A Model of Asset Pricing with Market Impact Costs and Transactions Costs

Ming Guo and Hui Ou-Yang*

March 30, 2005

Rough Draft

*Guo is with the Department of Economics, Duke University, Durham, NC 27708. E-mail: mg22@duke.edu. Ou-Yang is with the Fuqua School of Business, Duke University, Durham, NC 27708-0120. E-mail: huiou@duke.edu. We thank Pete Kyle and seminar participants at Duke for their comments.

Abstract

This paper develops an equilibrium model of asset pricing in which the investor incurs both fixed and proportional costs in trading and must consider the impact of his trades on the stock price. The market maker, who sets the stock price competitively, also incurs both fixed costs and proportional costs in processing the investor's trade. A nonlinear equilibrium stock price is obtained. In particular, the stock price depends on the investor's order flow, on the fixed cost multiplied by the inverse of the investor's order flow, as well as on the proportional cost multiplied by the signed trade size (positive for purchases and negative for sales). The trade size increases when the fixed cost increases and it decreases when the proportional cost and market impact cost increase. It is demonstrated that over a short period of time, positive autocorrelations between the investor's order flows as well as between the stock returns arise whereas over a longer period of time, negative autocorrelations arise. Using the ticker-by-ticker trades from Trades and Automated Quotes (TAQ) database for the time period between 1993 and 2002, we test the nonlinear pricing function as well as estimate the three cost components. Our empirical results confirm the nonlinear pricing function, thus discovering a new component or the fixed cost component in the pricing function.

1 Introduction

Previous asset pricing models with transactions costs focus mostly on the impact of transactions costs on the asset prices and trading strategies.[1] Under certain conditions, Constantides (1986), Heaton and Lucus (1996), Vayanos (1998), and Huang (2003) demonstrate that transactions costs have only a small impact on asset prices, whereas Amihud and Mendelson (1986) and Lo, Mamaysky, and Wang (2001) conclude that even small costs can lead to a significant illiquidity discount in asset prices.[2] However, two important features are missing in all those models. First, they consider a competitive equilibrium, thus ignoring the investor's strategic trading behavior or the impact of his trades on equilibrium stock prices. Second, they do not link transactions costs to the equilibrium stock prices in a direct manner. In addition, the previous equilibrium models typically consider only one type of cost, either the fixed cost or the proportional cost.

In contrast, there has been an extensive empirical literature documenting the effect of order flows on the stock price changes as well as linking directly asset prices to transactions costs. For example, Glosten and Harris (1988), Hasbrouch (1991), Madhavan and Smith (1991), and Brennan and Subrahmanyam (1998) find that asset prices respond positively to order flows and are related nonlinearly to proportional costs multiplied by signed trade size (positive for purchases and negative for sales), among other factors. Because the asset price structure is the foundation of an asset pricing model, it is crucial for theoretical models to be able to generate a pricing function that is consistent with the empirical findings. Otherwise, all of the subsequent results may not be sound. In other words, it is perhaps of first-order importance to derive equilibrium asset prices that are consistent with the empirical results.

The objective of this paper is to develop a parsimonious equilibrium model of asset pricing that incorporates the investor's strategic trading behavior as well as derives empirically sound asset prices. The paper studies how rational traders decide when to place orders for liquidity

[1]See, for example, Amihud and Mendelson (1986), Constantinides (1986), Duffie and Sun (1990), Foster and Viswanathan (1990), Dumas (1992), Heaton and Lucus (1996), Korn (1998), Schroeder (1998), Vayanos (1998, 2003), Leland (2000), Lo, Mamaysky, and Wang (2001), and Huang (2003).

[2]Heaton and Lucus and Huang also point out the possibility that costs can affect asset prices significantly.

reasons and provides potential explanations for certain asset pricing anomalies in an integrated fashion. In this model, there is one risk-free bond and one risky stock available for trading. There are two types of economic agents, a risk-averse market maker who sets the stock price competitively,[3] and a risk-averse investor who trades the stock strategically. It is assumed that trading in the bond market is free of transactions costs and that trading in the stock incurs costs. The investor incurs transactions costs in trading and the market maker incurs costs for processing the order flows submitted by the investor.

In the spirit of Kyle (1985), the market maker in our model sets the price competitively based not only on the investor's demand for the risky stock but also on the market maker's costs of processing the orders. The investor's transactions costs affect his demand and thus the stock price. The investor is endowed with a non-tradable, continuous stream of risky labor income and would like to hedge this income by trading on the risky stock. With transactions costs, he cannot hedge the income risk continuously.

Under certain regularity conditions, we demonstrate that there exists an equilibrium in our model in which the equilibrium stock price is of a nonlinear form. When transactions costs are small, which allows expansion techniques,[4] we obtain closed-form solutions for the stock price, the investor's trading strategies, and other equilibrium properties. In particular, we show that the equilibrium stock price depends linearly and positively on the investor's holding of the stock, depends nonlinearly on the market maker's proportional cost of processing the order flows multiplied by signed trade size (positive for purchases and negative for sales), as well as on the fixed cost multiplied by the inverse of the investor's order flow. Because the market maker is risk averse, she requires compensation for holding inventory in order to clear the order from the investor, which is the reason that the investor's demand affects the stock price in equilibrium. Because the market maker incurs both fixed and proportional costs for processing the investor's order flows, she passes these costs onto the investor in equilibrium. This is the reason that these costs affect the price in their particular ways.

[3]One may interpret the competitive market maker as one of the infinite number of identical market makers.

[4]See, for example, Korn (1998) and Lo, Mamaysky, and Wang (2001) for a description of small costs.

Empirical studies indicate that there exist positive autocorrelations in short-horizon stock returns and negative autocorrelations in long-horizon stock returns.[5] The empirical studies of Hasbrouck (1991), Foster and Viswanathan (1993), and others have also documented the predictability in the investor's order flows, that is, over a short period of time, there is a positive autocorrelation in the investor's order flows. We demonstrate that the presence of market impact costs and transactions costs naturally lead to these results. Even with market impact costs but without transactions costs, the investor breaks up large trades into small ones but can still trade continuously, so the positive autocorrelation in order flows will only be instantaneously lived. In the presence of transactions costs, however, it is not feasible for the investor to trade continuously so the positive autocorrelation in order flows can last. More specifically, the investor trades only when the deviation of his holdings in the stock from an ideal position reaches a lower or an upper boundary, where the ideal position is defined as the optimal position without transactions costs. For example, if the investor's position reaches the upper boundary due to a change in his income, then he reduces his holdings to a target level. The investor will not trade again until his holdings hit either the upper boundary or the lower boundary. It is found that after a sale order, it is more likely for the investor's holdings to hit the upper boundary than the lower one, resulting in positive autocorrelations in trading.

It is demonstrated that due to the positive autocorrelation in the order flow as well as to the market impact cost, the expected stock returns exhibit positive serial correlations. Over the long run, if the investor's income stream is mean reverting, then the order flows and the expected returns of the stock will exhibit negative serial correlations. Both the market impact cost and the transactions cost are essential to generate these results. These results may provide a rational explanation for the empirical findings on the autocorrelations as well as on the momentum effect [see, e.g., Jagedeesh and Titman (1993)]. It is also found that the market maker's processing costs may affect the stock price and the liquidity premium significantly, in the order of the 3/4 power of the costs.

[5]See, for example, Conrad and Kaul (1988) and Lo and Mackinlay (1988) for positive autocorrelations and Fama and French (1987), Lo and Mackinlay (1988), and Poterba and Summers (1988) for negative autocorrelations.

Consistent with the empirical findings of Brennan and Subrahmanyam (1995), the trade size in our model is negatively related to both the market impact cost and the proportional cost. The investor breaks up the trade to reduce market impact costs and with proportional costs, he trades less aggressively per trade but more frequently in equilibrium. We further predict that trade size is positively related to fixed costs. Also consistent with the empirical findings of Dufour and Engle (2000), our model obtains that when the volatility of the dividend process increases, the investor trades more frequently and the market impact costs of his trades become greater. These results arise because the market maker and investor are risk averse. When the dividend process becomes more volatile, it is less effective for the investor's trading to hedge his income risk. Therefore, the investor trades more frequently, and the risk-averse market maker demands a higher risk premium from the investor by making the price more sensitive to the investor's trade.

We use the ticker-by-ticker trades from the Trades and Automated (TAQ) database to test the nonlinear pricing function as well as estimate the various costs. Our empirical results confirm the nonlinear stock price structure derived from our theoretical model. Our empirical study extends those of Glosten and Harris (1988), Hasbrouch (1991), Madhavan and Smidt (1991), Brennan and Subrahmanyam (1998), and others in two dimensions. First, our study utilizes ten years of high frequency data whereas other studies typically use one year of data. Second, our study discovers a new component or the fixed cost component in the pricing function, in addition to the market impact cost and proportional cost. We also find that the market impact cost remains stable over time whereas both the fixed cost and the proportional cost decrease over time. Due to lack of data, previous studies cannot examine the time series pattern of the costs.

The rest of this paper is organized as follows. Section 2 specifies the notation and assumptions of the model. Section 3 solves a benchmark case in which transactions costs are absent. Section 4 presents the equilibrium stock price function in the general model. Section 5 obtains closed-form solutions to a special case in which transactions costs are small. Section 6 presents the empirical estimation of the pricing function. Section 7 concludes the papers. Technical proofs are provided in several proofs.

2 The Model

We consider an economy over a continuous-time horizon $[0, \infty)$. The following assumptions characterize our economy.

Assumption 1. There are two assets: a risk-free bond and a risky stock. The bond yields a constant rate of return $1 + r$ $(r > 0)$. At time t, each unit of stock generates a flow of dividend at an instantaneous rate of D_t:

$$dD_t = -\alpha_D(D_t - \bar{D})dt + b_D dB_{1t}, \tag{1}$$

where $-\alpha_D(D_t - \bar{D})$ represents the expected growth rate and $\bar{D}$ is the steady-state level of the dividend rate. For tractability, assume that the dividend process D_t and all other parameters in it are observable to all parties in the economy.

Assumption 2. There are two types of agents: an investor and a representative market maker. At time t, the investor receives an income stream at an instantaneous rate of

$$dX_t = -a_X X_t dt + b_X Z_t dB_{2t}, \tag{2}$$

$$dZ_t = -a_Z Z_t dt + b_Z dB_{3t}, \tag{3}$$

where the coefficients a and b are positive real constants and where B_{2t} and B_{3t} are standard Brownian motion processes. For simplicity, assume that the correlations are given by $Cov(dB_{1t}, dB_{3t}) = \eta$ and $Cov(dB_{1t}, dB_{2t}) = Cov(dB_{2t}, dB_{3t}) = 0$. Note that the steady-state level of the income process X_t is given by zero.

Because of the risky income stream the investor would like to trade in the stock to hedge the income risk. As shall be seen later, the investor's demand for stock vanishes if the income stream becomes risk free. Assume that the income process X_t is known to the investor only or that the representative market maker does not observe this process.

Assumption 3. Transactions in the bond market are costless, whereas transactions in the stock market incur costs. Specifically, the investor incurs trading costs and the market maker incurs costs for processing the order flows submitted by the investor. The cost functions are of the following form:

$$K_i(\Delta y) = A_i|\Delta y| + B_i, (A, B > 0), \tag{4}$$

where Δy denotes the trade size, where $i = I$ or $i = M$, with I and M representing investor and market maker, and where B_i and A_i denote the fixed cost component and the proportional cost component, respectively. For tractability, we have followed the literature by assuming that the proportional cost is proportional to the number of shares traded rather than the dollar value traded.[6]

Assumption 4. In the spirit of Kyle (1985) and Subrahmanyam (1991), the competitive market maker sets the stock price based on the observed current and lagged order flows sent by the investor. She earns a zero expected utility at every time t, that is, $E[U_m|\mathcal{F}_M(t)] \equiv E_t[U_m] = 0$, where $\mathcal{F}_M(t) = \{D_\tau, y_\tau, P_\tau : \tau \leq t\}$ denotes the market maker's information set with y_t and Δy_t being the number of shares held and the order flows submitted by the investor, respectively.Specifically, the market maker's expected utility is given by

$$0 = E_t\left[U_m\right] = E_t\left[y_{t_k}\left(P_{t_k} - \int_{s=0}^{\infty} \exp(-rs)D_{t_k+s}ds - \frac{B_M}{\Delta y_{t_k}} - \frac{A_M|\Delta y_{t_k}|}{\Delta y_{t_k}}\right)\right] - \frac{\gamma_m}{2}\sigma_p^2 y_{t_k}^2, \tag{5}$$

where γ_m is the market maker's risk aversion coefficient, t_k is the time of kth trading, P_{t_k} denotes the stock price, and $\sigma_p^2 = Var\left[(P_t - \int_{s=0}^{\infty} \exp(-rs)D_{t+s}ds)|\mathcal{F}_M(t)\right] = \frac{b_D^2}{2r(r+\alpha_D)^2}$.

This type of utility function has been used by Leland and Pyle (1977) and Subrahmanyam (1991). For a risk-neutral market maker or $\gamma_m = 0$, this assumption means that she earns a zero expected net profit as in Kyle (1985). To obtain an empirically relevant, nonlinear equilibrium stock price in a tractable manner, this assumption abstracts from the market maker's dynamic maximization problem. Because she must earn a zero net expected utility, she passes her processing costs onto the investor via the equilibrium stock price. The stock price depends

[6] See, for example, Vayanos (1998) for the same assumption.

on the market impact cost because the risk-averse market maker demands a risk premium for clearing the investor's orders. Equivalently, the investor's trades affect the stock price.

Assumption 5. The investor's utility function is of the exponential form, $-\exp[-\rho t - \gamma C_t]$, where γ is the investor's absolute risk aversion coefficient, where ρ is the investor's time preference parameter, and C_t is his consumption rate. Therefore, the investor chooses the optimal time as well as the optimal amount of stock to maximize his expected utility over an infinite time horizon:

$$E_0\left[\int_0^\infty - \exp\left[-\rho t - \gamma C_t\right] dt\right], \tag{6}$$

where $E_t[\cdot]$ denotes the expectation conditional on the investor's information set $\mathcal{F}(t) = \{D_\tau, X_\tau, P_\tau, Z_\tau : \tau \leq t\}$.

3 An Economy in the Absence of Transactions Costs

First consider a benchmark without transactions costs for trading the risky stock but with strategic trading. In this case trading takes place continuously. Even this seemingly simple benchmark case has not been solved in the previous literature. It shall be seen that this case does not generally admit closed-form solutions.

3.1 The Equilibrium Stock Price

Following Kyle (1985), we assume that the market maker employs a linear pricing rule given by

$$P_t = \frac{\bar{D}}{r} + \frac{D_t - \bar{D}}{\alpha_D + r} + \lambda y_t, \tag{7}$$

where y_t is the investor's order for the stock and where λ resembles the Kyle lambda that reflects the investor's market impact cost. Equation (5) for the market maker's expected utility can then be written as

$$0 = E_t\left[y_t\left(P_t - \int_{s=0}^\infty \exp(-rs)D_{t+s}ds\right)\right] - \frac{\gamma_m}{2}\sigma_p^2 y_t^2, \tag{8}$$

where

$$\sigma_p^2 = Var\left[(P_t - \int_{s=0}^\infty \exp(-rs)D_{t+s}ds)|\mathcal{F}_M(t)\right] = \frac{b_D^2}{2r(r+\alpha_D)^2}.$$

Evaluating the expectation and comparing the coefficients of y_t^2 on both sides, we have

$$\lambda = \frac{1}{2}\gamma_m \sigma_p^2. \tag{9}$$

Note that λ goes to zero when the market maker is risk neutral. It means that the investor's order will not have any impact on the stock price because the risk-neutral market maker does not require a risk premium.

3.2 The Investor's Dynamic Maximization Problem

In the absence of transactions costs, the investor trades continuously to hedge his continuous income shock. The investor's problem is to choose both his holdings y_t and the consumption rate C_t to maximize his expected utility given in Equation (6). Following the market microstructure literature, we confine the trading strategy to be a linear function of the state variable Z_t.[7] Conjecture that y_t is given by

$$y_t = kZ_t. \tag{10}$$

Applying Ito's lemma to y_t yields

$$dy_t = -a_Z y_t dt + kb_Z dB_{3t}.$$

The wealth of the investor's portfolio is then given by

$$dW_t = (rW_t - C_t + X_t)dt + y_t dQ_t, \tag{11}$$

where dQ_t is the investor's excess return given by

$$dQ_t = dP_t - rP_t dt + D_t dt = \left[-\lambda(a_Z + r)y_t\right] dt + \frac{b_D}{\alpha_D + r} dB_{1t} + k\lambda b_Z dB_{3t}.$$

Appendix A solves the investor's maximization problem using the dynamic programming approach. We present the results in the following proposition.

[7]It can be shown that it is redundant to include the level of income X_t due to the investor's exponential utility function. Further, it can be shown that if the diffusion term of the income process X_t is a constant, then the investor will hold the same number of shares of the stock throughout, that is, there will not be any trading.

Proposition 1. *There exists a linear equilibrium in which the stock price is given by Equation (7), the investor's optimal consumption rate is given by*

$$C_t = \frac{1}{\gamma}(r\gamma W_t + g_0 + g_1 X_t + 0.5 g_2 Z_t^2 - logr), \tag{12}$$

and his optimal trading strategy is given by $y_t = kZ_t$, where k satisfies the following equation:

$$0 = b_Z g_2 \left(\lambda k b_Z + \frac{\eta b_D}{\alpha_D + r}\right) + k\left[2(a+r)\lambda + r\gamma\sigma_y^2\right]. \tag{13}$$

Here the coefficients g_0, g_1, and g_2 are constants that satisfy the set of nonlinear equations given in the appendix. When the market maker is risk neutral, the relevant coefficients admit solutions in closed form:

$$g_2 = \frac{\sqrt{(0.5r + a_Z)^2 - b_X^2 b_Z^2(1-\eta^2)g_1^2} - (0.5r + a_Z)}{b_Z^2(1-\eta^2)}, \qquad g_1 = \frac{r\gamma}{r + a_X},$$

$$0 = (r - \rho) + r(g_0 - logr) - r\gamma - \frac{1}{2}b_Z^2 g_2^2, \qquad k = -\frac{\eta b_Z(\alpha_D + r)}{r\gamma b_D} g_2.$$

Remark 1. If η (the correlation coefficient between the X_t process and the Z_t process) vanishes, then we can show that $k = 0$ or that the investor's demand is zero. Intuitively, the investor cannot hedge his income shock using an unrelated security. In addition, closed-form solutions do not exist when the market maker is risk averse, which illustrates that the market impact cost greatly complicate the problem.

Remark 2. To examine the impact of the market impact cost on the investor's trading and expected utility, Panel A and Panel B in Figure 3 plot g_2 and k again λ (change in γ_M), respectively. As expected, when λ increases, g_2 and k both decrease. In other words, the market impact cost reduces both the investor's expected utility and the investor's trade in hedging his risky income stream.

Appendix B solves a benchmark case in which the market maker observes the investor's income stream process X_t and its diffusion process Z_t. In this case, the market maker can simply sets the equilibrium stock price in terms of Z_t so that the investor's trade y_t will not affect the price in a direct manner. As shall be seen, the investor's trading as reflected in k and his expected utility as reflected in g_2 are higher in this competitive equilibrium than those in the strategic equilibrium in which the investor's trades affect the stock price directly.

4 An Economy in the Presence of Transactions Costs

When there are transactions costs, the investor cannot hedge his income shock continuously. The investor must choose an optimal time to hedge. The equilibrium in this case is defined as follow. (1) The stock price is efficient at the time of trading t_k, that is,

$$0 = E_t\left[y_{t_k}\left(P_t - \int_{s=0}^{\infty}\exp(-rs)D_{t+s}ds - \frac{B_m}{\Delta y_{t_k}} - \frac{A_m|\Delta y_{t_k}|}{\Delta y_{t_k}}\right)\right] - \frac{\gamma}{2}\sigma_p^2 y_{t_k}^2, \tag{14}$$

where σ_p^2 is given by

$$\sigma_p^2 = Var\left[(P_t - \int_{s=0}^{\infty}\exp(-rs)D_{t+s}ds)|\mathcal{F}_M(t)\right] = \frac{b_D^2}{2r(r+\alpha_D)^2}. \tag{15}$$

(2) The investor chooses consumption C_t, the time of trading t_k, and the quantity of trading Δy_{t_k} to maximize his expected utility:

$$E_0\left[\int_0^{\infty} -e^{-\varphi t - \gamma C_t}\right]dt,$$

subject to

$$M_t = \int_0^t \left[rM_s - C_s + \theta_s D_s\right]ds - \sum_{k:0\le t_k \le t}\left[P_{t_k}\Delta y_k + K(\Delta y_k)\right],$$
$$W_t = M_t + \theta_t P_t,$$

where Δy_k is the investor's trade size at trading time t_k, M_t is his position in bond at time t, and θy_k is his position in stock at time t_k.

4.1 The Nonlinear Price Function

The price efficiency as given in Equation (14) determines the equilibrium stock price.

Theorem 1. *When $A_i > 0$ and $B_i > 0$, under some regularity conditions, there exists a nonlinear price function of the following form:*

$$P_k = \frac{\bar{D}}{r} + \frac{D - \bar{D}}{\alpha_D + r} + \lambda y_k + A_m sign(\Delta y_k) + \frac{B_m}{\Delta y_k}, \tag{16}$$

where Δy_k *denotes the trade size at kth trading.*

It is easy to verify that this price function satisfies the price efficiency condition (14). The first three terms correspond to the equilibrium stock price without transactions costs. $A_m sign(\Delta y_k)$ captures the impact of the market maker's proportional costs on the price: the price is higher for a buy order and lower for a sell order. $B_m \Delta y_k$ represents the market maker's fixed cost per share, which also means that the price is higher for a buy order and lower for a sell order. Although the investor's trading costs do not appear directly in the pricing function, they affect the trade size Δy_k and λ, thus influencing the stock price indirectly. Because the competitive market maker must always achieve a zero net expected utility, she passes all the processing costs onto the investor. Even if the investor himself does not incur any costs in trading, he will still not trade continuously because he indirectly bears the marker maker's costs. We next discuss the investor's dynamic problem that determines his trade size Δy.

4.2 The Investor's Maximization Problem

In the non-trading region, we have that the value of the investor's bond position evolves as

$$dM_t = (rM_t + X_t - C + D_t y_t)dt,$$

where y_t is the investor's holdings in the stock. The investor's Bellman equation is

$$\begin{aligned} 0 &= -\exp\left[-\rho t - \gamma C\right] + J_t + J_M\{rM - C + X + Dy\} + J_D\left[-\alpha_D(D - \bar{D})\right] + \frac{1}{2}J_{DD}\sigma_D^2 \\ &- aXJ_X + \frac{1}{2}J_{XX}b_X^2 Z^2 - a_Z Z J_Z + \frac{1}{2}b_Z^2 J_{ZZ} + J_{DZ}\eta b_D b_Z. \end{aligned} \tag{17}$$

Conjecture that the investor's value function is of the form:

$$\begin{aligned} J(M, D, y, X, Z, t) &\equiv \exp(-\rho t) I(M, D, y, X, Z) \\ I(M, D, y, X, Z) &= -\exp\left\{-r\gamma\left[M + y\left(\bar{P} + \frac{1}{2}\lambda y\right)\right]\right\} \\ &\times \exp\{-g_0 - g_1 X - \frac{1}{2}g_2 Z^2 - V(y, Z)\Big\}, \end{aligned} \tag{18}$$

where $\bar{P} = \frac{\bar{D}}{r} + \frac{D - \bar{D}}{\alpha_D + r}$ and M is the investor's position in bond.

With transactions costs, the investor cannot trade continuously. Instead, the investor does not trade until his holding in the stock reaches an upper boundary $y_t^* + Z_u$ or a lower boundary $y_t^* + Z_d$. When it reaches the upper boundary the investor sell an amount of Δy_t to reduce his position to the upper target position $y_t^* + Z_{mu}$, whereas when his position hits the lower boundary, he buys an amount of Δy_t to the lower target position $y_t^* + Z_d$, where Δy_t represents the trade size. To solve for the trade size, we need to impose the value-matching condition, that is, the investor is indifferent between trading and not trading:

$$V(\bar{y} + \Delta y, Z) = V(\bar{y}, Z) + r\gamma \left[(B + B_m) + |\Delta y|(A + A_m) + \frac{1}{2}\lambda(\Delta y)^2 \right]. \tag{19}$$

where $\bar{y}$ is a point on the boundary, Δy is the trade size, and $\bar{y} + \Delta y$ is the corresponding point on the target. We also impose the smooth-pasting condition, that is, the derivatives of the investor's value function are equal in both regions:

$$\begin{aligned} \frac{\partial V(\bar{y} + \Delta y, Z)}{\partial y} &= -r\gamma(A_I + A_M) + r\gamma\lambda\Delta y, \quad \text{if} \quad \Delta y < 0, \\ \frac{\partial V(\bar{y} + \Delta y, Z)}{\partial y} &= r\gamma(A_I + A_M) + r\gamma\lambda\Delta y, \quad \text{if} \quad \Delta y > 0. \end{aligned} \tag{20}$$

The next theorem discusses the existence of a solution for the investor's dynamic maximization problem.

Theorem 2. *Let $y^* + Z_d(Z), y^* + Z_{md}(Z), y^* + Z_{mu}(Z), y^* + Z_u(Z)$ be the solutions to (19), and (20), where $V(y, Z)$ satisfies (17) for $y \in [Z_d(Z), Z_U(Z)]$. Then, under certain conditions, there exists a set of optimal solutions to the investor's value function, his optimal consumption policy, and his trading strategy.*

Proof. See Appendix C.

Remark. Theorems 1 and 2 establish the equilibrium stock price function as well as the existence of the solutions to the investor's maximization problem under general cost functions. Because they do not offer specific solutions, it is difficult to make sharp predictions about the effects of market impact costs and transactions costs on the investor's trading strategies. In the next section we solve a case in which the market maker is close to risk neutral and the transactions costs are small so that the expansion techniques apply.

5 Small Market Impact Costs and Small Transactions Costs

For tractability, we assume that $\eta = 1$ following Lo, Mamaysky, and Wang (2001). We scale the cost function as follows:

$$K_i(\varepsilon, \Delta y) = A_{1,i}\varepsilon^{0.75}|\Delta y| + B_{1,i}\varepsilon, \tag{21}$$

where $0 < \varepsilon \ll 1$, $A_{1,i}$ and $B_{1,i}$ are real positive constants, and Δy denotes the trade size in the stock. We also scale λ as $\lambda = \lambda_0 \varepsilon^{0.5}$. Accordingly, we scale y as

$$y = y^{**} + \varepsilon^{0.25} Y_1, \tag{22}$$

where the time index is omitted for expositional simplicity.

Substituting the value function (18) into the investor's Bellman equation (17), we obtain

$$\begin{aligned} 0 &= (r-\rho) - rlogr + \frac{1}{2}r^2\gamma\lambda y^2 + rV + r(-\gamma X + g_0 + g_1 X + 0.5 g_2 Z^2) \\ &+ \frac{(r\gamma\sigma_D)^2}{2(\alpha_D + r)^2}y^2 + a_X g_1 X + \frac{1}{2}b_X^2 Z^2 g_1^2 + a_Z Z(g_2 Z + V_Z) + \frac{1}{2}b_Z^2\left[(g_2 Z + V_Z)^2 - g_2 - V_{ZZ}\right] \\ &+ \frac{r\gamma y}{\alpha_D + r}\eta(g_2 Z + V_Z)b_D b_Z. \end{aligned} \tag{23}$$

or

$$\begin{aligned} 0 &= \left[(r-\rho) - rlogr + rg_0 - \frac{1}{2}b_Z^2 g_2\right] + (-r\gamma + rg_1 + a_X g_1)X + rV + \frac{1}{2}(V_Z^2 - V_{ZZ}) \\ &+ \left[a_Z Z + \frac{r\gamma y}{\alpha_D + r}b_D b_Z \eta + b_Z^2 g_2 Z\right]V_Z + \frac{1}{2}\left(\frac{r\gamma b_D}{\alpha_D + r}\right)^2 y^2 + \frac{1}{2}r^2\lambda y^2 + \frac{1}{2}rg_2 Z^2 \\ &+ \frac{1}{2}b_X^2 Z^2 g_1^2 + a_Z g_2 Z^2 + \frac{1}{2}b_Z^2 g_2^2 Z^2 + \frac{r\gamma y}{\alpha_D + r}\eta g_2 Z b_D b_Z. \end{aligned} \tag{24}$$

Suppose that $\eta = 1$. Using the expression for g_0 and g_2 in Proposition 1 for the risk-neutral market maker and without transactions costs, we can simplify the above equation:

$$\begin{aligned} 0 &= rV + \frac{1}{2}b_Z^2\left[V_Z^2 - V_{ZZ}\right] + \left[a_Z Z + \frac{r\gamma b_D b_Z}{\alpha_D + r}(y - y^{**})\right]V_Z \\ &+ \frac{1}{2}\left(\frac{r\gamma b_D}{\alpha_D + r}\right)^2 (y - y^{**})^2 + \frac{1}{2}r^2\lambda y^2 \\ &\equiv a_1(V_Z^2 - V_{ZZ}) + a_2(y - y^{**})V_Z + a_3 V + a_4(y - y^{**})^2 + a_5 Z V_Z + a_6 y^2, \end{aligned} \tag{25}$$

where y^{**} denotes the investor's optimal demand in the absence of both market impact costs and transactions costs.

The value-matching condition is given by

$$\begin{aligned}
&\exp(-\rho t)\exp -\{r\gamma\left[M+\bar{y}(\bar{P}+0.5\lambda\bar{y})\right]+g_0+g_1X+0.5g_2Z^2+V(\bar{y},Z)\}\\
&=\exp(-\rho t)\exp -\{r\gamma\left[M+(\bar{y}+\Delta y)(\bar{P}+0.5\lambda(\bar{y}+\Delta y))\right]+g_0+g_1X+0.5g_2Z^2+V(\bar{y}+\Delta y,Z)\}\\
&\times\exp r\gamma\{|\Delta y|(A+A_m)+(B+B_m)+\Delta y\left[\bar{P}+\lambda(\bar{y}+\Delta y)\right]\},
\end{aligned} \tag{26}$$

where $\bar{P}=\frac{\bar{D}}{r}+\frac{D-\bar{D}}{r+\alpha_D}$, $\bar{y}$ is a point on the boundary, Δy is the trade size, and $\bar{y}+\Delta y$ is the corresponding point on the target. By rearrangement, we have

$$V(\bar{y}+\Delta y,Z)=V(\bar{y},Z)+r\gamma\left[(B+B_m)+|\Delta y|(A+A_m)+\frac{1}{2}\lambda(\Delta y)^2\right]. \tag{27}$$

The smooth-pasting conditions give

$$\begin{aligned}
\frac{\partial V(\bar{y}+\Delta y,Z)}{\partial y}&=-r\gamma(A+A_m)+r\gamma\lambda\Delta y,\quad \text{if}\quad \Delta y<0,\\
\frac{\partial V(\bar{y}+\Delta y,Z)}{\partial y}&=r\gamma(A+A_m)+r\gamma\lambda\Delta y,\quad \text{if}\quad \Delta y>0.
\end{aligned} \tag{28}$$

In the region of $Y_1\geq 0$, the value-matching condition becomes

$$d_1(Y_1^b)^4+G_1(Z)(Y_1^b)^2=d_1(Y_1^T)^4+G_1(Z)(Y_1^T)^2-r\gamma A_1(Y_1^b-Y_1^T)-r\gamma B-\frac{1}{2}r\gamma\lambda_0(Y_1^b-Y_1^T)^2, \tag{29}$$

where superscript b denotes the upper boundary and superscript T denotes the upper target. The smooth-pasting conditions yield

$$4d_1(Y_1^b)^3+2G_1(Z)Y_1^T=4d_1(Y_1^T)^3+2G_1(Z)Y_1^T=-r\gamma\left[A_1+\lambda_0(Y_1^d-Y_1^T)\right]. \tag{30}$$

Conjecture that $V(Y_1,Z)$ is of the form:

$$V(Y_1,Z)=H_2(Z)\varepsilon^{0.5}+H_3(Z)\varepsilon^{0.75}+H_4(Y_1,Z)\varepsilon+H_5(Y_1,Z)\varepsilon^{1.25}+H_6(Y_1,Z)\varepsilon^{1.25}+o(\varepsilon^{1.5}). \tag{31}$$

We have

$$V_Z=H_{\{2,Z\}}\varepsilon^{0.5}+\left[H_{\{3,Z\}}-\beta_1H_{\{4,Y_1\}}\right]\varepsilon^{0.75}+\left[H_{\{4,Z\}}-\beta_1H_{\{5,Y_1\}}\right]\varepsilon+o(\varepsilon), \tag{32}$$

$$V_{ZZ} = \left[H_{\{2,ZZ\}} + \beta_1^2 H_{\{4,Y_1Y_1\}} \right] \varepsilon^{0.5} + \left[H_{\{3,ZZ\}} + \beta_1 H_{\{4,ZY_1\}} + \beta_1^2 H_{\{5,Y_1Y_1\}} \right] \varepsilon^{0.75}$$

$$+ \Big[H_{\{4,ZZ\}} + \beta_1 H_{\{5,ZY_1\}}$$

where $H_{\{i,j\}}$ denotes the partial derivative of H_i with respect to j and $H_{\{i,jk\}}$ denotes the second order partial derivative of H_i with respect to j and k, and $\beta_1 = -\frac{\eta b_Z(\alpha_D + r)}{r\gamma b_D} g_2$.

Substituting the above expressions into the investor's Bellman equation (25) and comparing the coefficients of $\varepsilon^{0.5}$ on both sides, we have

$$\beta_1^2 H_{\{4,Y_1Y_1\}} = \frac{a_3}{a_1} H_2 + \frac{a_4}{a_1} Y_1^2 + \frac{a_5}{a_1} H_{\{2,Z\}} Z - H_{\{2,ZZ\}} + \frac{a_6}{a_1} Z^2. \tag{34}$$

$H_4(Y_1, Z)$ has the form

$$H_4(Y_1, Z) = d_1 {Y_1}^4 + G_1(Z) Y_1^2 + G_2(Z) Y_1 + G_3(Z), \tag{35}$$

where $d_1 = \frac{a_4}{12 a_1 {\beta_1}^2}$. From the value-matching conditions and smooth-pasting conditions, we have $G_2(Z) = 0$. As a result, $H_4(Y_1, Z)$ is an even function with respect to Y_1 , that is, $H_4(Y_1, Z) = H_4(-Y_1, Z)$. Thus, we can focus on the case of $Y_1 \geq 0$.

In the region of $Y_1 \geq 0$, the value-matching condition is

$$\begin{aligned} d_1 (Y_1^b)^4 + G_1(Z) (Y_1^b)^2 &= d_1 (Y_1^T)^4 + G_1(Z) (Y_1^T)^2 \\ -r\gamma A_1 (Y_1^b - Y_1^T) - r\gamma B_1 &- \frac{r\gamma\lambda_0}{2} (Y_1^b - Y_1^T)^2, \end{aligned} \tag{36}$$

where superscript b denotes the upper boundary and superscript T denotes the upper target. The smooth-pasting conditions give

$$4 d_1 (Y_1^b)^3 + 2 G_1(Z) Y_1^b = 4 d_1 (Y_1^T)^3 + 2 G_1(Z) Y_1^T = -r\gamma A_1 - r\gamma\lambda_0 (Y_1^b - Y_1^T). \tag{37}$$

Let $Y_1^b + Y_1^T \equiv x_1$ and $Y_1^b - Y_1^T \equiv x_2$. x_2 then denotes the trade size or the amount of trading required to bring the investor's holding from the boundary level to the target level. Equation (37) now reduces to

$$G_1(Z) = -2 \frac{d_1 {Y_1^b}^3 - d_1 {Y_1^T}^3}{Y_1^b - Y_1^T} = -\frac{d_1 (3x_1^2 + x_2^2)}{2} \tag{38}$$

and

$$\frac{4 d_1 {y_1^b}^3 + r\gamma A_1 + r\gamma\lambda_0 (Y_1^b - Y_1^T)}{4 d_1 {y_1^T}^3 + r\gamma A_1 + r\gamma\lambda_0 (Y_1^b - Y_1^T)} = \frac{y_1^b}{y_1^T}. \tag{39}$$

By rearrangement, we have

$$r\gamma A_1 \quad = \quad d_1(x_1^2 - x_2^2)x_1 - r\gamma\lambda_0 x_2.$$

From equation (37), we also have

$$-r\gamma A_1(Y_1^b - y_1^T) - r\gamma\lambda_0(Y_1^b - Y_1^T)^2 = 4d_1[(Y_1^b)^4 - (y_1^T)^4] + 2G_1(Z)[(Y_1^b)^2 - (Y_1^T)^2].$$

Using the above expression for $-r\gamma A_1(Y_1^b - Y_1^T)$ and the expression for $G_1(Z)$, the value-matching condition then reduces to

$$3d_1[(Y_1^b)^4 - (Y_1^T)^4] + G_1(Z)[(Y_1^b)^2 - (Y_1^T)^2] - r\gamma B + \frac{1}{2}r\gamma\lambda_0(Y_1^b - Y_1^T)^2 = 0 \qquad (40)$$

or

$$r\gamma B = d_1 x_2^3 x_1 + \frac{1}{2}r\gamma\lambda_0 x_2^2. \qquad (41)$$

By rearrangement, the above equations reduce to

$$x_1 = \frac{r\gamma B_1 - 0.5r\gamma\lambda_0 x_2^2}{d_1 x_2^3}, \qquad (42)$$

$$r\gamma A_1 x_2 + r\gamma\lambda_0 x_2^2 = d_1 x_1^3 x_2 - r\gamma B_1. \qquad (43)$$

Remark. Equation (42) implies that x_1 is a decreasing function of x_2 and we can also prove that $\frac{\partial x_1^3 x_2}{\partial x_2} < 0$. From Equation (43), since the left hand side is an increasing function of x_2 and the right hand side is a decreasing function of x_2, there exists a unique set of solutions (x_1, x_2). The following proposition gives the relations between the investor's trade size and costs.

Proposition 2. *The trade size decreases with the proportional cost and the market impact cost but increases with the fixed cost. Mathematically, we have* $\frac{\partial x_2}{\partial A_m} < 0$, $\frac{\partial x_2}{\partial \lambda} < 0$, *and* $\frac{\partial x_2}{\partial B_m} > 0$.

When the market impact cost and the proportional cost increase, the investor incurs more costs per share of trade, so he reduces the trade size in equilibrium. With a fixed cost, the investor's cost is independent of his trade size, so he trades more in order to bring his holdings closer to the desired position.

We next use these results to examine the impact of various costs on the stock price and return as well as the serial correlations between the investor's trading strategies and the stock returns.

5.1 The Impact of Costs on the Stock Price and Return

For tractability, we assume that $\alpha_D = 0$ or that the dividend process follows a random walk. Start with the equilibrium stock price function (16):

$$P_t = \frac{D_t}{r} + \lambda y_t + \frac{B}{\Delta y_{kt}} + Asign(\Delta y_{kt}).$$

From the expansion solution, we know that Δy_t, A, B, and λ are proportional to $\epsilon^{1/4}$, $\epsilon^{3/4}$, ϵ, and $\epsilon^{1/2}$, respectively. Thus, the price is proportional to the cost in the order of 3/4 power. The dollar return defined as the stock price difference between the time interval Δt is given by

$$\Delta P_t = \sum_{n=1}^{N} \frac{\Delta D_{nt}}{r} + \sum_{n=1}^{N} \lambda \Delta y_{nt} + B\left[\frac{1}{\Delta y_N} - \frac{1}{\Delta y_1}\right] + A\left[sign(\Delta y_n) - sign(\Delta y_1)\right]. \tag{44}$$

It can be seen that the return is also proportional to $\epsilon^{3/4}$, which suggests that trading costs can have first-order effects on stock returns.

5.2 The Autocorrelations of the Investor's Trade Size and the Expected Returns

Without transactions costs the investor trades continuously, so the serial correlations of the investor's trading $\Delta y_t = y_t - y_{t-1}$ (signed trade) depend on those of the $\Delta Z_t = Z_t - Z_{t-1}$ process. If the ΔZ_t process does not exhibit positive autocorrelation as in the current paper, then Δy_t will not exhibit positive autocorrelations. With transactions costs, however, the investor trades only when his holdings reach either the upper boundary or the lower boundary. For example, if the investor's holdings hit the upper boundary, he sells a certain amount so that his holdings reduce to the upper target level. The investor's next trade will not occur until his holdings hit the two boundaries. Because the difference between the upper target position and the upper boundary is less than that between the upper target and the lower boundary, it is more likely that a sell order will be followed by a sell order, resulting in positive autocorrelations. Similarly, if the investor's holdings hit the lower boundary, he buys a ceratin amount to bring his holdings up to the lower target level. Because the lower target is closer to the lower boundary than to the upper boundary, it is more likely that the investor will buy again, resulting in positive autocorrelations. From the expression for the stock return (44), a positive autocorrelation of the investor's singed trade leads to a positive autocorrelation for the stock return process. Over a longer time period, if the Z_t process is mean reverting, then negative autocorrelations will arise. Figures 2 and 3 plot the autocorrelations of the investor's signed trade size and the expected return against horizon Δ, respectively.

The next theorem formalizes the results regarding serial correlations. Its proof is given in Appendix D.

Theorem 3. *Both the investor's signed trade size and the expected stock return processes can be positively autocorrelated.*

Remark. Notice that transactions costs alone can lead to positive autocorrelations in the investor's trading, but they are not sufficient to generate positive autocorrelations in the stock return. The reason is that the components in the equilibrium stock price, which are associated with the transactions costs, are of temporary effects or the autocorrelations related to these components decay to zero much faster than those related to the market impact cost. In addition, the market impact cost alone cannot lead to positive autocorrelations in trading and stock returns. In the absence of transactions costs, trading takes place continuously and its serial correlations are not positive. This theorem sheds light on a rational explanation for the momentum effect and is consistent with the results of the empirical studies that find positive autocorrelations for investors' trading strategies and stock returns.

5.3 Risk, Market Impact, and Trading Frequency

As Biais, Glosten, and Spatt (2004) note, Engle and Dufour (2000) have established that in volatile times, trades and orders are more frequent, the price impact of trades is greater, and the positive autocorrelation of signed trades increases. Our model, which engogenizes the investor's trading frequency, makes the explanation of these empirical results possible.

With transactions costs, trading takes place only at discrete time τ_k when trading gains outweigh trading costs. In this subsection we first determine the investor's trading frequency or the expected time between investor's two trades and then investigate how the risk of the dividend process or the dollar return process of the stock affects the investor's trading frequency as well as the impact of the investor's trading on the stock price. For tractability of proofs, this subsection assumes that $a_Z = 0$ or that the Z_t process follows a random walk.[8]

The market impact cost as reflected by coefficient λ is proportional to b_D, the volatility of the dividend process:

$$\lambda = \frac{\gamma b_D^2}{4r(r + \alpha_D)^2}.$$

Because the market maker is risk averse, she requires a higher risk premium for sharing risk with the investor so the market impact cost of the investor's trades is higher. The next proposition relates the average trading interval between two trades to the trade size, among other determinants.

Proposition 3. *The average time interval between the investor's two consecutive trades is given by*

$$E[T_k] = \frac{x_1 x_2}{b_Z^2 k^2} = \frac{r\gamma B_1 - 0.5 r\gamma\lambda_0 x_2^2}{d_1 x_2^2 b_Z^2 k^2}, \tag{45}$$

where k is defined in Proposition 1, x_2 denotes the investor's trade size and x_1 and x_2 are determined by Equations (42) and (43).

It can be seen that after controlling for certain exogenous parameters such as r, γ, B_1, and d_1, the average time interval between two trades is inversely related to the trade size. Because the trade size x_2 must be solved numerically, we are unable to obtain an analytical relation between the volatility of the dividend process and the average interval between trades. However, many numerical calculations yield the same pattern, that is, when the dividend process becomes more volatile, the average time interval becomes shorter or the investor trades more frequently for liquidity reasons. The intuition for this result is that because the investor is risk averse, he would like to trade more frequently to hedge the dividend risk that is correlated with the risk of his income streams.

Figure 4 plots λ and $E[T_k]$ against b_D. As expected, λ increases with and $E[T_k]$ decreases

[8]Numerical simulations confirm that the results obtained for $a_Z = 0$ also hold for the Z_t processes with positive a_Z values.

with b_D. This figure also plots the volatility of the absolute dollar return, ΔP_t, against b_D and finds a monotonically positive relation between the two volatilities.[9] Indirectly, this figure shows that λ and $E[T_k]$ would also decrease and increase with the dollar return volatility, respectively.

6 Empirical Estimation

Equation (16) determines the stock price dynamics, which captures both the market impact cost and the transactions costs. Note that the previous theoretical literature has not captured the nonlinear feature resulted from the transactions costs. Empirically, Glosten and Harris (1988), Hasbrouch (1991), Madhavan and Smith (1991), and Brennan and Subrahmanyam (1998) have discovered that asset prices respond positively to order flows and are related nonlinearly to proportional costs multiplied by signed trade size. These studies, however, have not uncovered the nonlinear component caused by the fixed cost. In addition, the studies use small data sets that typically cover a one year of tick-by-tick trades. In this section, we confirm the previous studies by using a larger data set as well as confirms the fixed cost related nonlinear term in the stock price.

Because there is a unit root in the dividend process, we set $\alpha_D = 0$. The price change process is then given by

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + K/\Delta y_t - K/\Delta y_{t-1} + \epsilon_t, \tag{46}$$

where Δy_t is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$, that is, if the trade size Δy_t is positive, $v_t = 1$, otherwise, $v_t = -1$.

We estimate the market impact cost λ, the fixed cost K, and the proportional cost k, based on equation (**??**) for each stock and each quarter using tick-by-tick data. To take into account potential misspecifications, we include an intercept term in the regression specification.[10] We assume that $\lambda_i \sim N(\mu_1, \sigma_1^2)$, $K_i \sim N(\mu_1, \sigma_2^2)$, and $k_i \sim N(\mu_3, \sigma_3^2)$, where subscript i denotes

[9] The volatility of the dollar return is defined as the standard deviation of the daily price change.

[10] Glosten and Harris (1988) estimate a simpler price function in the absence of the fixed cost component $\frac{K}{\Delta y_t}$. Here, for simplicity, we assume that ϵ_t is identically and independently distributed.

stock i. We do cross-section test of the following hypothesis:

$$H0 : \mu_1 > 0, \quad H0 : \mu_2 > 0, \quad H0 : \mu_3 > 0. \tag{47}$$

The relevant statistics follow t distributions, which are given by

$$Z_1 = \frac{\sqrt{T}\bar{\mu}_1}{S_1} \sim t(T-1), \quad Z_2 = \frac{\sqrt{T}\bar{\mu}_2}{S_2} \sim t(T-1), \quad Z_3 = \frac{\sqrt{T}\bar{\mu}_3}{S_3} \sim t(T-1), \tag{48}$$

where T is the number of cross-sectional stocks in our sample for each quarter. $\bar{\mu}_1 = \frac{1}{T}\sum_{i=1}^{T} \mu_{1,i}$, $\bar{\mu}_2 = \frac{1}{T}\sum_{i=1}^{T} \mu_{2,i}$, $\bar{\mu}_3 = \frac{1}{T}\sum_{i=1}^{T} \mu_{3,i}$, $S_1 = \frac{1}{T-1}\sum_{i=1}^{T} (\mu_{1,i} - \bar{\mu}_1)^2$, $S_2 = \frac{1}{T-1}\sum_{i=1}^{T} (\mu_{2,i} - \bar{\mu}_2)^2$, and $S_3 = \frac{1}{T-1}\sum_{i=1}^{T} (\mu_{3,i} - \bar{\mu}_3)^2$.

6.1 Data

We consider a sample of stocks traded on the NASDAQ . We use ticker-by-ticker trades contained in the Trades and Automated Quotes (TAQ) database to estimate the cost parameters λ, k, and K for each stock. The time period is from January 1, 1993 through December 31, 2002. We use the University of Chicago's Center for Research in Security Prices (CRSP) daily database to determine the relevant stocks for the estimation of the above three types of costs. We estimate our model within each quarter. Following Stoll (2000), we clean the data and exclude stocks for the following reasons:

1. ADRs and REITs.
2. Closing price below 2 dollars on at least one day.
3. Not traded each day.
4. Stock split.

Thus, this sample consists of all common stocks that had ticker names on both the CRSP and the TAQ and were traded for every three-month period.

Following Stoll (2000) and Chordia, Roll, and Subrahmanyam (2001, 2002), quotes from the exchange other than the exchange of listing are excluded, and the price and quote data must occur between 9:30AM and 4:00PM. We omit the overnight price changes from the analyses to avoid mixing the price change series with those at the opening. Following the convention of Lee and Ready (1991), for each transaction the quote preceding the transaction by at least 5 seconds is associated with the trade. We classify trades as buyer-or-seller initiated trades as follows. If a transaction occurs above (below) the matched quote mid-points, it is regarded as a purchase (sell). Following Brennan and Subrahmanyam (1998), if a transaction occurs exactly at the midpoint of the bid and ask, it is signed using the previous transaction price according to the tick test (i.e. buys (sells) if the sign of the last nonzero price change is positive (negative)). We exclude the price when there is no opening quote. Thus, the first trade after the opening time is ignored. We also exclude the trades with negative prices. In addition, as in Salka (2003), only quotes that satisfy the following filter conditions are retained: the bid-ask spread is positive and below five dollars, the bid-ask spread divided by the midpoint of the quoted bid and ask is less than 10% if the midpoint is greater than or equal to 50 dollars, and quoted spread is less than 25% for midpoints less than 50 dollars. These conditions guarantee the use of reasonable quotes in our analysis. The average number of stocks for each quarter in our sample is around 2000 (see tables 1 through 6).

6.2 Results

Tables 1 and 2 give the average value of market impact cost λ and the corresponding t statistics Z_1 for each quarter. Table 1 presents the results from 1993 through 1997 and table 2 presents the results from 1998 through 2002. In table 1, there are only two quarters in which the t ratios are not significantly positive at 5% significance level. The average market impact cost λ is $1.44e-06$. In table 2, there are 4 quarters in which the t ratios are not significantly positive. The average market impact cost is $8.44e-07$. Overall, the market impact costs λ is significantly positive at 1% significance level for most of the quarters. Figure 5 plots the average market impact cost and the corresponding t ratio against time t. It shows that the market impact cost remains stable over time. Specifically, the average market impact cost is $8.44e-07$ from 1998 to 2002. However, if we exclude two quarters with negative λ, which are not significant from zero, then the average market impact cost has the same magnitude as in the period from 1993 to 1997.

Tables 3 and 4 offer the average value of the fixed cost K and the corresponding t statistics Z_2 for each quarter. Table 3 presents the results from 1993 to 1997 and table 4 presents the results from 1998 to 2002. In table 3, the t ratios are not significantly positive at 5% significant level for only one quarter from 1993 to 1997. The average fixed cost K is 4.1093 dollars. Somewhat different results arise in table 4, however, where there are 10 quarters in which the t ratios are not significantly positive at 5% significance level from 1998 to 2002. The average fixed cost is 0.3343 dollars. Overall, the fixed cost K is still significantly positive at 5% significance level for the most of the quarters in our sample. However, after 2000, this component of trading cost is not significantly positive. For some quarters, it is negative. Figure 6 plots the average fixed cost and the corresponding t ratio against time t. It shows that the fixed cost decreases over time quite dramatically, especially after 2000.

Tables 5 and 6 present the average value of the proportional cost k and the corresponding t statistics Z_3 for each quarter. Table 5 contains the results from 1993 through 1997 and table 6 contains the results from 1998 through 2002. In table 5, all of the t ratios are significantly positive at 1% significance level. The average proportional cost k is 0.1209 dollars per share. Similarly, In table 6, the t ratios are significantly positive at 1% significance level for each quarter from 1998 to 2002. The average proportional cost is 0.06556 dollars. Overall, the proportional cost k is significantly positive at 1% significance level from 1993 to 2002. Figure 7 shows that the proportional cost decreases over time quite dramatically. The value is about 0.1388 dollars per trade at the beginning of 1993. however, it decreases to about 0.03 dollars at the end of 2002.

In summary, our empirical study confirms the nonlinear stock price function derived from our theoretical model. The market impact is stable over time. However, the fixed cost and the proportional cost components decrease over time. The proportional cost is still significantly positive for all the quarters in our sample, but the fixed cost is insignificantly different from zero after year 2000.

7 Conclusion

This paper develops a simple equilibrium model with both market impact costs and transactions costs. Because the market maker is risk averse, she demands a risk premium for clearing the market. Consequently, the investor's trades affect the stock price. In other words, the investor must behave strategically in trading. Because the market maker incurs costs in processing the investor's order flows, she passes the costs onto the investor by taking them into account in pricing the stock in equilibrium. As a result, the stock price depends positively and linearly on the dividend payment and the investor's holdings in the stock, as in a model without transactions costs. In addition, the stock price depends nonlinearly on the fixed and proportional costs of the market maker. Our empirical study supports this equilibrium price structure.

To obtain tractable solutions for the investor's trading strategies, we solve a special case in which both the transactions costs and market impact costs are small so that the expansion techniques can be applied. We demonstrate that the investor does not trade continuously. Instead, the investor trades only when his holdings reach either an upper boundary or a lower boundary. When the investor's holdings hit the upper (lower) boundary, he trades so that his holdings reach an upper (lower) target level rather than a perfect hedging position to save costs. The investor's trade size, which allows the investor's holdings from the boundary positions to the target positions, decreases when the market impact cost and the proportional cost increase and it increases when the fixed cost increases.

Due to the properties of trading regions, the investor's trades may be positively autocorrelated. In addition, the market impact cost, coupled with the transactions costs, makes the stock return positively autocorrelated. Mathematically, after a sell order by the investor, his holdings reach the upper target level. It is more likely for the investor's holdings to reach the upper boundary than the lower boundary, where trading takes place. As a result, it is more likely for the investor to sell again, resulting in positive autocorrelations in both the investor's trades and the expected stock returns. Over a longer period of time, the investor's traders and the expected stock returns may exhibit negative autocorrelations due to the mean-reverting properties of the investor's holdings.

As part of the model solution, we endogenize the investor's trading frequency or the time interval of his two consecutive two trades. We find that the investor trades more frequently in more risky environment. For tractability, we assume that the diffusion terms of the dividend process and the investor's income stream process are constant. As a result, the expected investor's trading frequency is a constant. Using descriptive statistics, Biais, Hillion, and Spatt (1995) find that trading frequencies may be positively autocorrelated. It would be of interest to study this result by extending the current model.

Last, we test empirically the stock price structure developed from our theoretical model. Our empirical work supports the price structure. The previous empirical literature discovers the components associated with the market impact cost and the proportional cost in the stock price structure using small data sets. Our empirical study extend the previous literature by employing a much larger data set as well as uncovers a new component associated with the fixed cost in the stock price function.

In short, this paper discovers theoretically and empirically a new component in the stock price structure particular.

A Proof of Proposition 1

Conjecture that the investor's value function is of the form:

$$J(W,X,Z,t) = -\exp\left[-\rho t - r\gamma W - g_0 - g_1 X - \frac{1}{2}g_2 Z^2\right].$$

The investor's Bellman equation is then given by

$$\begin{aligned} 0 &= \max_{C,y} -\exp\left[-\rho t - \gamma C\right] + J_t + J_W\{rW - C + X - y^2\left[\lambda(a+r)\right]\} + \frac{1}{2}J_{WW}y^2\sigma_y^2 \\ &- a_x X J_X + \frac{1}{2}J_{XX}b_X^2 Z^2 - a_Z Z J_Z + \frac{1}{2}J_{ZZ}b_Z^2 + J_{WZ}b_Z(\lambda k b_Z + \eta\frac{b_D}{\alpha_D + r})y, \end{aligned} \quad (49)$$

where $J_W = -r\gamma J$, $J_{WW} = r^2\gamma^2 J$, $J_X = -g_1 J$, $J_t = -\rho J$, $J_{XX} = g_1^2 J$, $J_Z = -g_2 Z J$, $J_{ZZ} = [g_2^2 Z^2 - g_2]J$, $J_{WZ} = r\gamma g_2 Z J$, and $\sigma_y^2 = \frac{b_D^2}{(\alpha_D+r)^2} + \lambda^2 k^2 b_Z^2 + 2\eta k\lambda b_Z\frac{b_D}{\alpha_D+r}$.

The first-order conditions (FOCs) with respect to C and y yield

$$J_W = \gamma\exp\left[-\rho t - \gamma C_t\right], \quad (50)$$

$$0 = -J_W\left[2\lambda(a_Z + r)y\right] + J_{WW}\sigma_y^2 y + (\lambda k b_Z + \eta\frac{b_D}{\alpha_D + r})b_Z J_{WZ}, \quad (51)$$

respectively. Thus, we have

$$y = -\frac{(\lambda k b_Z + \eta\frac{b_D}{\alpha_D+r})b_Z g_2 Z}{2\lambda(a_Z + r) + r\gamma\sigma_y^2}. \quad (52)$$

Comparing equation (52) with $y = kZ$, we have

$$k = -\frac{(\lambda k b_Z + \eta\frac{b_D}{\alpha_D+r})b_Z g_2}{2\lambda(a_Z + r) + r\gamma\sigma_y^2}. \quad (53)$$

The Bellman equation is then reduced to

$$\begin{aligned} 0 &= (r-\rho) + r(g_0 + g_1 X + 0.5 g_2 Z^2 - \log r) - r\gamma X + a g_1 X \\ &+ 0.5 b_X^2 g_1^2 Z^2 + a_z g_2 Z^2 + 0.5 r\gamma b_Z g_2(\lambda k b_Z + \eta\frac{b_D}{\alpha_D + r})kZ^2 + 0.5(g_2^2 Z^2 - g_2)b_Z^2. \end{aligned}$$

Comparing the coefficients of constant , X, and Z^2 on both sides, we then have the solutions for g_0, g_1, and g_2, which satisfy

$$\begin{aligned} 0 &= (r-\rho) + r(g_0 - logr) - \frac{1}{2}b_Z^2 g_2, \\ g_1 &= \frac{r\gamma}{(r + a_X)}, \\ 0 &= (\frac{1}{2}r + a_Z)g_2 + \frac{1}{2}b_X g_1^2 + + \frac{1}{2}b_Z^2 g_2^2 + \frac{1}{2}r\gamma b_Z g_2 k(\lambda k b_Z + \eta\frac{b_D}{\alpha_D + r}) \end{aligned}$$

and k satisfies the following equation:

$$0 = b_Z g_2 \left(\lambda k b_Z + \frac{\eta b_D}{\alpha_D + r} \right) + k \left[2(a_Z + r)\lambda + r\gamma\sigma_y^2 \right].$$

When the market maker is risk neutral, we can reduce the above equations to

$$\begin{aligned} 0 &= (r + 2a_Z)g_2 + b_X^2 g_1^2 + b_Z^2(1-\eta^2)g_2^2, \\ g_1 &= \frac{r\gamma}{r + a_X}, \\ 0 &= (r - \rho) + r(g_0 - logr) - \frac{1}{2} b_Z^2 g_2^2. \end{aligned} \tag{54}$$

Note that g_2 satisfies a quadratic function that admits two solutions. Given the conjectured form of the value function, the larger, positive solution yields a higher expected utility for the investor. Thus, g_2 is give by

$$g_2 = \frac{\sqrt{(0.5r + a_Z)^2 - b_X^2 b_Z^2 (1-\eta^2) g_1^2} - (0.5r + a_Z)}{b_Z^2(1-\eta^2)},$$

where $(0.5r + a_Z)^2 - b_X b_Z (1-\eta)^2 g_1^2 \geq 0$. Because the solutions for this benchmark case form the foundation for our solutions to the small cost case, we next show that our solutions for g_1 and the larger positive g_2 are limiting solutions from a model of finite horizon $[0, T]$ in which $T \to \infty$.

Consider the model of finite horizon $[0, T]$. Conjecture that the investor's value function is of the form:

$$J(W, X, Z, t) = -\exp\left[-\rho t - f(t)W - g_0(t) - g_1(t)X - \frac{1}{2} g_2(t) Z^2 \right] - \exp[-\gamma W(T)],$$

with boundary conditions

$$g_0(T) = g_1(T) = g_2(T) = 0, \quad f(T) = \gamma.$$

Conjecture that the stock price function is given by

$$P_t = h_0(t) + h_1(t) D_t,$$

with boundaries $h_0(T) = h_1(T) = 0$. The excess dollar return is then given by

$$\begin{aligned} dQ &= dP - rPdt + Ddt, \\ &= \left\{ \left[h_0'(t) - r h_0(t) + \alpha_D h_1(t) \bar{D} \right] + \left[h_1'(t) - (r + \alpha_D) h_1(t) + 1 \right] \right\} D_t dt + h_1(t) \sigma_D dB_1. \end{aligned}$$

Since P_t is the present value of future dividends, the expected excess dollar return must be zero. We then have

$$
\begin{aligned}
0 &= h_0^{'}(t) - rh_0(t) + \alpha_D h_1(t)\bar{D}, \\
0 &= h_1^{'}(t) - (r + \alpha_D)h_1(t) + 1.
\end{aligned}
$$

Solving the above ODEs with the boundary conditions and letting $T \to \infty$ yields

$$
\begin{aligned}
h_0(t) &= \frac{\bar{D}}{r} - \frac{\bar{D}}{r + \alpha_D}, \\
h_1(t) &= \frac{1}{\alpha_D + r}.
\end{aligned}
$$

The investor's Bellman equation is given by

$$
\begin{aligned}
0 &= -\exp\left[-\rho t - \gamma C_t\right] + J_t + J_W\{rW - C + X\} + \frac{1}{2}J_{WW}y^2\sigma_y^2 \\
&- aXJ_X + \frac{1}{2}J_{XX}b_X^2 Z^2 - a_Z Z J_Z + \frac{1}{2}J_{ZZ}b_Z^2 + J_{WZ}b_Z\left(\eta\frac{b_D}{\alpha_D + r}\right)y,
\end{aligned}
\tag{55}
$$

where $J_W = -f(t)J$, $J_{WW} = f(t)^2 J$, $J_X = -g_1(t)J$, $J_t = -\left[\rho + f^{'}(t)W + g_0^{'} + g_1^{'}X(t) + \frac{1}{2}Z(t)^2\right]J$, $J_{XX} = g_1(t)^2 J$, $J_Z = -g_2(t)ZJ$, $J_{ZZ} = [g_2(t)^2 Z^2 - g_2(t)]J$, $J_{WZ} = f(t)g_2(t)ZJ$, and $\sigma_y^2 = h_1^2(t)b_D^2$.

The FOCs give:

$$
\begin{aligned}
J_W &= \gamma\exp\left[-\rho t - \gamma C_t\right], \\
0 &= J_{WW}\sigma_y^2 y + (\eta\frac{b_D}{\alpha_D + r})b_Z J_{WZ}.
\end{aligned}
$$

Thus, we have

$$
y = -\frac{\frac{\eta b_D b_Z}{\alpha_D + r}J_{WZ}}{J_{WW}\sigma_y^2} = -\frac{\eta b_Z g_2(t)}{(\alpha_D + r)f(t)h_1^2(t)b_D}Z. \tag{56}
$$

Plugging the FOCs into the Bellman equation (55) obtains

$$
\begin{aligned}
0 &= \left[\frac{f(t)}{\gamma} - \rho\right] + \frac{f(t)}{\gamma}\left[g_0 + g_1 X + 0.5g_2 Z^2 - \frac{\log f(t)}{\gamma}\right] - f(t)X + ag_1 X \\
&+ 0.5b_X^2 g_1(t)^2 Z^2 + a_z g_2(t)Z^2 - \frac{1}{2}f^2(t)h_1^2(t)b_D^2 y^2 + 0.5(g_2^2 Z^2 - g_2)b_Z^2 \\
&- \left[f^{'}(t)W + g_0^{'}(t) + g_1^{'}(t)X + \frac{1}{2}g_2^{'}Z^2\right] - f(t)\left(r - \frac{f}{\gamma}\right)W.
\end{aligned}
$$

Comparing the coefficients of W, X, and Z^2 on both sides yields

$$0 = -f^{'}(t) - f(t)\left[r - \frac{f(t)}{\gamma}\right], \tag{57}$$

$$0 = -g_1^{'} + \frac{f(t)g_1(t)}{\gamma} + ag_1(t) - f(t), \tag{58}$$

$$0 = 0.5g_2^{'}(t) + \frac{1}{2\gamma}f(t)g_2(t) + a_Z g_2 + 0.5b_X^2 g_1^2(t) - 0.5\frac{\eta^2 b_Z^2 g_2^2}{(\alpha_D + r)^2 h_1^2(t)} + 0.5g_2^2 b_Z^2. \tag{59}$$

We first solve $f(t)$. The equation (64) yields

$$\frac{\gamma df}{f(f - r\gamma)} = dt,$$

where $f(T) = \gamma$ and $0 < \gamma < 1$. Integrating both sides from $[0, T]$ yields

$$r(T - t) = log\frac{(r - r\gamma)}{\gamma} - log\frac{f(t) - r\gamma}{f(t)}.$$

Letting $T \to \infty$ obtains

$$f(t) = r\gamma.$$

We then solve for $g_1(t)$. Plugging the expression for $f(t)$, where $T \to \infty$, into equation (58) yields

$$-g_1^{'} + rg_1(t) + ag_1(t) - r\gamma = 0.$$

Let $H_1(t) = g_1(t)e^{-(r+a)t}$, we then have $H_1^{'} = -r\gamma e^{-(r+a)t}$ with boundary condition $H_1(T) = 0$. Therefore,

$$H_1(T) - H_1(t) = \frac{r\gamma}{r + a}\left[e^{-(r+a)T} - e^{-(r+a)t}\right].$$

Let $T \to \infty$. We then have

$$g_1(t) = \frac{r\gamma}{r + a}.$$

We next solve for $g_2(t)$. Plugging the expressions for $f(t)$ and $g_1(t)$, where $T \to \infty$, into equation (59) yields

$$-\frac{1}{2}g_2^{'}(t) + (r + 2a_Z)g_2 + b_X^2 + b_Z^2(1 - \eta^2)g_2^2 = 0.$$

Let $(r + 2a_Z)g_2 + b_X^2 + b_Z^2(1 - \eta^2)g_2^2 \equiv a_1(y - y_1)(y - y_2)$, where $y_1 < y_2 < 0$. Thus

$$\int \frac{dy}{(y - y_1)(y - y_2)} = \int a_1 dx.$$

Integrating both sides and letting $T \to \infty$, we arrive at

$$g_2(t) = y_2.$$

Last, we show that the other trading strategy $dy = \theta dt$, where θ is an unknown variable to be determined, is not optimal for the investor.

$dy = \theta dt$ **is not an optimal trading strategy**

Proof: We proceed by contradiction. Suppose that $dy = \theta dt$ is the investor's optimal trading strategy, the excess dollar return is then given by

$$dQ = (\lambda\theta - ry)dt + \frac{b_D}{\alpha_D + r} dB_1, \tag{60}$$

and the wealth process is given by

$$dW = (rW - C + X + \lambda\theta y - ry^2)dt + \frac{b_D y}{\alpha_D + r} dB_1. \tag{61}$$

Conjecture that the investor's value function is of the form:

$$J(W, y, X, t) = -\exp\left[-\rho t - r\gamma W - g_0 - g_1 X - \frac{1}{2} g_2 Z^2 - \frac{1}{2} h_2 y^2 - h_3 Z y\right].$$

The investor's Bellman equation is given by

$$\begin{aligned} 0 &= -\exp[-\rho t - \gamma C] + J_t + J_W\{rW - C + X + \lambda y\theta - ry^2\} + \frac{1}{2} J_{WW} y^2 \sigma_y^2 \\ &- a_X X J_X + \frac{1}{2} J_{XX} b_X^2 Z^2 + J_{WZ} b_Z \left(\eta \frac{b_D}{\alpha_D + r}\right) y + J_y \theta - a_Z Z J_Z + \frac{1}{2} J_{ZZ} b_Z^2, \end{aligned} \tag{62}$$

where $J_W = -r\gamma J$, $J_{WW} = r^2\gamma^2 J$, $J_X = -g_1 J$, $J_t = -\rho J$, $J_{XX} = g_1^2 J$, $J_Z = -(g_2 Z + h_3 y)J$, $J_{ZZ} = [(g_2 Z + h_3 y)^2 - g_2^2]J$, $J_{WZ} = r\gamma(g_2 Z + h_3 y)J$, $J_y = -(h_2 y + h_3 Z)J$, and $\sigma_y^2 = \frac{b_D^2}{(\alpha_D + r)^2}$. The FOC with respect to consumption C gives

$$C = -\frac{1}{\gamma} logr + rW + \frac{1}{\gamma}(g_0 + g_1 X + 0.5 g_2 Z^2 + 0.5 h_2 y^2 + h_3 Z y). \tag{63}$$

To guarantee the existence of a finite θ, the FOC with respect to θ gives

$$-r\gamma\lambda y - h_2 y - h_3 Z = 0. \tag{64}$$

We discuss the solution in different cases. The first case is that $r\gamma\lambda + h_2 \neq 0$. Solving equation (64) yields $y = -\frac{h_3}{r\gamma\lambda+h_2}Z$. However, it means that

$$dy = -\frac{h_3}{r\gamma\lambda + h_2}(-a_Z Z dt + b_Z dB_3), \tag{65}$$

which has a Brownian motion component, contradicting the supposed trading strategy $dy = \theta dt$. The second case is that $r\gamma\lambda + h_2 = 0$ and $h_3 = 0$. Plugging them back into the Bellman equation and comparing the coefficient of y^2 yields

$$r^2\gamma + \frac{1}{2}rh_2 + \frac{(r\gamma)^2 b_D^2}{2(\alpha_D + r)^2} = r^2\gamma + \left(1 - \frac{1}{2r^2}\right)\frac{(r\gamma)^2 b_D^2}{2(\alpha_D + r)^2} \neq 0, \tag{66}$$

for general parameter values. Therefore, there does not exist an optimal strategy in the form of $dy = \theta dt$.

B A Competitive Equilibrium

Conjecture that

$$y_t = kZ_t. \tag{67}$$

Assume that market makers employ a linear pricing rule given by

$$P_t = \frac{\bar{D}}{r} + \frac{D_t - \bar{D}}{\alpha_D + r} + \lambda k Z, \tag{68}$$

where Δy_t is the order sent by investors.

In the following part, we omit subscript 't'. Applying Ito's lemma to y_t yields

$$dy = -a_Z k Z dt + k b_Z dB_3.$$

The budget constraint of the investor is of the form:

$$dW = (rW - C + X)dt + ydQ, \tag{69}$$

where $dQ = dP - rPdt + Ddt = [-\lambda(a_Z + r)kZ]\,dt + \frac{b_D}{\alpha_D+r}dB_1 + k\lambda b_Z dB_3$.

Conjecture that the investor's value function is of the form:

$$J(W, X, Z, t) = -\exp\left[-\rho t - r\gamma W - g_0 - g_1 X - \frac{1}{2}g_2 Z^2\right].$$

The investor's Bellman equation is then given by

$$
\begin{aligned}
0 &= -\exp\left[-\rho t - \gamma C\right] + J_t + J_W\{rW - C + X - y\left[k\lambda(a+r)\right]\} + \frac{1}{2}J_{WW}y^2\sigma_y^2 \\
&- aXJ_X + \frac{1}{2}J_{XX}b_X^2Z^2 - a_Z Z J_Z + \frac{1}{2}J_{ZZ}b_Z^2 + J_{WZ}b_Z(\lambda k b_Z + \eta\frac{b_D}{\alpha_D + r})y, \qquad (70)
\end{aligned}
$$

where $J_W = -r\gamma J,\ J_{WW} = r^2\gamma^2 J,\ J_X = -g_1 J,\ J_t = -\rho J,\ J_{XX} = g_1^2 J,\ J_Z = -g_2 Z J,$ $J_{ZZ} = [g_2^2 Z^2 - g_2]J,\ J_{WZ} = r\gamma g_2 Z J$, and $\sigma_y^2 = \frac{b_D^2}{(\alpha_D+r)^2} + \lambda^2 k^2 b_Z^2 + 2\eta k\lambda b_Z \frac{b_D}{\alpha_D+r}$.

The first-order conditions (FOCs) give:

$$
\begin{aligned}
J_W &= \gamma \exp\left[-\rho t - \gamma C_t\right], \qquad (71) \\
0 &= -J_W\left[k\lambda(a_Z + r)Z\right] + J_{WW}\sigma_y^2 y + (\lambda k b_Z + \eta\frac{b_D}{\alpha_D + r})b_Z J_{WZ}. \qquad (72)
\end{aligned}
$$

Thus, we have

$$
y = -\frac{(\lambda k b_Z + \eta\frac{b_D}{\alpha_D+r})b_Z g_2 + \lambda k(a_Z + r)}{r\gamma\sigma_y^2}Z. \qquad (73)
$$

Comparing equation (67) with equation (73), we have

$$
k = -\frac{(\lambda k b_Z + \eta\frac{b_D}{\alpha_D+r})b_Z g_2 + \lambda k(a_Z + r)}{r\gamma\sigma_y^2}. \qquad (74)
$$

The Bellman equation is then reduced to

$$
\begin{aligned}
0 &= (r-\rho) + r(g_0 + g_1 X + 0.5 g_2 Z^2 - \log r) - r\gamma X + a g_1 X \\
&+ 0.5 b_X^2 g_1^2 Z^2 + a_z g_2 Z^2 + 0.5(g_2^2 Z^2 - g_2)b_Z^2 - \frac{1}{2}(r\gamma)^2\sigma_y^2 k^2 Z^2.
\end{aligned}
$$

Comparing the coefficients of constant, X, and Z^2 on both sides, we then have the solutions for g_0, g_1, and g_2, which satisfy

$$
\begin{aligned}
0 &= (r-\rho) + r(g_0 - logr) - \frac{1}{2}b_Z^2 g_2, \qquad (75) \\
g_1 &= \frac{r\gamma}{(r + a_X)}, \qquad (76) \\
0 &= \left(\frac{1}{2}r + a_Z\right)g_2 + \frac{1}{2}b_X g_1^2 + \frac{1}{2}b_Z^2 g_2^2 - \frac{1}{2}(r\gamma)^2\sigma_y^2 k^2. \qquad (77)
\end{aligned}
$$

where k is given in equation (74).

Remark. If $\eta = 0$, we see that the investor's demand for stock is zero. Intuitively, the investor cannot hedge his income shock using the stock. Notice that there is no closed-form solution for k.

C Proof of Theorems 2

The proof consists of three steps. We first define the quasi-variational inequalities (QVI) for the optimization problem. We next show that under some regularity conditions, there exists a solution to the QVI. The third step is to show that the solution is optimal. The proof is similar to Korn (1998), Cadenilias and Zapataro (2000), and Lo, Maminsky and Wang (2002).

C.1 Quasi-variational Inequalities (QVI)

Definition (QVI) We say that a function $J(M, D, y, X, Z, t)$ satisfies the quasi-variational inequalities for optimization problem for all δ and C,

Definition A.1 The value function in this case satisfies the quasi-variational inequalities conditions

$$\begin{aligned} D[J] - \exp\left[-\rho t - \gamma C(t)\right] &\leq 0, \\ T[J] &\leq J, \\ \left[J - T[J]\right]\left(sup_{c(t)}\{D[J] - \exp\left[-\rho t - \gamma C(t)\right]\}\right) &= 0 \end{aligned}$$

where $D[J] = J_t + J_M\{rM - C + X + Dy\} + J_D\left[-\alpha_D(D - \bar{D})\right] + \frac{1}{2}J_{DD}\sigma_D^2 - aXJ_X + \frac{1}{2}J_{XX}b_X^2Z^2 - a_ZZJ_Z + \frac{1}{2}b_Z^2J_{ZZ} + J_{DZ}\eta b_D b_Z$, $T[.]$ is defined as

$$T[J] = Sup_{\Delta} J\left[M(t) - \Delta(t)(\bar{P} + \lambda(y + \delta)) - (B + B_m) - \delta(A + A_m) - \frac{1}{2}\lambda\delta^2, y + \delta, Z(t)\right].$$

Perhame (1984) has studies the existence and uniqueness of a similar QVI. A solution J of the QVI problem separates the state space $S = (M, D, y, X, Z, t)$ into two disjoint regions: a continuation (non-trading) region NT and a trading region T:

$$\begin{aligned} T &\equiv \{s : J = T[J], D[J] - \exp\left[-\rho t - \gamma C(t)\right] \leq 0\}, \\ NT &\equiv \{s : J \geq T[J], D[J] - \exp\left[-\rho t - \gamma C(t)\right] = 0\}. \end{aligned}$$

Definition A.2 Let J be a solution of the QVI. The following mixed classical impulse stochastic control:

$$(C, T_k, \delta_k) = (c;\ \tau_1,\ \tau_2,\ \cdots,\ \tau_n,\ \cdots;\ \delta_1,\ \delta_2,\ \cdots,\ \delta_n,\ \cdots)$$

is called the QVI-control associated with it if

$$
\begin{aligned}
C &= arg\sup\{D[J] - \exp\left[-\rho t - \gamma C\right]\} \quad a.s.\forall S \in NT, \\
\tau_k &= \inf\{t > \tau_{k-1} : J = T[J]\}, \\
\delta_k &= argmaxT[J],
\end{aligned}
$$

where C is consumption policy.

In the non-trading region, the informed investor's position in the stock does not adjust. The investor chooses optimal consumption policy:

$$
0 = \sup_{\{C(t)\}} (D[J] - \exp -\rho t - \gamma C) = 0. \tag{78}
$$

Suppose $J = \exp(-\rho t)I$. The optimal consumption is then given by

$$
C = -\frac{1}{\gamma}\left[Lnr - LnI_M\right]. \tag{79}
$$

Conjecture that the investor's value function is of the form:

$$
\begin{aligned}
J(M, D, y, X, Z, t) &\equiv \exp\{-\rho t\} I(M, D, y, X, Z), \\
I(M, D, y, X, Z) &= -\exp\{-r\gamma\left[M + y(\bar{P} + 0.5\lambda y)\right] - g_0 - g_1 X - 0.5 g_2 Z^2 - V(y, Z)\},
\end{aligned}
$$

where $\bar{P} = \frac{\bar{D}}{r} + \frac{D-\bar{D}}{\alpha_D + r}$, M is the investor's position in bond, g_0 and g_2 are defined for the risk-neutral market maker and without transactions costs, and the price is given in equation (16). The QVI condition is then given by

$$
G[V(Z, y)] = a_1(V_Z^2 - V_{ZZ}) + a_2(y - y^{**})V_Z + a_3 V + a_4(y - y^{**})^2 + a_5 Z V_Z + a_6 y^2 \leq 0, \tag{80}
$$

where a_1 through a_6 are defined in equation (25).

$$
T\left[V(Z, \bar{y} + \Delta) - r\gamma\{(B + B_m) + (A + A_m)|\Delta y| + \frac{1}{2}(\Delta y)^2\}\right] \leq V(Z, \bar{y}), \tag{81}
$$

$$
G[V(Z, X)]\left\{\left[V(Z, \bar{y} + \Delta) - r\gamma\{(B + B_m) + (A + A_m)|\Delta y| + \frac{1}{2}(\Delta y)^2\right\}\right] - V(Z, \bar{y})\} = 0. \tag{82}
$$

Lemma 1. Let $F(y, Z)$ be a solution of the PDE with free boundary conditions for $y \in [KZ + D_B(Z), KZ + U_B(Z)]$ so that equation (80) holds for $KZ + D_B(Z) \leq \bar{y} \leq KZ + U_B(Z)$

$$
\begin{aligned}
&\left\{\left[F(Z, \bar{y} + \Delta) - r\gamma\{(B + B_m) + (A + A_m)|\Delta y| + \frac{1}{2}(\Delta y)^2\}\right] - F(Z, \bar{y})\right\} < 0, \\
&\text{If } F' - r\gamma[(A + A_m)sign(\Delta_y) + \lambda(\Delta_y)] = 0.
\end{aligned} \tag{83}
$$

Proof: Define $F(z)$ as

$$V(y,Z) \equiv \begin{cases} F(y^*,Z) - r\gamma[(B+B_m) + (y^*-y)(A+A_m) + \frac{\lambda}{2}(y^*-y)^2] \\ \text{if} \quad y < KZ + D_B(Z) \\ F(y,Z) \quad \text{if} \quad KZ + D_B(Z) \leq y \leq KZ + U_B(Z) \\ F(KZ+U_T(Z),Z) - r\gamma[(B+B_m) - (y^*-y)(A+A_m) + \frac{\lambda}{2}(y^*-y)^2] \\ \text{if} \quad y > KZ + U_B(Z) \end{cases}, \tag{84}$$

where y^* is the the solution to $max_{y_1}\left[F(y_1,Z) - r\gamma[(A+A_m)|y-y_1| + \frac{\lambda}{2}(y-y_1)^2\right]$ for any y. It is a solution to the QVI conditions (80)-(82). Now consider the proof of (81). Given the definition of y^*, it is optimal for the investor to trade whenever $y \leq KZ + D_B(Z)$ or $y \geq KZ + D_B(Z)$. As a result, we have $T[.] = V(Z,\bar{y})$ for any $y \leq KZ + D_B(Z)$ or $y \geq KZ + D_B(Z)$. Moreover, given the assumption about $F^{'}(Z,y)$ in $[KZ + D_B(Z), KZ + U_B(Z)]$, the optimal strategy for the investor is not to trade. Combining the above results leads to lemma 1.

C.2 Admissibility of the QVI

Now we show that the QVI given by the above solution is an admissible policy. We assume that the variance of the noise trader's demand follows a martingale. The QVI policies exist. The trading policy is totally captured by $kZ + U_B(Z), KZ + U_T(Z), KZ + D_T(Z), KZ + D_B(Z)$ and the trading time and trading amount are defined in the main text. We next show that the QVI-policy is admissible.

Lemma 2 *The QVI policy is admissible.*

Proof. Let us consider the case that $a = 0$. It can be seen that the QVI policies satisfy conditions (1)-(4) in definition 1. We only need to prove condition (5), namely,

$$N(S) \equiv E_0\left[\exp n(s)\right] < \infty.$$

There exists a constant v so that $\exp(1)\int_{\tau=0}^{v} P(\tau)d\tau < 0.5$, where $P(\tau)$ denotes the probability that the investor trades at time τ. Rearrangement yields

$$\begin{aligned} N(S) &= \int_{\tau=0}^{S} P(\tau)exp(1)N(S-\tau)d\tau \\ &= \int_{\tau=0}^{v} P(\tau)exp(1)N(S-\tau)d\tau + \int_{\tau=v}^{S} P(\tau)exp(1)N(S-\tau)d\tau \\ &\leq \frac{N(S)}{2} + N(S-v)\exp(1). \end{aligned}$$

Therefore, we have

$$N(S) \leq 2exp(1)N(S-v).$$

By iteration, since $N(0)$ is finite, $N(Tv)$ is finite, where T is finite. Since $N(S)$ is monotonically increasing, $N(S)$ is also finite for a finite S.

C.3 Optimality of the QVI-Policy

The last step is to prove that the QVI-policy is optimal. We first show that two conditions hold.

$$\lim_{t\to\infty} \exp(-rt)E_0\left[J(W,X,Z,y,t)\right] = 0, \tag{85}$$

and

$$E_0\left[\int_0^t (J_D \frac{y\sigma_D}{\alpha_D + r} + J_X Z\sigma_X + J_Z\sigma_Z)^2\right] < \infty. \tag{86}$$

Lemma 3. *For any admissible policy, equation (85) is satisfied.*

Proof. Suppose $t \in (t_{\tau_i}, t_{\tau_{i+1}})$, where t_{τ_i} denotes the ith trading. The Bellman equation yields:

$$0 = \exp -(\rho t - C_t) + E_t\left[dJ(W,X,y,t)\right]. \tag{87}$$

Because the optimal consumption policy satisfies

$$-\exp[-(\rho t - C_t)] = rJ, \tag{88}$$

taking the conditional expectation with respect to time t_{τ_i} yields

$$(\exp -r(t - t_{\tau_i}))J(M,D,X,y,Z,t_{\tau_i}) = E_{t_{\tau_i}}\left[J(M,D,X,y,Z,t)\right]. \tag{89}$$

From the value-matching condition, we have

$$J(M,D,X,y,Z,t_{\tau_i -}) = J(M,D,X,y,Z,t_{\tau_i}). \tag{90}$$

Let $\tau = \; t_{\tau_1}, \;\; ..., \;\; , t_{\tau_i}$. We then have

$$E\left[J(M,D,X,y,Z,t)|\mathcal{F}_0,\tau\right] = \exp(-rt)J(M,D,X,y,Z,0). \tag{91}$$

Taking the unconditional expectation with respect to τ yields

$$E\left[J(M,D,X,y,Z,t)|\mathcal{F}_0\right] = \exp(-rt)J(M,D,X,y,Z,0). \tag{92}$$

Because $J(M, D, X, y, Z, 0)$ is finite, we have

$$\lim_{t\to\infty} \exp(-rt) E_0 \left[J(M, D, X, y, Z, t) \right] = 0. \tag{93}$$

Lemma 4. *Equation (86) is satisfied. In other words,* $E\left[(J_D\sigma_D + J_X Z\sigma_X + J_Z\sigma_Z)^2\right] < \infty$ *is satisfied, where* $J_D = -r\gamma\frac{y\sigma_D}{\alpha_D+r}J$, $J_X = -g_1 J$, *and* $J_Z = -(g_2 Z + V_Z)J$.

Proof. Since $J = e^{-\rho t} I(M, D, y, X, Z, t)$. We only need to prove that $E_0[(I_D\sigma_D + I_X Z\sigma_X + I_Z\sigma_Z)^2] < \infty$, where $I_D = -r\gamma\frac{y\sigma_D}{\alpha_D+r}I$, $I_X = -g_1 I$ and $I_Z = -(g_2 Z + V_Z)I$. Since at any time t, $y - KZ \le \bar{Z}$ and $V_Z < \bar{F}$ (constant), we have

$$\begin{aligned}
&E_0\left[r\gamma\frac{y}{\alpha_D+r}\sigma_D + g_1 + (g_2 Z + V_Z)\right]^2 = E_0\left[r\gamma\frac{(y-KZ+KZ)}{\alpha_D+r}\sigma_D + g_1 + (g_2 Z + V_Z)\right]^2 \\
\le\;& 2E_0\left[r^2\gamma^2\frac{1}{(\alpha_D+r)^2}(y-KZ)^2\sigma_D^2 + r^2\gamma^2\left(\frac{KZ}{\alpha_D+r}\right)^2\sigma_D^2 + g_1^2 + g_2^2 Z^2 + V_Z^2\right] \\
\le\;& 2E_0\left[r^2\gamma^2\frac{1}{(\alpha_D+r)^2}\bar{Z}^2\sigma_D^2 + r^2\gamma^2\left(\frac{KZ}{\alpha_D+r}\right)^2\sigma_D^2 + g_1^2 + g_2^2 Z^2 + \bar{F}^2\right].
\end{aligned}$$

Since Z follows a normal distribution, $E_0\left[r\gamma\frac{y\sigma_D}{\alpha_D+r} + g_1 + (g_2 Z + V_Z)\right]^2 < \infty$. Given that $E_0\left[e^{-r\gamma W - g_0 - g_1 X - \frac{1}{2}g_2 Z^2}\right] < \infty$ and $V(y, Z) \ge \underline{V}$ at time t, $E_0\left[e^{-r\gamma W - g_0 - \frac{1}{2}g_2 Z^2 - V(y,Z)}\right] < \infty$ follows. By Holder's inequality,

$$\begin{aligned}
0 \;&\le\; E_0[(I_D\sigma_D + I_X Z\sigma_X + I_Z\sigma_Z)^2] \\
&\le\; \left(E_0[I^4] E_0[r\gamma\frac{y\sigma_D}{\alpha_D+r} + g_1 Z\sigma_X + (g_2 Z + V_Z)\sigma_Z]^2\right)^{1/2} < \infty,
\end{aligned} \tag{94}$$

we then arrive at

$$E_0\left[\int_0^t (J_D\frac{y\sigma_D}{\alpha_D+r} + J_X Z\sigma_X + J_Z\sigma_Z)^2\right] < \infty. \tag{95}$$

Lemma 5. *Suppose there exists a solution* $J^*(M, D, X, y, Z, t)$ *that satisfies the transversality conditions:*

$$\begin{aligned}
&E_0[\int_0^t (J_D\frac{y\sigma_D}{\alpha_D+r} + J_X Z\sigma_X + J_Z\sigma_Z)^2] < \infty, \\
&E_{t\to\infty}\left[J(M_t, D_t, X_t, y_t, Z_t, t)\right] = 0.
\end{aligned} \tag{96}$$

For each process (M_t, X_t, y_t, Z_t) *corresponding to admissible impulse controls, we have*

$$J(W_t, \theta_t, I_t, t) \le J^*(W_t, \theta_t, I_t, t), \forall (M_t, D_t, X_t, y_t, Z_t) \in (-\infty, \infty)^5.$$

Further, if there exists an admissible QVI-control, then it is an optimal impulse control and J^* is identical to the value functions.

Proof: We prove that the solution to the QVI, $J^*(M_0, D_0, X_0, y_0, Z_0, 0) = H(S_0)$, gives an upper bound on the value function of the optimization problem. From the QVI conditions, we have

$$\begin{aligned}
& \int_0^{t\wedge\tau_k} -\exp\left[-\varphi s - \gamma C_s\right] ds + \exp\left[-\varphi t \wedge \tau_k\right] H(t\wedge\tau_k) \\
= & \int_0^{t\wedge\tau_k} -\exp\left[-\varphi s - \gamma C_s\right] ds + H(S_0) + \int_0^{t\wedge\tau_k} \exp\left[-\varphi s\right](D[H] - \varphi H) ds \\
+ & \int_0^{t\wedge\tau_k} \exp\left[-\varphi s\right]\left[H_D \frac{y\sigma_D}{\alpha_D + r} dB_1 + H_X Z\sigma_X dB_2 + J_Z\sigma_Z\right] + \sum_{0\le k\le t} \exp\left[-\varphi t_k\right]\left[H(S_{t_k}) - H(S_{t_k^-})\right] \\
\le & \ H(S_0) + \int_0^{t\wedge\tau_k} \exp\left[-\varphi s\right]\left[H_D \frac{y\sigma_D}{\alpha_D + r} dB_1 + H_X Z\sigma_X dB_2 + J_Z\sigma_Z\right],
\end{aligned}$$

where $D[J] = H_M\{rM - C + X + Dy\} + H_D\left[-\alpha_D(D - \bar{D})\right] + \frac{1}{2}H_{DD}\sigma_D^2 - aXH_X + \frac{1}{2}H_{XX}b_X^2 Z^2 - a_Z Z H_Z + \frac{1}{2}b_Z^2 H_{ZZ} + H_{DZ}\eta b_D b_Z$. Taking expectations on both sides, we get

$$\begin{aligned}
& E_0\left[\int_0^{t\wedge\tau_k} -\exp\{-\varphi s - \gamma C_s\} ds\right] + E_0\left[\exp\{-\varphi(t\wedge\tau_k)\} H(t\wedge\tau_k)\right] \\
\le & \ H(S_0) + E_0 \int_0^{t\wedge\tau_k} \exp\left[-\varphi s\right]\left[H_D \frac{y\sigma_D}{\alpha_D + r} dB_1 + H_X Z\sigma_X dB_2 + J_Z\sigma_Z\right].
\end{aligned}$$

Applying equation (96) yields

$$\begin{aligned}
& E_0\left[\int_0^{t\wedge\tau_k} -\exp\{-\varphi s - \gamma C_s\} ds\right] + E_0\left[\exp\{-\varphi(t\wedge\tau_k)\} H(t\wedge\tau_k)\right] \\
\le & \ H(S_0).
\end{aligned}$$

Let $k \to \infty$ and take the limit of $t \to \infty$. Applying equation (96) again yields

$$E_0\left[\int_0^{\infty} -\exp\{-\varphi s - \gamma C_s\} ds\right] \le H(S_0)$$

for all admissible policies. Then we get

$$J(W, \theta, I, 0) \equiv \sup_{(c,\delta)\in\Theta} E_0\left[\int_0^{\infty} -\exp\{-\varphi s - \gamma C_s\} ds\right] \le H(S_0), \tag{97}$$

and the equality holds for the QVI policy. Thus the QVI policy is the optimal policy within the admissible set and the solution to the QVI policies gives the value function.

Remark. These two conditions rule out any arbitrage opportunities such as the "Ponzi schemes" and they are similar to those in Korn (1998) and Liu (2004).

Q.E.D.

D Proof of Theorem 3 and Proposition 3

Given Proposition 3, we know the trade size is close to a constant when the trading costs are small. As a result, the optimal trading strategy is symmetric: the distance from the upper boundary to the desired position is equal to the distance from the desired position to the lower boundary, and the distance from the upper target to the desired position is equal to the distance from the desired position to the lower target. Suppose that the signed trade size is given by x. When the trade size $|x|$ is a constant, the trades is captured by the following Markov chain

$$\begin{aligned} Prob(b_t|b_{t-1}) &= Prob(s_t|s_{t-1}) = h, \\ Prob(b_t|s_{t-1}) &= Prob(s_t|b_{t-1}) = 1-h, \end{aligned}$$

where $0 < h < 1$, $prob(b_t)$ denotes that the trade at tth trade is buy-initiated and $prob(s_t)$ denotes that the trade at $t-1$th trade is sell-initiated.

Since $lim_{n\rightarrow\infty}\begin{pmatrix} h & 1-h \\ 1-h & h \end{pmatrix}^n = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$, we have $Prob(b_t) = Prob(S_t) = 0.5$. Therefore, τ, the first-order autocorrelation of x, is given by

$$\tau = \frac{E[x_t, x_{t-1}]}{E[x_t^2]} = 2h - 1. \tag{98}$$

We now determine the value of h. The equivalent problem is a boundary problem with the upper boundary given by $\bar{b} = (x_1 - x_2)/k > 0$ and the lower boundary given by $\underline{b} = -(x_1 + x_2)/k$, where k, x_1 and x_2 are defined in the main text. By assumption, we have $dZ = b_Z dB_3$ and $Z_0 = 0$. Let $T_b = T_{\delta_0}$ be the first time that the process of Z hits $\bar{b}$ or $\underline{b}$, namely, $T_b = \{inf t > 0 : Z_t = \bar{b} or Z_t = \underline{b}\}$. Given the solution is symmetric, we have that $h = Prob(Z(T_b = \bar{b}))$. Let $T \wedge n = min\{T, n\}$. Since Z_t is a martingale, it can be shown that $f_t = Z_t^2 - tb_Z^2$ is a martingale. By iterated expectations, we have

$$E_t[f_{t+t_1}] = E_t[Z_{t+t_1}^2 - (t+t_1)b_Z^2] = Z_t^2 - tb_Z^2 = f_t. \tag{99}$$

Thus,

$$E[Z^2(T \wedge n)] = E[t \wedge n] b_Z^2 \leq max[\bar{b}, \underline{b}]^2 = \underline{b}^2,$$

or

$$E[T] = \lim_{n\rightarrow\infty} E[T \wedge n] \leq \bar{b}^2 / b_Z^2. \tag{100}$$

Let $\mu = Prob[X(T_b = \bar{b})]$. Since $E[T]$ is finite, Z_t is a martingale. Applying the optimal stopping time theorem yields

$$E[Z_T] = Z_0 = \mu\bar{b} + (1-\mu)\underline{b}.$$

Rearranging yields $\mu = \frac{-\underline{b}}{\bar{b}-\underline{b}} > 0.5$. Thus, $h = \mu > 0.5$. From equation (98), we know that the signed trade size of the investor is positively autocorrelated. Hence, we have

$$E[T_k] = \frac{x_1 x_2}{b_Z^2 k^2}.$$

References

Amihud, Y., and H. Mendelson, 1986, Liquidity and stock returns, *Financial Analysts Journal*, 42, 43-48.

B. Biais, P. Hilton, C. Spatt, 1995, An empirical analysis of the limit order book and the order flow in the Paris Bourse, *Journal of Finance*, 50, 1655.

Brennan, M.J., and A. Subrahmanyam, 1998, *Journal of Business*, 71, 1-25.

Cadenillas, A., and F. Zapatero, 2000, Classical and impulse stochastic control of the exchange rate using interest rates and reserves, *Mathematical Finance*, 10, 141-156.

Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam, 2001, Market liquidity and trading activity, *Journal of Finance*, 56, 501-530.

Chordia Tarun, Richard Roll and Avanidhar Subrahmanyam, 2002, Order Imbalance, Liquidity and Market Returns, *Journal of Financial Economics*, 65, 111-130.

Conrad, J. and K. Gaul, 1988, Time-variation in expected returns, *Journal of Business* 61, 409-425.

Constantinides, G.M., 1986, Capital market equilibrium with transaction costs, *Journal of Political Economy*, 94, 842-862.

Duffie, D., and T. Sun, 1990, Transaction costs and portfolio choice in a discrete time setting, *Journal of Economic Dynamic and Control*, 14, 35-51.

Dufour, A., and R. Engle, 2000, Time and the price impact of a trade, *Journal of Finance*, 55, 2467-2498.

Dumas, B., and E. Luciano, 1991, An exact solution to a dynamic portfolio choice problem under transactionc costs, *Journal of Finance*, 46, 577-595.

Fama, E.F., and K.R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246-273.

Foster, D., and S. Viswanathan, 1990, A theory of the interday variation in volume, variance, and trading costs in securities markets, *Review of Financial Studies*, 3, 593-624.

Glosten, L., and L. Harris, 1988, Estimating the components of the bid-ask spread, *Journal of Financial Economics*, 21, 123-142.

Guo, M., 2003, The optimal trading by an active investor, Working paper, Duke University.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance*, 46, 179-207.

Heaton, J., and D.J. Lucas, 1996, Evaluating the effects of incomplete markets on risk sharing and asset pricing, 1996, *Journal of Political Economy*, 104, 443-487.

Huang, M., 2003, Liquidity shocks and equilibrium liquidity premia, *Journal of Economic Theory*, 109, 104-129.

Korn, R., 1998, Portfolio optimization with strictly positive transition costs and impulse control, *Finance and Stochastic*, 2, 85-114.

Kyle, A.S., 1985, Continuous auctions and insider trading, *Econometrica*, 53, 1315-1336.

Lee, C., and Ready, M. 1991, Inferring trade direction from intradaily data. *Journal of Finance*, 46, 733-46.

Leland, H. E., 2000, Optimal portfolio implementation with transactions costs and capital gains taxes, working paper, U.C. Berkeley.

Leland, H. E., and D. Pyle, 1977, Information asymmetries, financial structure, and financial intermediation, *Journal of Finance*, 32, 371-387.

Lo, A.W., and A.C. MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 3, 175-205.

Madhaven, A., and S. Smidt, 1991, A Bayesian model of intraday specialist pricing, *Journal of Financial Economics*, 30, 99-134.

Lo, A.W., H. Mamaysky, and J. Wang, Asset prices and trading volume under fixed transactions costs, working paper, MIT and Yale.

Poterba, J.M., and L.H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, *Journal of Financial Economics* 22, 27-59.

Sadka, R., 2003, Liquidity risk and asset pricing, Working paper, Northwestern University.

Schroeder, M., 1998, Optimal portfolio selection with fixed transactions costs, working paper, Michigan State.

Stoll, H. R., Friction, 55, 4, 1479-1514.

Subrahmanyam, A., 1991, Risk aversion, market liquidity, and price efficiency, *Review of Financial Studies*, 4, 417-441.

Vayanos, D., 1998, Transactions costs and asset prices: A dynamic equilibrium model, *Review of Financial Studies*, 11, 1-58.

Vayanos, D., 2003, Flight to quality, flight to liquidity, and the pricing of risk, working paper, MIT.

Figure 1: The effects of the market impact cost λ (r = 0.03, $\gamma = 0.1$, $a_Z = 1$, $a_X = 0.01$, $\sigma_D = 0.2$, $\sigma_X = 10$, $\sigma_Z = 5$, and $\eta = 1$)

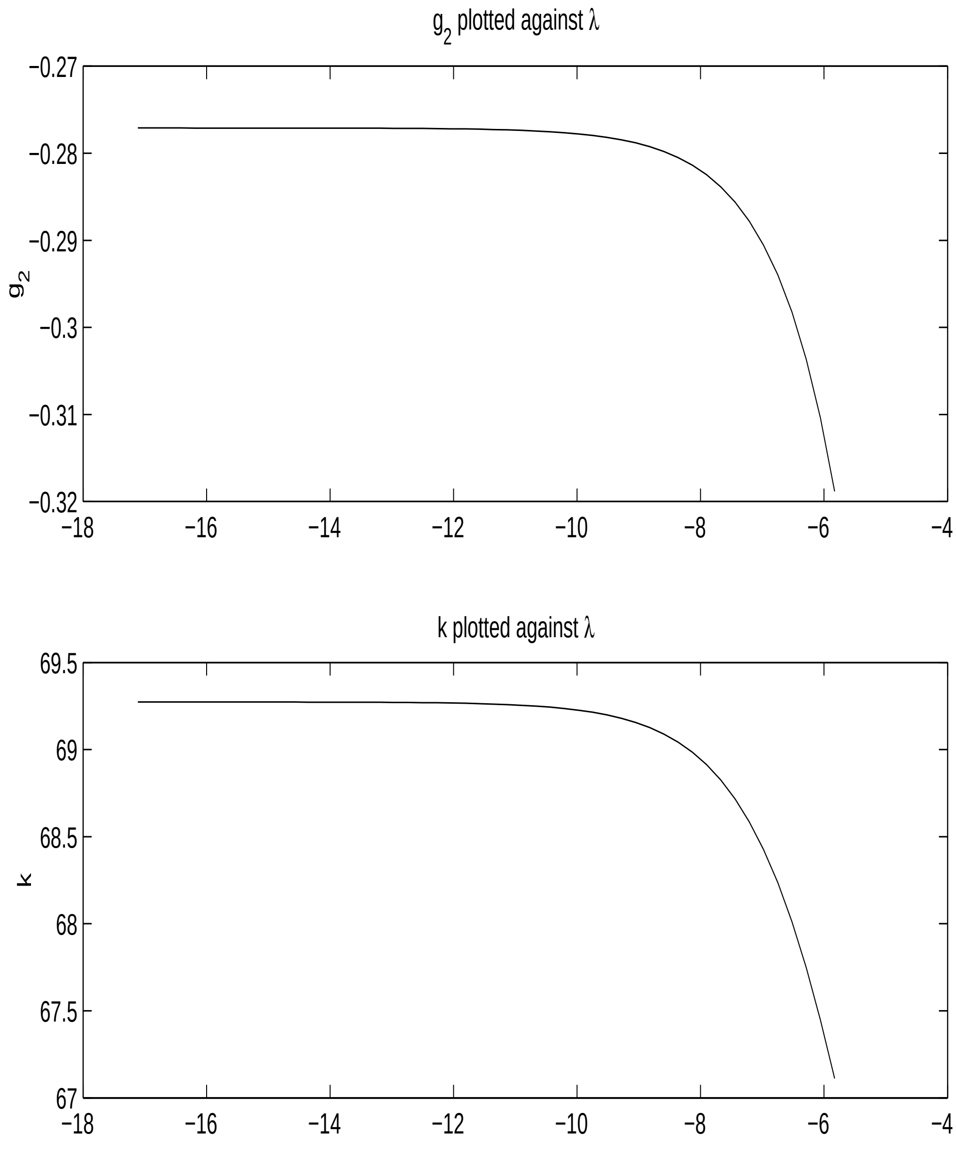
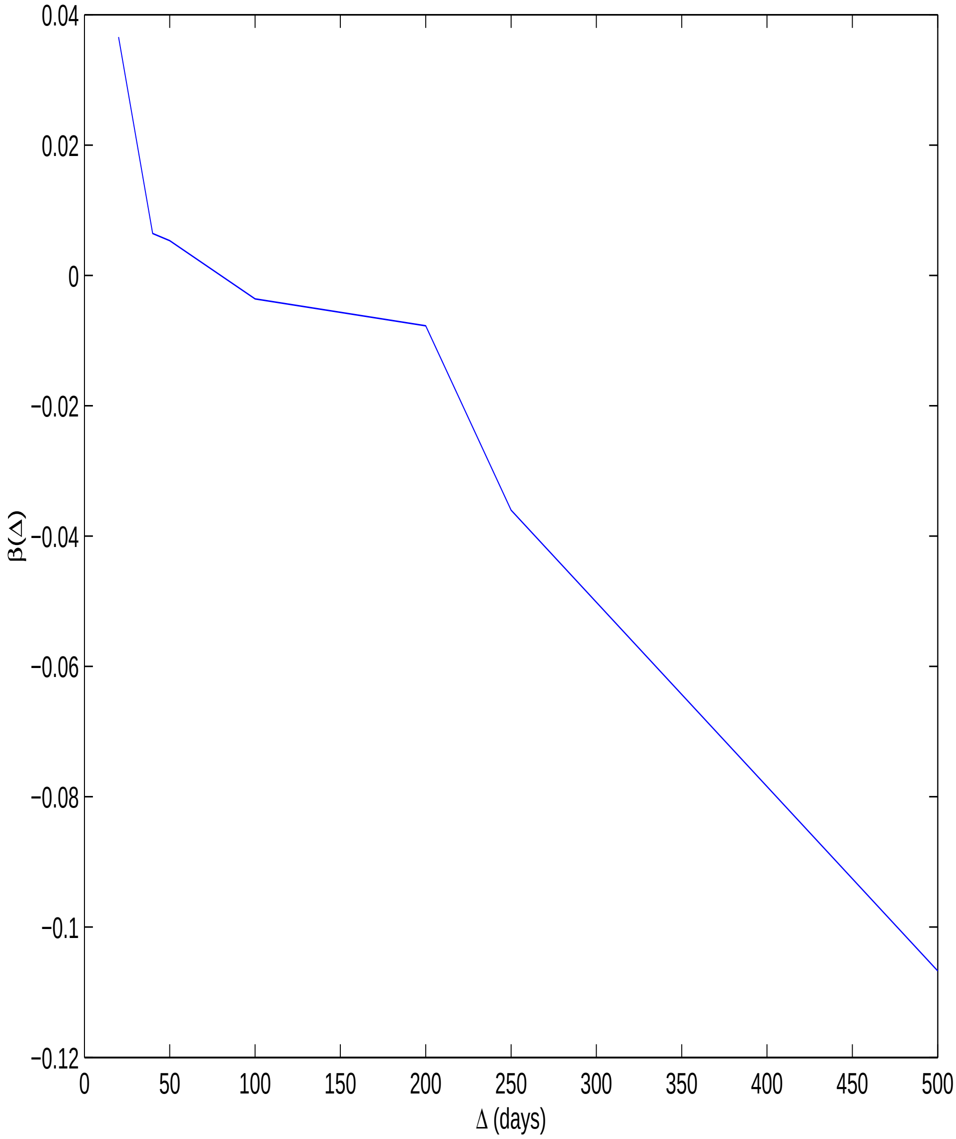


Figure 2: Order flow plotted against Δ ($r = 0.04$, $\gamma = 0.01$, $\sigma_X = 25$, $\sigma_Z = 50$, $a_X = 0.1$, $a_Z = 1$, $\sigma_D = 0.2$, $\epsilon = 0.000855$, $\eta = 1$, $\gamma_m = 10^{-6}$, $A = 50$, and $B = 500$)



.

Figure 3: Price change plotted against Δ ($r = 0.04$, $\gamma = 0.01$, $a_Z = 1$, $a_X = 0.1$, $\sigma_D = 0.5$, $\sigma_X = 25$, $\sigma_Z = 50$, $\epsilon = 0.000855$, $\eta = 1$, $\gamma_m = 10^{-6}$, $A = 50$, and $B = 500$)

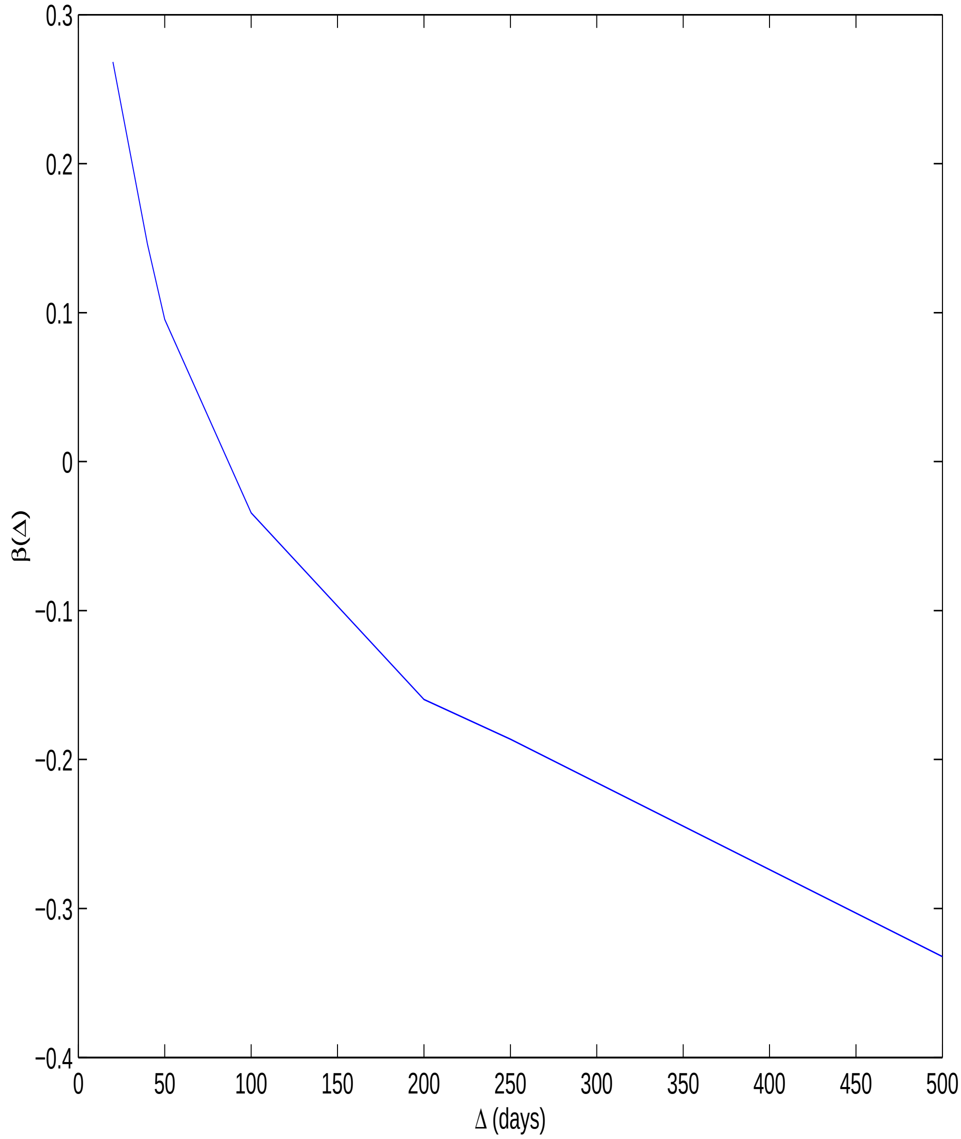


Figure 4: Market impact cost, expected trading interval (duration), and return volatility plot against σ_D ($r = 0.04$, $\gamma = 0.01$, $a_Z = 1$, $a_X = 0.1$, $\sigma_D = 0.5$, $\sigma_X = 25$, $\sigma_Z = 50$, $\epsilon = 0.000855$, $\eta = 1$, $\gamma_m = 10^{-6}$, $A = 50$, and $B = 500$)

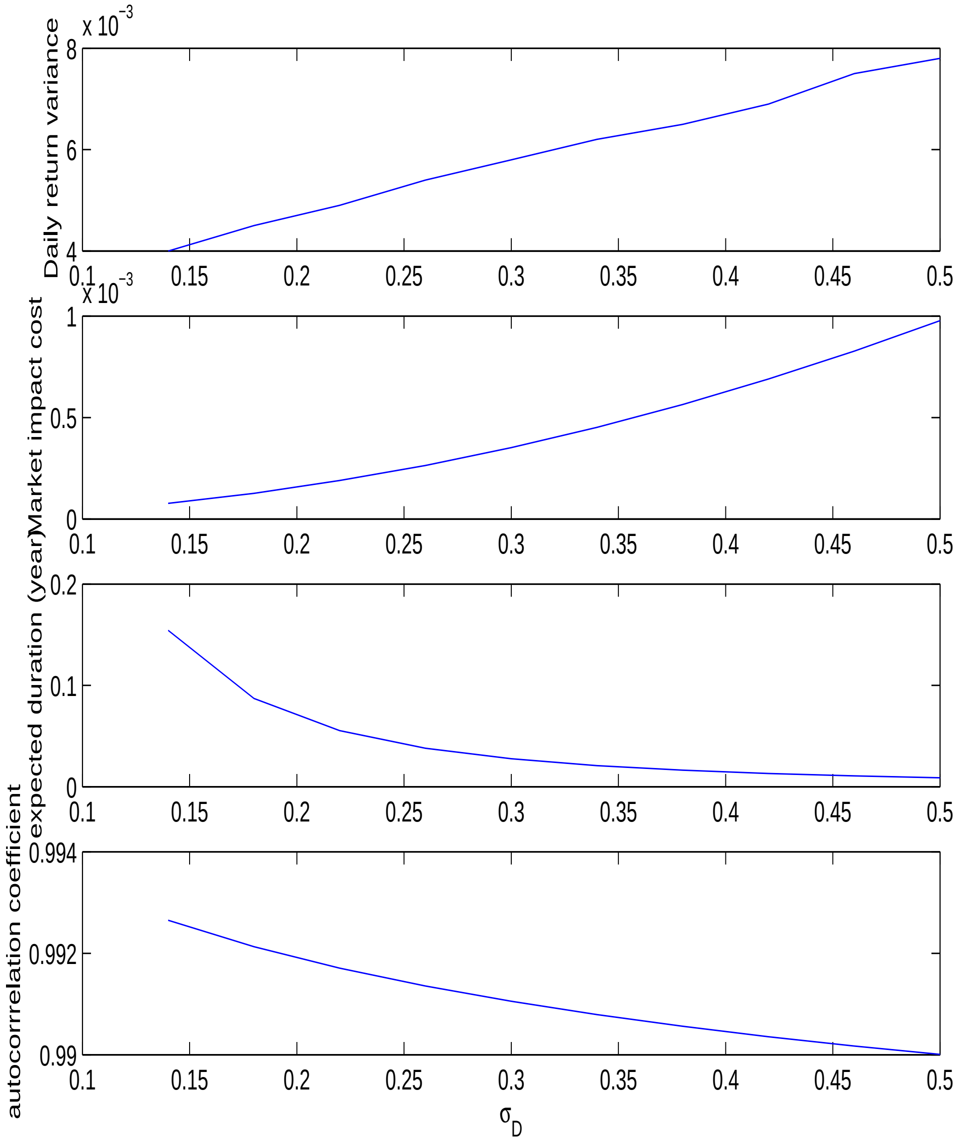


Table 1: The estimates of λ, 1993 - 1997

Quarter	Number of firms	λ	t ratio
1993/1	1409	9.12e-07**	4.8
1993/2	1402	-2.60e-07	-0.22
1993/3	1457	1.31e-06**	6.55
1993/4	1590	1.37e-06**	10.36
1994/1	1818	1.71e-06**	7.8
1994/2	1636	1.96e-06**	4.95
1994/3	1628	1.76e-06**	5.05
1994/4	1653	1.27e-06**	11.38
1995/1	1750	1.17e-06**	2.37
1995/2	1918	2.52e-06**	2.91
1995/3	2030	5.04e-07	0.56
1995/4	2083	1.79e-06**	2.53
1996/1	2190	1.27e-06**	7.85
1996/2	2485	1.10e-06**	4.78
1996/3	2204	1.60e-06**	12.69
1996/4	2404	3.02e-06*	1.8
1997/1	2451	1.12e-06**	8.4
1997/2	2274	1.62e-06**	10.53
1997/3	2495	1.34e-06**	12.73
1997/4	2363	1.62e-06**	8.99
mean	1962	1.44e-06	6.3405

λ is the estimated market impact cost from the following specification:

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + \frac{K}{\Delta y_t} - \frac{K}{\Delta y_{t-1}} + \epsilon_t,$$

where $\Delta y_t = y_t - y_{t-1}$ is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$. We assume that ϵ_t is identically and independently distributed.

Table 2: The estimates of λ, 1998 - 2002

Quarter	Number of firms	λ	t ratio
1998/1	2487	1.36e-06**	11.51
1998/2	2426	-1.10e-06	-0.47
1998/3	2133	1.81e-06**	8.33
1998/4	2115	1.40e-06**	2.92
1999/1	2308	2.09e-06**	6.16
1999/2	2325	2.00e-06**	8.82
1999/3	2317	1.62e-06**	4.43
1999/4	2405	2.58e-06**	8.04
2000/1	2724	2.80e-06**	10.85
2000/2	2491	2.27e-06**	11.54
2000/3	2380	1.58e-06**	5.45
2000/4	1951	2.15e-06**	7.62
2001/1	1879	2.42e-06**	11.8
2001/2	1814	2.23e-06**	13.8
2001/3	1718	1.95e-06**	19.06
2001/4	1649	-9.80e-06	-0.83
2002/1	1706	-7.30e-06	-0.64
2002/2	1802	1.99e-06**	11.38
2002/3	1542	1.91e-06	1.39
2002/4	1493	2.92e-06**	7.52
mean	2083	8.44e-07	7.434

λ is the estimated market impact cost from the following specification:

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + \frac{K}{\Delta y_t} - \frac{K}{\Delta y_{t-1}} + \epsilon_t,$$

where $\Delta y_t = y_t - y_{t-1}$ is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$. We assume that ϵ_t is identically and independently distributed.

Figure 5: The average λ plots against time: 1993- 2002

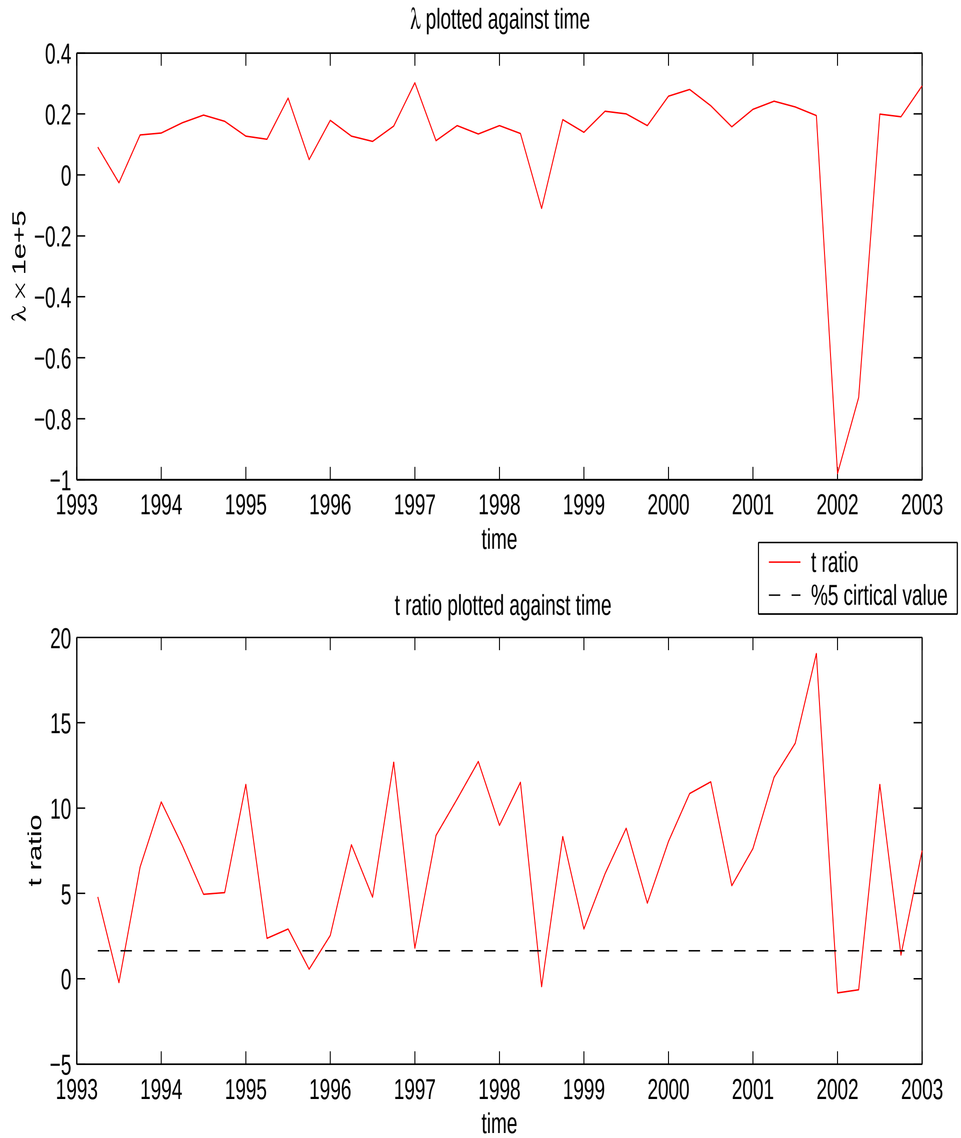


Table 3: The estimates of K, 1993 - 1997

Quarter	Number of firms	K	t ratio
1993/1	1409	5.2777**	12.11
1993/2	1402	2.3284**	0.86
1993/3	1457	5.6931**	28
1993/4	1590	5.309**	13.88
1994/1	1818	6.0898**	20.67
1994/2	1636	5.3529**	19.54
1994/3	1628	4.4573**	14.71
1994/4	1653	4.4548**	36.01
1995/1	1750	4.9247**	6.34
1995/2	1918	3.0777**	3.4
1995/3	2030	3.983**	22.07
1995/4	2083	3.7073**	5.07
1996/1	2190	4.9769**	5.29
1996/2	2485	3.628**	17.97
1996/3	2204	3.7136**	26.47
1996/4	2404	4.0874**	8
1997/1	2451	3.1031**	8.35
1997/2	2274	3.0872**	19.86
1997/3	2495	2.641**	19.99
1997/4	2363	2.2933**	23.19
mean	1962	4.10931**	15.589

K is the estimated fixed cost per trade from the following specification:

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + \frac{K}{\Delta y_t} - \frac{K}{\Delta y_{t-1}} + \epsilon_t,$$

where $\Delta y_t = y_t - yt - 1$ is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$. We assume that ϵ_t is identically and independently distributed.

Table 4: The estimates of K, 1998 - 2002

Quarter	Number of firms	K	t ratio
1998/1	2487	2.0384*	18.26
1998/2	2426	1.1939*	2.82
1998/3	2133	1.2833*	7.61
1998/4	2115	0.6085*	3.04
1999/1	2308	1.5517*	1.66
1999/2	2325	0.52**	3.27
1999/3	2317	0.0748	1.12
1999/4	2405	-0.269	-1.08
2000/1	2724	0.2319**	2.48
2000/2	2491	-0.003	-0.02
2000/3	2380	-0.307	-2.49
2000/4	1951	-0.204	-3.38
2001/1	1879	-0.248	-6.89
2001/2	1814	-0.408	-4.92
2001/3	1718	0.0784*	2.04
2001/4	1649	0.1839**	2.58
2002/1	1706	0.0455	0.31
2002/2	1802	0.0813	1.51
2002/3	1542	0.1745**	2.56
2002/4	1493	0.0594	1
mean	2083	0.334325	1.574

K is the estimated fixed cost from the following specification:

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + \frac{K}{\Delta y_t} - \frac{K}{\Delta y_{t-1}} + \epsilon_t$$

where $\Delta y_t = y_t - y_{t-1}$ is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$. We assume that ϵ_t is identically and independently distributed.

Figure 6: The average K plots against time: 1993- 2002

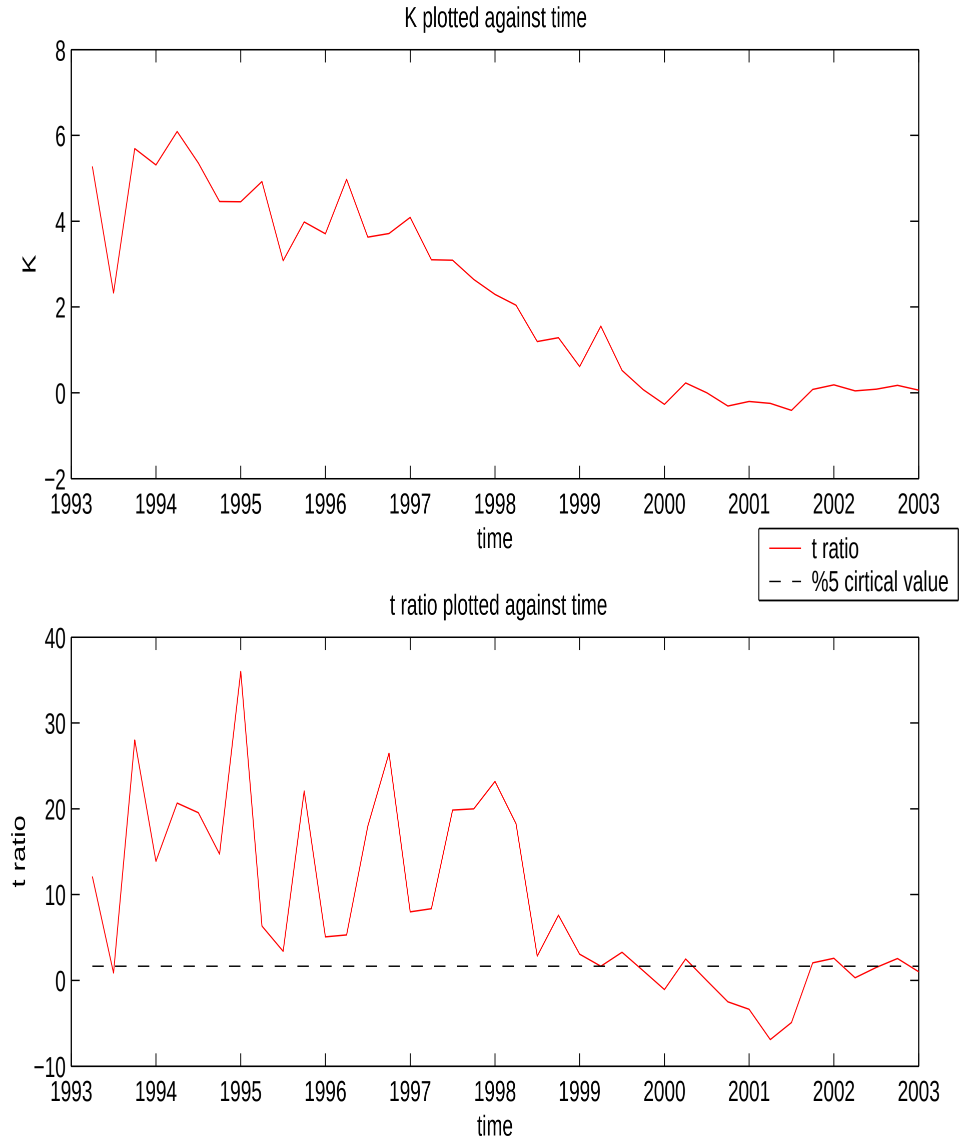


Table 5: The estimates of k, 1993 - 1997

Quarter	Number of firms	k	t ratio
Quarter Number of firms t ratio			
1993/1	1409	0.1388**	81.76
1993/2	1402	0.1375**	70.78
1993/3	1457	0.1326**	80.8
1993/4	1590	0.1327**	88.85
1994/1	1818	0.1314**	85.09
1994/2	1636	0.1276**	83.68
1994/3	1628	0.1227**	81.03
1994/4	1653	0.116**	93.41
1995/1	1750	0.1148**	65.74
1995/2	1918	0.1161**	60.66
1995/3	2030	0.1234**	92.66
1995/4	2083	0.1258**	67.32
1996/1	2190	0.1245**	80.04
1996/2	2485	0.1304**	84.19
1996/3	2204	0.1235**	89.34
1996/4	2404	0.1143**	55.51
1997/1	2451	0.1135**	88.65
1997/2	2274	0.1048**	85.43
1997/3	2495	0.0977**	86.08
1997/4	2363	0.0901**	87.19
mean	1962	0.12091**	80.41

k is the estimated proportional cost from the following specification:

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + \frac{K}{\Delta y_t} - \frac{K}{\Delta y_{t-1}} + \epsilon_t$$

where $\Delta y_t = y_t - y_{t-1}$ is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$. We assume that ϵ_t is identically and independently distributed.

Table 6: The estimates of k, 1998 - 2002

Quarter	Number of firms	k	t ratio
1998/1	2487	0.0812**	81.86
1998/2	2426	0.081**	28.95
1998/3	2133	0.0847**	63.83
1998/4	2115	0.0849**	72.39
1999/1	2308	0.0806**	44.38
1999/2	2325	0.0778**	62.38
1999/3	2317	0.0706**	77.19
1999/4	2405	0.0847**	54.57
2000/1	2724	0.107**	65.11
2000/2	2491	0.0886**	73.6
2000/3	2380	0.0712**	38.65
2000/4	1951	0.0711**	82.82
2001/1	1879	0.0578**	80.92
2001/2	1814	0.0396**	38.96
2001/3	1718	0.045**	61.68
2001/4	1649	0.0401**	63.42
2002/1	1706	0.0365**	28.26
2002/2	1802	0.0355**	55.76
2002/3	1542	0.0426**	7.04
2002/4	1493	0.0305**	47.06
mean	2083	0.06555	56.4415

k is the estimated proportional cost from the following specification:

$$\triangle P_t = \lambda \Delta y_t + k(v_t - v_{t-1}) + \frac{K}{\Delta y_t} - \frac{K}{\Delta y_{t-1}} + \epsilon_t$$

where $\Delta y_t = y_t - y_{t-1}$ is the order flow, $\epsilon = D_t - D_{t-1}$, and $v_t = sign(\Delta y_t)$. We assume that ϵ_t is identically and independently distributed.

Figure 7: The average k plots against time: 1993- 2002