\psi_N \tau)} Q$$

for  $\xi_N \in \left(\psi_N \tau \left(1 - \frac{\lfloor \frac{t}{\tau} N \rfloor}{N}\right), \psi_N \tau \left(1 + \frac{1}{N} - \frac{\lfloor \frac{t}{\tau} N \rfloor}{N}\right)\right)$ . Clearly,  $\frac{q_{\lfloor \frac{t}{\tau} N \rfloor}}{\tau/N}$  converges to the expression  $\tilde{q}_t$  given in (36), which is the negative of the first derivative of  $\tilde{Q}_t$  in (57).

Thus,  $\tilde{l}_{[0,\tau]}(Q) = \lim_{N \rightarrow \infty} L(Q, N)$  remains to be proved. Employing  $Q_n$  and  $\tilde{Q}_t$  as in (60) and (57), respectively, it follows that

$$\sum_{n=1}^N Q_n (Q_n - Q_{n+1}) \rightarrow - \int_0^\tau \tilde{Q}_t d\tilde{Q}_t = \frac{Q^2}{2}, \quad (64)$$

$$\frac{\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2}{N/\tau} \sum_{n=1}^N Q_n^2 = (\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2) \sum_{n=1}^N Q_n^2 \frac{\tau}{N} \rightarrow (\lambda^2 \sigma_{\eta\tau}^2 + \sigma_{\varepsilon\tau}^2) \int_0^\tau \tilde{Q}_t^2 dt, \quad (65)$$

and

$$\frac{N\theta}{\tau} \sum_{n=1}^N (Q_n - Q_{n+1})^2 = \theta \sum_{n=1}^N \frac{Q_{n+1} - Q_n}{\tau/N} (Q_{n+1} - Q_n) \rightarrow \theta \int_0^\tau \left( \frac{d\tilde{Q}_t}{dt} \right)^2 dt \quad (66)$$

as  $N \rightarrow \infty$ . But this shows that  $L(Q, N) \rightarrow \tilde{l}_{[0,\tau]}(Q)$  and the proof is complete.  $\square$

## A.5 Proofs of Section 6

**Proof of Proposition 7.** To solve (39) for  $V(f, g) = f(\tau)$  we only need to consider the problem

$$\tilde{L}(Q) = \inf_{\tilde{q}_t \in \tilde{M}[\Theta_{[0,\tau]}(Q)]} E \left[ \int_0^\tau \tilde{q}_t^T \Lambda \tilde{q}_t dt \right] \quad (67)$$

subject to (37). This is due to the constraint  $\tilde{Q}_\tau = 0$  and the equality  $E \left[ \int_0^\tau \tilde{Q}_\tau^T \Gamma dB_t \right] = 0$ . The solution of (67) can be obtained by applying the following Lagrange approach. Solve for all  $\kappa \in \mathbf{R}^M$

$$\tilde{L}(Q, \kappa) = \inf_{\substack{\tilde{q}_t \text{ is continuous} \\ M(H_t)\text{-Markov control}}} E \left[ \int_0^\tau \tilde{q}_t^T \Lambda \tilde{q}_t dt + \kappa^T \tilde{Q}_\tau \right] \quad (68)$$

subject to (37). Note that the optimal policy of (68) may not satisfy the constraint  $\int_0^\tau \tilde{q}_t dt = Q$ . However, any deviation from this constraint would alter the objective  $\tilde{L}(Q, \kappa)$  through the term  $\kappa^T \tilde{Q}_\tau$ , where  $\kappa$  has the interpretation of a Lagrange multiplier. If we are able to determine a vector  $\kappa^*$  such that the solution  $\tilde{q}_t(\kappa^*)$  of (68) satisfies  $\int_0^\tau \tilde{q}_t(\kappa^*) dt = Q$ , then we have also found a solution of (67).

The Hamilton-Jacobi-Bellman equation for (68) entails

$$\inf_{\tilde{q}_s} \left[ (1 + \phi_{\tilde{C}}) \tilde{q}_s^T \Lambda \tilde{q}_s + \phi_t - \phi_{\tilde{Q}}^T \tilde{q}_s + \frac{1}{2} \tilde{Q}_s^T \Gamma^2 \tilde{Q}_s \phi_{\tilde{C}\tilde{C}} \right] = 0, \quad (69)$$

$s \in [0, \tau]$ , for a smooth candidate function  $\phi : [0, \tau] \times \mathbf{R}^M \times \mathbf{R} \rightarrow \mathbf{R}$ , where  $\phi_t$ ,  $\phi_{\tilde{C}}$ ,  $\phi_{\tilde{C}\tilde{C}}$ , and

$\phi_{\tilde{Q}} \triangleq (\phi_{\tilde{Q}_1}, \phi_{\tilde{Q}_2}, \dots, \phi_{\tilde{Q}_M})$  denote the partial derivatives of  $\phi$  with respect to time and the state variables, respectively. By virtue of (69), the optimal trading rate is given by

$$\tilde{q}_s = \frac{1}{2(1 + \phi_{\tilde{C}})} \Lambda^{-1} \phi_{\tilde{Q}}. \quad (70)$$

Hence, by substituting (70) into (69), we obtain that the candidate function  $\phi$  must solve the boundary value problem

$$\phi_t - \frac{1}{4(1 + \phi_{\tilde{C}})} \phi_{\tilde{Q}}^T \Lambda \phi_{\tilde{Q}} + \frac{1}{2} \tilde{Q}_s^T \Gamma^2 \tilde{Q}_s \phi_{\tilde{C}\tilde{C}} = 0 \quad (71)$$

with

$$\phi(\tau, \tilde{Q}, \tilde{C}) = \kappa^T \tilde{Q} \quad \text{for all } \tilde{Q} \in \mathbf{R}^M \text{ and } \tilde{C} \in \mathbf{R}.$$

It can be easily verified that  $\phi(t, \tilde{Q}, \tilde{C}) = \frac{1}{4} \kappa^T \Lambda^{-1} \kappa (t - \tau) + \kappa^T \tilde{Q}$  solves (71) and therefore  $\tilde{q}_s(\kappa) = \frac{1}{2} \Lambda^{-1} \kappa$ . As a consequence, the unique  $\kappa^*$  that implies  $\int_0^\tau \tilde{q}_t(\kappa^*) dt = Q$  equals  $\kappa^* = \frac{2}{\tau} \Lambda Q$ . The solution of (39) for  $V(f, g) = f(\tau)$  is thus found:  $\tilde{q}_s(\kappa^*) = \frac{1}{\tau} Q$  as stated in Proposition 7. The value function  $\tilde{L}(Q)$  can be calculated from (38) and the equality  $\phi^*(0, Q, 0) = \frac{1}{\tau} Q^T \Lambda Q$ , which completes the proof.  $\square$

**Proof of Proposition 8.** Consider the problem

$$\tilde{L}(Q) \triangleq \inf_{\tilde{q}_t \in \hat{M}[\Theta_{[0, \infty)}(Q)]} E \left[ \int_0^\infty [\tilde{q}_t^T \Lambda \tilde{q}_t + \frac{R}{2} \tilde{Q}_t^T \Gamma^2 \tilde{Q}_t] dt \right]. \quad (72)$$

It is easy to see that a solution of (72) also solves (42). So we only need to study here (72). From

its Hamilton-Jacobi-Bellman equation

$$\inf_{\tilde{q}_s} \left[ (1 + \phi_{\tilde{C}}) \tilde{q}_s^T \Lambda \tilde{q}_s + \phi_t - \phi_{\tilde{Q}}^T \tilde{q}_s + \frac{1}{2} (R + \phi_{\tilde{C}\tilde{C}}) \tilde{Q}_s^T \Gamma^2 \tilde{Q}_s \right] = 0 \quad (73)$$

we derive that a solution  $\phi : [0, \infty) \times \mathbf{R}^M \times \mathbf{R} \rightarrow \mathbf{R}$  has to satisfy

$$\phi_t - \frac{1}{4(1 + \phi_{\tilde{C}})} \phi_{\tilde{Q}}^T \Lambda \phi_{\tilde{Q}} + \frac{1}{2} (R + \phi_{\tilde{C}\tilde{C}}) \tilde{Q}_s^T \Gamma^2 \tilde{Q}_s = 0 \quad (74)$$

and

$$\lim_{t \rightarrow \infty} \phi(t, \tilde{Q}_t, \tilde{C}_t) = 0.$$

Straightforward computations show that  $\tilde{q}_t$  as given in (70) solves (73) and that the function  $\phi(t, \tilde{Q}, \tilde{C}) = \sqrt{\frac{R}{2}} \tilde{Q}^T \Gamma \Lambda^{1/2} \tilde{Q}$  meets both equations in (74). In view of the state equations  $d\tilde{Q}_t = -\tilde{q}_t dt$ , we conclude that

$$\frac{d\tilde{Q}_t}{dt} = -\sqrt{\frac{R}{2}} \tilde{Q}^T \Gamma \Lambda^{1/2} \tilde{Q}_t.$$

The solution of this differential equation system yields at once (43) and (44). Finally, note that the loss function in (45) follows from  $\phi(0, Q, 0) = \sqrt{\frac{R}{2}} Q^T \Gamma \Lambda^{1/2} Q$ .  $\square$

## References

- Back, K., 1992, Insider Trading in Continuous Time, *Review of Financial Studies*, Vol.5, No.3, 387-409.
- Back, K., and H. Pedersen, 1998, Long-lived Information and Intraday Patterns, *Journal of Financial Markets*, Vol.1, Issue 3-4, 385-402.
- Beneš, V.A., A. Shepp, and H.S. Witsenhausen, 1980, Some Solvable Stochastic Control Problems, *Stochastics* 4, 39-83.
- Bertsimas, D., and A.W. Lo, 1998, Optimal Control of Execution Costs, *Journal of Financial Markets* 1, 1-50.
- Chordia, T., R. Roll, and A. Subrahmanyam, Market Liquidity and Trading Activity, *Working Paper*, Andersen School UCLA.
- Cvitanic, J., and I. Karatzas, 1992, Convex Duality in Constrained Portfolio Optimization, *Annals of Applied Probability*, Vol.2, No.4, 767-818.
- Dutta, P.K., and A. Madhavan, 1995, Price Continuity Rules and Insider Trading, *Journal of Financial and Quantitative Analysis*, Vol.30, No.2, 199-221.
- Easley, D., and M. O'Hara, 1992, Time and the Process of Security Price Adjustment, *Journal of Finance*, Vol.47, No.2, 577-605.
- Hausman, J.A., A.W. Lo, and A.C. MacKinlay, 1992, An Ordered Probit Analysis of Transaction Stock Prices, *Journal of Financial Economics*, Vol.31, 319-379.
- Huberman, G., and W. Stanzl, 2000, Arbitrage-free Price-Update and Price-Impact Functions, *Mimeo*, Yale School of Management.

- Karatzas, I., and S.E. Shreve, 1986, Equivalent Models for Finite-Fuel Stochastic Control, *Stochastics*, Vol.18, 245-276.
- Karatzas, I., D. Ocone, H. Wang, and M. Zervos, 2000, Finite-Fuel Singular Control with Discretionary Stopping, forthcoming in *Stochastics*.
- Keim, D.B., and A. Madhavan, 1996, The Upstairs Market for Large Block Transactions: Analysis and Measurement of Price Effects, *Review of Financial Studies*, Vol.9, 1-36.
- Kempf, A., and O. Korn, 1999, Market Depth and Order Size, *Journal of Financial Markets*, Vol.2, No.1, 29-48.
- Kyle, A.S., 1985, Continuous Auctions and Insider Trading, *Econometrica*, Vol.53, No.6, 1315-1336.
- Merton, R.C., 1990, Continuous-time Finance, *Blackwell Publishers*.
- Musiela, M., and M. Rutkowski, 1997, Martingale Methods in Financial Modelling, *Applications of Mathematics*, Springer Verlag.
- Lo, A.W., and A.C. MacKinlay, 1990, When are Contrarian Profits Due to Stock Market Overreaction?, *Review of Financial Studies*, Vol.3, 175-206.
- Oksendal, B., 1995, Stochastic Differential Equations, *Springer Verlag*, Chapter 11.
- Protter, P., 1990, Stochastic Integration and Differential Equations, *Springer Verlag*.
- Rao, M.M., 1984, Probability Theory with Applications, *Academic Press*.

## List of figures

Figure 1a: Optimal trading volume with different risk aversion

Figure 1b: Optimal change in price with different risk aversion

Figure 2a: Optimal trading volume with different price volatilities

Figure 2b: Optimal change in price with different price volatilities

Figure 3a: Optimal trading volume with different constant  $\lambda$ 's

Figure 3b: Optimal change in price with different constant  $\lambda$ 's

Figure 4a: Optimal trading volume with time-dependent  $\lambda_n$ 's

Figure 4b: Optimal change in price with time-dependent  $\lambda_n$ 's

Figure 5a: Optimal trading volume with different updating weights

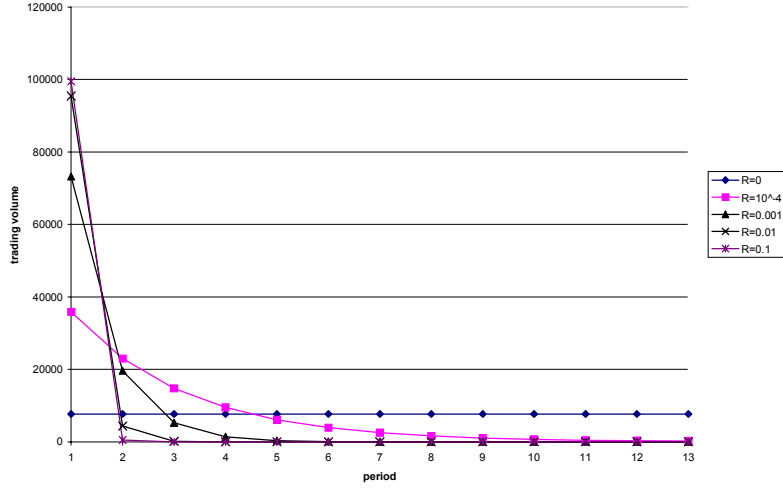
Figure 5b: Optimal change in price with different updating weights

Figure 6a: The loss function with transaction costs depending on risk aversion

Figure 6b: The loss function with transaction costs depending on the  $\lambda$ 's

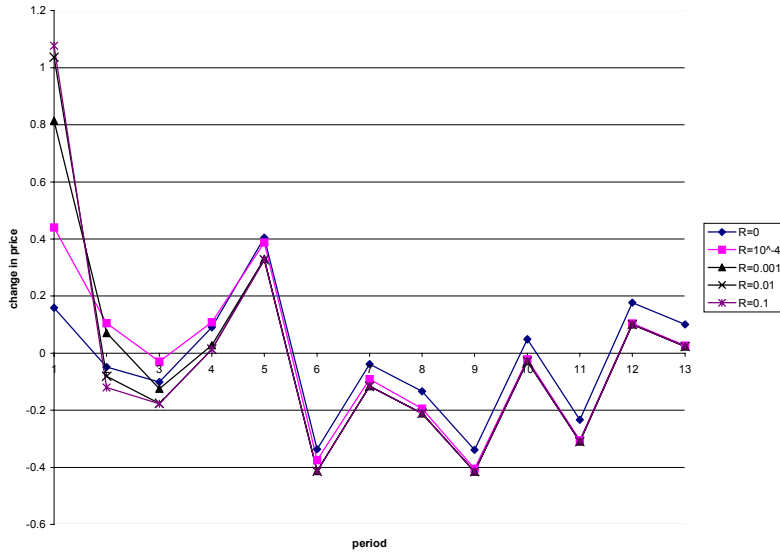
Figure 6c: The loss function depending on fixed transaction costs

# Figures



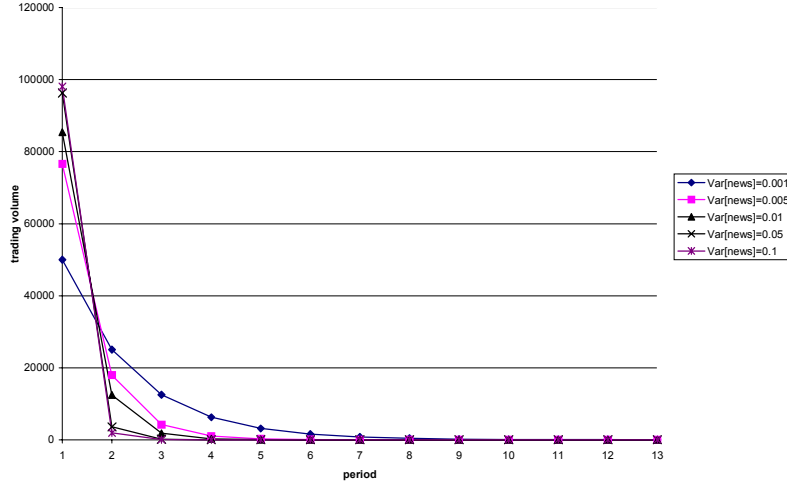
**Figure 1a:** Optimal trading volume with different risk aversion. Simulation values:  $N = 13$ ,

$$p_0 = 40, Q = 100,000, \lambda = 10^{-5}, \sigma_\varepsilon^2 = 0.02, \sigma_\eta^2 = 1000.$$

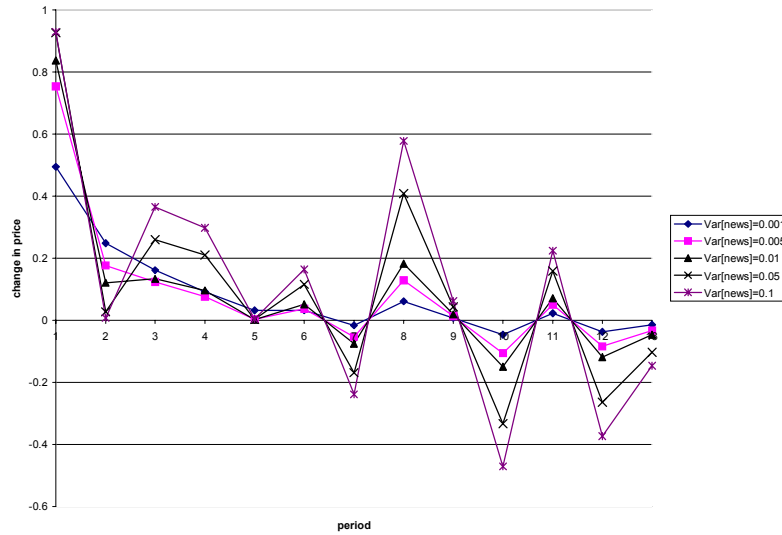


**Figure 1b:** Optimal change in price with different risk aversion. Simulation values:  $N = 13$ ,

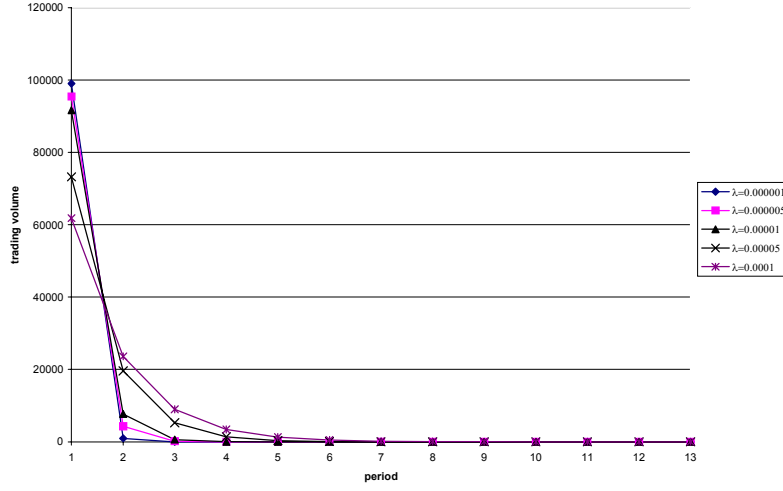
$$p_0 = 40, Q = 100,000, \lambda = 10^{-5}, \sigma_\varepsilon^2 = 0.02, \sigma_\eta^2 = 1000.$$



**Figure 2a:** Optimal trading volume with different price volatilities. Simulation values:  $N = 13$ ,  $p_0 = 40$ ,  $Q = 100,000$ ,  $\lambda = 10^{-5}$ ,  $R = 0.005$ ,  $\sigma_\eta^2 = 1000$ ; the legen box shows the different values used for the variance of  $\varepsilon$ .

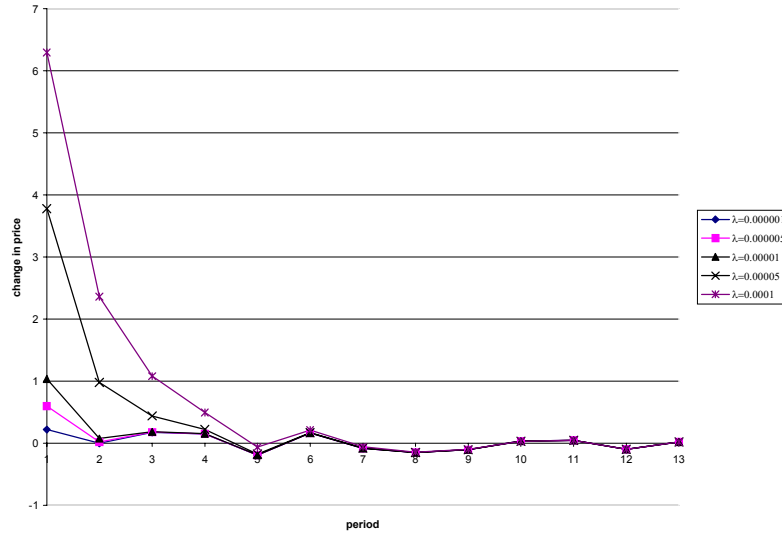


**Figure 2b:** Optimal change in price with different price volatilities. Simulation values:  $N = 13$ ,  $p_0 = 40$ ,  $Q = 100,000$ ,  $\lambda = 10^{-5}$ ,  $R = 0.005$ ,  $\sigma_\eta^2 = 1000$ ; the legend box shows the different values used for the variance of  $\varepsilon$ .



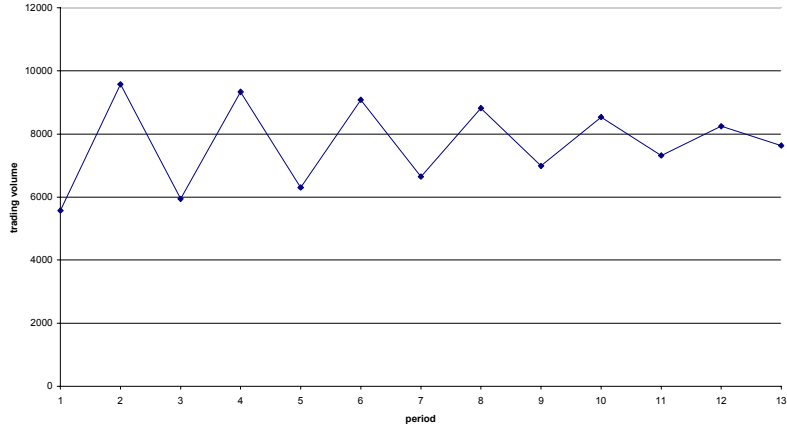
**Figure 3a:** Optimal trading volume with different constant  $\lambda$ 's. Simulation values:  $N = 13$ ,

$$p_0 = 40, Q = 100,000, R = 0.005, \sigma_\varepsilon^2 = 0.02, \sigma_\eta^2 = 1000.$$

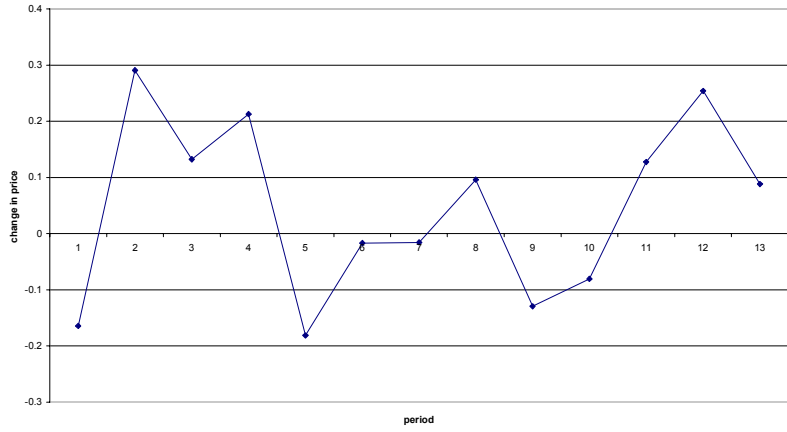


**Figure 3b:** Optimal change in price with different constant  $\lambda$ 's. Simulation values:  $N = 13$ ,

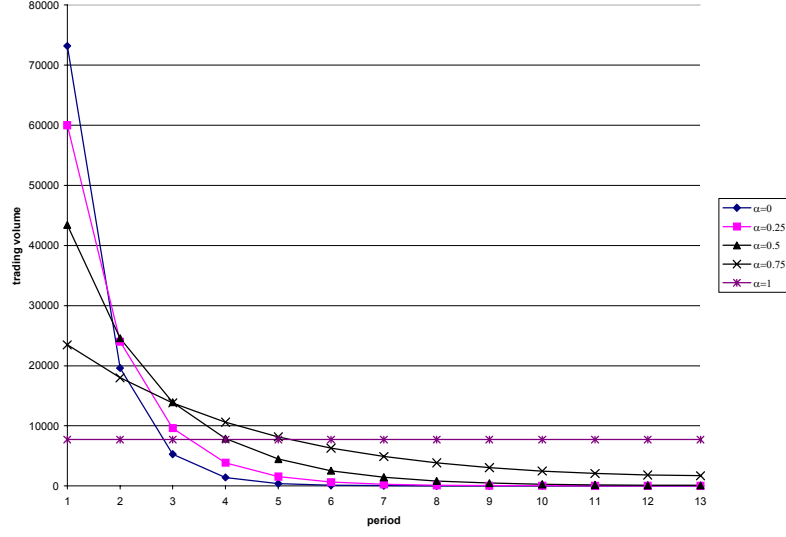
$$p_0 = 40, Q = 100,000, R = 0.005, \sigma_\varepsilon^2 = 0.02, \sigma_\eta^2 = 1000.$$



**Figure 4a:** Optimal trading volume with time-dependent  $\lambda_n$ 's. Simulation values:  $N = 13$ ,  $p_0 = 40$ ,  $Q = 100,000$ ,  $\lambda_n = 5.2 * 10^{-6}$  in odd periods,  $= 5 * 10^{-6}$  in even periods,  $\sigma_\varepsilon^2 = 0.02$ ,  $\sigma_\eta^2 = 1000$ .

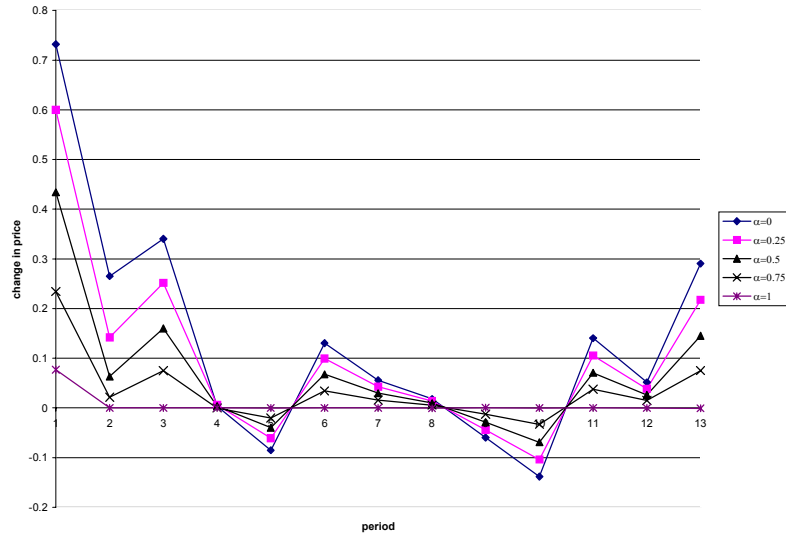


**Figure 4b:** Optimal change in price with time-dependent  $\lambda_n$ 's. Simulation values:  $N = 13$ ,  $p_0 = 40$ ,  $Q = 100,000$ ,  $\lambda_n = 5.2 * 10^{-6}$  in odd periods,  $= 5 * 10^{-6}$  in even periods,  $\sigma_\varepsilon^2 = 0.02$ ,  $\sigma_\eta^2 = 1000$ .



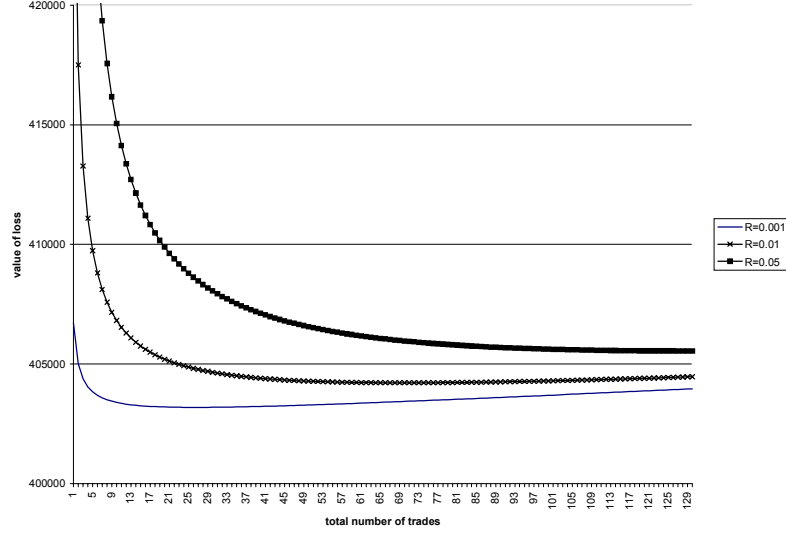
**Figure 5a:** Optimal trading volume with different updating weights. Simulation values:  $N = 13$ ,

$$p_0 = 40, Q = 100,000, \lambda = 10^{-5}, \sigma_\varepsilon^2 = 0.02, \sigma_\eta^2 = 1000.$$

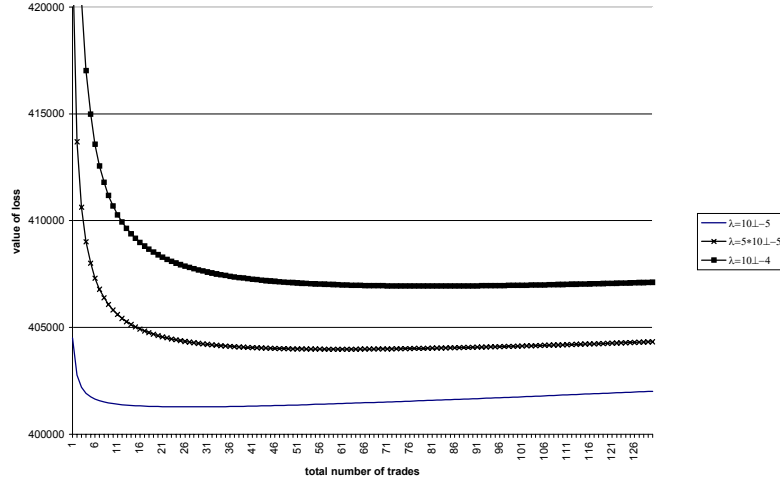


**Figure 5b:** Optimal change in price with different updating weights. Simulation values:  $N = 13$ ,

$$p_0 = 40, Q = 100,000, \lambda = 10^{-5}, \sigma_\varepsilon^2 = 0.02, \sigma_\eta^2 = 1000.$$

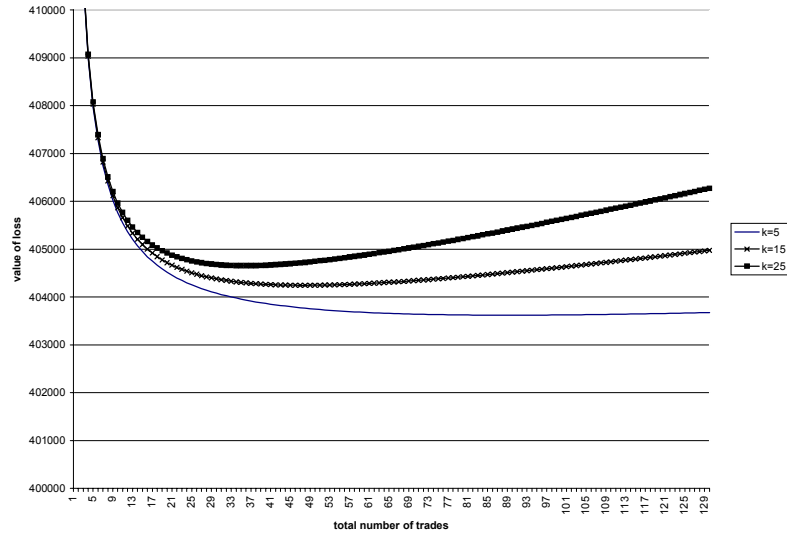


**Figure 6a:** The loss function with transaction costs depending on risk aversion. Simulation values:  $N = 130$ ,  $p_0 = 40$ ,  $Q = 10,000$ ,  $\lambda = 5 * 10^{-5}$ ,  $k = 10$ ,  $\sigma_\eta^2(N) = 1000/N$ ,  $\sigma_\varepsilon^2(N) = 0.02/N$ .



**Figure 6b:** The loss function with transaction costs depending on the  $\lambda$ 's. Simulation values:

$$N = 130, p_0 = 40, Q = 10,000, R = 0.01, k = 10, \sigma_\eta^2(N) = 1000/N, \sigma_\varepsilon^2(N) = 0.02/N.$$



**Figure 6c:** The loss function depending on fixed transaction costs. Simulation values:  $N = 130$ ,

$$p_0 = 40, Q = 10,000, R = 0.01, \lambda = 5 * 10^{-5}, \sigma_{\eta}^2(N) = 1000/N, \sigma_{\varepsilon}^2(N) = 0.02/N.$$