2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Dynamic portfolio choice with market impact costs

Andrew E.B. Lim and Poomyos Wimonkittiwat

***Abstract*— Illiquidity and market impact refer to the situation where it may be costly or difficult to trade a desired quantity of assets over a desire period of time. In this paper, we formulate a simple model of dynamic portfolio choice that incorporates liquidity effects. The resulting problem is a stochastic linear quadratic control problem where liquidity costs are modeled as a quadratic penalty on the trading rate. Though easily computable via Riccati equations, we also derive a multiple time scale asymptotic expansion of the value function and optimal trading rate in the regime of vanishing market impact costs. This expansion reveals an interesting but intuitive relationship between the optimal trading rate for the "illiquid" problem and the classical Merton model for dynamic portfolio selection in perfectly liquid markets. It also gives rise to the notion of a "liquidity time scale" which shows how trading horizon and market impact costs affect the optimal trading rate.**

I. Introduction

Market impact and illiquidity refer to situations where it may be costly or difficult to trade a desired quantity of assets in a desired period of time. In this paper, we present a simple model of portfolio choice with liquidity effects. Our model is a linear-quadratic stochastic control problem where the objective is the sum of quadratic utility of terminal wealth and a new quadratic penalty on the trading rate which accounts for liquidity effects. It is a generalization of the classical Merton problem for dynamic portfolio choice [1] in which liquidity effects are ignored (see also the recent papers [2], [3], [4], [5], [6], [7], [8] for other models of liquidity in portfolio choice and risk management).

Several highlights of our paper are as follows. Firstly, we show that our model is equivalent to another problem where the objective is to trade at rate such that the risky asset holding tracks the optimal holding of the perfectly liquid Merton problem with minimal cost. This gives an intuitively appealing relationship between two approaches for accounting for illiquidity. Secondly, we derive a multiple time scale asymptotic expansion of the value function and optimal trading rate in the case of vanishing market impact costs, which allows us to understand the relationship between the "illiquid problem" and the perfectly liquid Merton problem. The asymptotic expansion shows that to a first order, the investor trades to decrease the gap between his/her current position and optimal holding for the Merton problem, and that the trading rate is increasing in the volatility of the risky asset, the risk-aversion of the investor, the market depth, and the remaining trading time. Additionally, the asymptotic expansion leads to the introduction of a "liquidity time scale" $(T-t)/\sqrt{\lambda}$ which captures the intuitive notion that the (remaining) trading horizon $T-t$ should be evaluated not only in "calendar time", but also in the context of market impact costs λ and the ease of trading.

An outline of the paper is as follows. We introduce the classical Merton problem in Section II for optimal dynamic portfolio selection in perfectly liquid markets. Our model of portfolio selection with market impact costs, which are modeled as a penalty on the trading rate, is introduced in Section III. This model is shown to be equivalent to another portfolio choice problem in Section IV, where the goal is to track the optimal portfolio of the (perfectly liquid) Merton problem with minimal market impact costs. In Section V, we derive a multiple time scale asymptotic expansion of the value function and optimal portfolio for the market impact problems in the regime of vanishing liquidity cost, which leads to the important and intuitive notion of the liquidity time scale. Examples are presented in Section VI.

Due to limitations in space, longer proofs have not been included. The interested reader is referred to [9].

II. Portfolio Selection Problem in Liquid Market

We recall the classical Merton problem [1] for frictionless markets.

Asset Dynamics

For simplicity, we consider a market with one risky asset and one risk-free asset. Our results can be extended to multiple assets with no essential difficulty. We model uncertainty using Brownian motion which is assumed to live on a filtered probability space $(\Omega, \mathscr{F}, P, \{\mathscr{F}_t\})$ over a finite time horizon $[0,T]$. The risky asset price $s(t)$ is assumed to follow geometric Brownian motion

$$ds(t) = \mu s(t)dt + \sigma s(t)dw(t), \qquad (1)$$

with expected return μ and volatility σ. The risk-free asset price process $s_0(t)$ satisfies

$$ds_0(t) = rs_0(t)dt$$

with risk-free rate of return r.

This research was supported by the NSF CAREER Award DMI-0348746 and the NSF Grants CMMI-0500503 and CMMI-1031637. The opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Support from the Coleman Fung Chair in Financial Modeling is also acknowledged.

Both authors are with the Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA. lim@ieor.berkeley.edu, poomyos@ieor.berkeley.edu

The Merton problem

Let $n(t)$ be the number of shares of the risky asset at time t. The wealth $x(t)$ of a self-financing investor satisfies

$$dx(t) = \{x(t)r + n(t)s(t)(\mu - r)\}dt + n(t)s(t)\sigma dw(t)$$

We rewrite the wealth dynamics in terms of the dollar value $\pi(t) \triangleq n(t)s(t)$ of the risky asset holding

$$dx(t) = \Big\{x(t)r + \pi(t)(\mu - r)\Big\}dt + \pi(t)\sigma dw(t). \quad (2)$$

The classical Merton problem maximizes expected utility of terminal wealth

$$\begin{cases} \sup\limits_{\pi(\cdot)} E\{\Phi(x(T))\} \\ \text{subject to:} \\ dx(t) = \{x(t)r + \pi(t)(\mu - r)\}dt + \pi(t)\sigma dw(t) \\ x(0) = x_0. \end{cases} \quad (3)$$

The value function $V(t,x)$ for (3) is the solution of the dynamic programming equations

$$\begin{cases} V_t + \sup\limits_{\pi}\Big\{xrV_x + \pi(\mu - r)V_x + \frac{1}{2}\pi^2\sigma^2 V_{xx}\Big\} = 0 \\ V(T,x) = \Phi(x(T)). \end{cases} \quad (4)$$

Explicit solutions for the Merton problem and the associated dynamic programming equation can be found when the utility function is of power, exponential, logarithmic and quadratic type. In the case of quadratic utility the Merton problem is a linear quadratic problem and we have the following result.

Proposition 2.1: The value function for the Merton problem (3) with quadratic utility function $\Phi(x(T)) = x(T) - \frac{\eta}{2}x(T)^2$ is

$$V_M(t,x) = -\frac{1}{2}A_M(t)x^2 + B_M(t)x + C_M(t) \quad (5)$$

where

$$A_M(t) = \eta e^{(2r - \frac{(\mu - r)^2}{\sigma^2})(T-t)},$$
$$B_M(t) = e^{(r - \frac{(\mu - r)^2}{\sigma^2})(T-t)},$$
$$C_M(t) = \frac{1}{2\eta}(1 - e^{-(\frac{\mu - r}{\sigma})^2(T-t)}).$$

The optimal investment policy is

$$\pi_M^*(t,x) = \frac{\mu - r}{\sigma^2}\Big(\frac{B_M(t)}{A_M(t)} - x\Big). \quad (6)$$

(The subscript M is added for future reference.)

III. Illiquid Portfolio Selection: Model I

In this section, we account for liquidity effects by formulating a modification of the Merton problem in which fast trading is expensive. In this model, trading rate is the control variable while the wealth and the risky asset position are the states.

Trading rate

Let $n(t)$ denote the number of shares in the risky asset, n_0 be the initial risky asset holding, and $\rho(t)$ denote the rate at which shares in the risky asset are being purchased at time t, which is controlled by the investor. The risky asset holding satisfies

$$dn(t) = \rho(t)dt, \; n(0) = n_0. \quad (7)$$

As in the Merton problem (3) let $\pi(t) \triangleq n(t)s(t)$ denote the dollar value of the risky asset holding. In contrast to the Merton problem, where $\pi(t)$ is the control variable, the assumption that $n(t)$ satisfies (7) means that $\pi(t)$ can only be controlled through the trading rate $\rho(t)$ and needs to be treated as a state. Ito's formula together with (1) and (7) imply that

$$d\pi(t) = \{\rho(t)s(t) + \pi(t)\mu\}dt + \pi(t)\sigma dw(t).$$

Defining $\bar{\rho}(t) \triangleq \rho(t)s(t)$, it follows that

$$d\pi(t) = \{\bar{\rho}(t) + \pi(t)\mu\}dt + \pi(t)\sigma dw(t) \quad (8)$$

where $\bar{\rho}(t)$ can be interpreted as the trading rate, in dollars per unit time, of the risky asset. The equation (8) tells us that changes in the dollar value of the risky asset holding equals the increase due to inflow at rate $\bar{\rho}(t) = \rho(t)s(t)$ per unit time from the money market account and the change $\pi(t)\mu dt + \pi(t)\sigma dw(t)$ due to fluctuations in the value of assets already being held. The self-financing condition implies that the wealth process $x(t)$ remains unchanged from the Merton problem and is given by (2).

Liquidity costs

To incorporate the idea that it is costly to trade large quantities of an asset in small periods of time, we consider the problem

$$\begin{cases} \sup\limits_{\bar{\rho}(\cdot)} E\Big\{-\int_0^T \frac{\lambda}{2}\bar{\rho}(t)^2 dt + \Phi(x(T))\Big\} \\ \text{subject to:} \\ dx(t) = \{x(t)r + \pi(t)(\mu - r)\}dt + \pi(t)\sigma dw(t) \\ d\pi(t) = \{\bar{\rho}(t) + \pi(t)\mu\}dt + \pi(t)\sigma dw(t) \\ \bar{\rho}(\cdot) \in \mathscr{A}, \, x(0) = x_0, \, \pi(0) = \pi_0 \end{cases} \quad (9)$$

where $\mathscr{A}$ is the class of admissible controls defined as

$$\mathscr{A} = \Big\{\bar{\rho} : [0,T] \times \Omega \to \mathbb{R} \;\Big|\; \text{such that } \bar{\rho}(\cdot) \text{ is } \{\mathscr{F}_t\}\text{-adapted and } E\int_0^T |\rho(t)|^2 dt < \infty\Big\}$$

Aside from the dynamics, which we have already discussed, a key modification relative to the Merton problem (3) is the introduction of the quadratic penalty on the trading rate

$$E\int_0^T \frac{\lambda}{2}\bar{\rho}(t)^2 dt \quad (10)$$

in the objective. Optimizing (9) involves a trade-off between maximizing expected utility $E[\Phi(x(T))]$ and minimizing the cost of trading. The quadratic $|\bar{\rho}(t)|^2$ in (10) means that marginal cost of trading increases in the trading rate. The *illiquidity coefficient* $\lambda > 0$ is large when illiquidity frictions are large. Similar models for liquidity costs have been proposed for dynamic models of active portfolio investment ([10], [11]). To our knowledge, this paper is the first in which such a model is used in the context of a Merton-type problem. It can be shown that the value function for (9) satisfies the dynamic programming equations

$$\begin{cases} V_t + \sup_{\bar{\rho}} \Big\{ -\frac{1}{2}\lambda\bar{\rho}^2 + xrV_x + \pi(\mu - r)V_x + \pi\mu V_\pi + \bar{\rho}V_\pi \\ +\frac{1}{2}\pi^2\sigma^2(V_{xx} + 2V_{x\pi} + V_{\pi\pi})\Big\} = 0 \\ V(T,x,\pi) = \Phi(x). \end{cases} \tag{11}$$

When the utility function is quadratic, the value function and the optimal trading rate can be characterized as follows.

Proposition 3.1: Suppose that the liquidity cost parameter λ is positive and the utility function is quadratic with risk-aversion parameter $\eta > 0$:

$$\Phi(x) = x - \frac{\eta}{2}x^2.$$

Then the value function for the portfolio selection problem (9) is

$$V(t,x,\pi) = -\frac{1}{2}\begin{bmatrix} x & \pi \end{bmatrix}\begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{12}(t) & A_{22}(t) \end{bmatrix}\begin{bmatrix} x \\ \pi \end{bmatrix} + \begin{bmatrix} B_1(t) & B_2(t) \end{bmatrix}\begin{bmatrix} x \\ \pi \end{bmatrix} + C(t), \tag{12}$$

and the optimal trading rate satisfies

$$\bar{\rho}^*(t,x,\pi) = \frac{1}{\lambda}(B_2(t) - A_{12}(t)x - A_{22}(t)\pi). \tag{13}$$

The coefficients

$$A(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{12}(t) & A_{22}(t) \end{bmatrix}, B(t) = \begin{bmatrix} B_1(t) \\ B_2(t) \end{bmatrix} \text{ and } C(t)$$

are the solutions to the system of ODEs:

$$\begin{cases} \dot{A}(t) + R'A(t) + A(t)R + \Sigma'A(t)\Sigma - \frac{1}{\lambda}A(t)SS'A(t) = 0 \\ A(T) = \begin{bmatrix} \eta & 0 \\ 0 & 0 \end{bmatrix} \end{cases} \tag{14}$$

$$\begin{cases} \dot{B}(t) + R'B(t) - \frac{1}{\lambda}A(t)SS'B(t) = 0 \\ B(T) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{cases} \tag{15}$$

$$\begin{cases} \dot{C} + \frac{1}{2\lambda}(S'B(t))^2 = 0 \\ C(T) = 0 \end{cases} \tag{16}$$

where $R = \begin{bmatrix} r & \mu - r \\ 0 & \mu \end{bmatrix}, \Sigma = \begin{bmatrix} 0 & \sigma \\ 0 & \sigma \end{bmatrix}, S = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Although the value function and optimal trading rate for (9) can be characterized and easily computed numerically when the utility function $\Phi(x)$ is quadratic, the expressions (12)-(16) are not particularly insightful. For instance, it is by no means clear how the presence of market impact effects in (9) modify the value function and optimal policy relative to the solution of the Merton problem. To clarify these issues we introduce an alternative formulation of the market impact problem which provides us a simple and intuitive relationship between the market impact problem and the Merton problem.

IV. Illiquid Portfolio Selection: Model II

Model II

Let $q : [0, T] \to (0, \infty)$ be a deterministic strictly positive function of time, $\xi(t)$ denote the trading rate, $x(t)$ the investor's wealth, and $\pi(t)$ his position in the risky asset. Consider the problem

$$\begin{cases} \inf_{\xi(\cdot)} E\Big\{\int_0^T \frac{\lambda}{2}\xi(t)^2 + \frac{1}{2}q(t)\{\pi(t) - \pi_M^*(t,x(t))\}^2 dt\Big\} \\ \text{subject to:} \\ dx(t) = \{x(t)r + \pi(t)(\mu - r)\}dt + \pi(t)\sigma dw(t) \\ d\pi(t) = \{\xi(t) + \pi(t)\mu\}dt + \pi(t)\sigma dw(t) \\ \xi(\cdot) \in \mathscr{A},\ x(0) = x_0,\ \pi(0) = \pi_0. \end{cases} \tag{17}$$

In this problem, the goal is to track the solution

$$\pi_M^*(t,x(t)) = \frac{\mu - r}{\sigma^2}\Big(\frac{B_M(t)}{A_M(t)} - x(t)\Big)$$

of the Merton problem (3) at minimal cost. Observe that the dynamics of this problem are identical to those of the illiquid problem (9). The value function $H(t,x,\pi)$ of (17) is the solution of the dynamic programming equations

$$\begin{cases} H_t + \inf_{\xi}\Big\{\frac{1}{2}\lambda\xi^2 + \frac{1}{2}q(t)\{\pi - \pi_M^*(t,x)\}^2 + xrH_x \\ +\pi(\mu - r)H_x + \pi\mu H_\pi + \xi H_\pi \\ +\frac{1}{2}\pi^2\sigma^2(H_{xx} + 2H_{x\pi} + H_{\pi\pi})\Big\} = 0 \\ H(T,x,\pi) = 0. \end{cases} \tag{18}$$

Linearity of $\pi_M(t,x(t))$ in $x(t)$ together with the form of the objective and state equations also imply that the value function for (17) is quadratic in the state variables (x, π).

Relationship to the Merton problem

Let

$$W(t,x,\pi) \triangleq V_M(t,x) - V(t,x,\pi) \tag{19}$$

denote the difference between the value functions for the Merton problem and the illiquid problem (9). Direct substitution of (5) and (12) gives

$$W(t,x,\pi) = \frac{1}{2}\begin{bmatrix} x \\ \pi \end{bmatrix}'\alpha(t)\begin{bmatrix} x \\ \pi \end{bmatrix} - \beta(t)'\begin{bmatrix} x \\ \pi \end{bmatrix} - \gamma(t), \tag{20}$$

where $\alpha(t),\beta(t)$ and $\gamma(t)$ satisfy

$$\begin{cases} \alpha(t) = A(t) - \begin{bmatrix} A_M(t) & 0 \\ 0 & 0 \end{bmatrix}, \\ \beta(t) = B(t) - \begin{bmatrix} B_M(t) \\ 0 \end{bmatrix}, \\ \gamma(t) = C(t) - C_M(t). \end{cases} \quad (21)$$

Since $A(t), B(t)$ and $C(t)$ satisfy (14)-(16), it follows that $\alpha(t),\beta(t)$ and $\gamma(t)$ solve

$$\begin{cases} \dot{\alpha}(t) + R'\alpha(t) + \alpha(t)R + \Sigma'\alpha(t)\Sigma - \frac{1}{\lambda}\alpha(t)SS'\alpha(t) \\ \quad + \begin{bmatrix} (\frac{\mu-r}{\sigma})^2 & \mu-r \\ \mu-r & \sigma^2 \end{bmatrix} A_M(t) = 0 \\ \alpha(T) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{cases} \quad (22)$$

$$\begin{cases} \dot{\beta}(t) + R'\beta(t) - \frac{1}{\lambda}\alpha(t)SS'\beta(t) + \begin{bmatrix} (\frac{\mu-r}{\sigma})^2 \\ \mu-r \end{bmatrix} B_M(t) = 0 \\ \beta(T) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{cases} \quad (23)$$

$$\begin{cases} \dot{\gamma} + \frac{1}{2\lambda}(S'\beta(t))^2 - \frac{1}{2}(\frac{\mu-r}{\sigma})^2 \frac{B_M(t)^2}{A_M(t)} = 0 \\ \gamma(T) = 0. \end{cases} \quad (24)$$

It is interesting to note that (22) is a Riccati equation. This suggests that $W(t,x,\pi)$ may be related to the value function of some linear quadratic control problem and it is natural to attempt constructing this problem from the equations (22)-(24) and to interpret it from a financial perspective. Along these lines, it can be shown that $W(t,x,\pi)$ is the value function of the linear-quadratic control problem

$$\inf_{\bar{\rho}(\cdot)} E\Big\{ \int_0^T \frac{\lambda}{2}\bar{\rho}(t)^2 + \frac{1}{2}A_M(t) \begin{bmatrix} x(t) - \frac{B_M(t)}{A_M(t)} \\ \pi(t) \end{bmatrix}' \times \begin{bmatrix} (\frac{\mu-r}{\sigma})^2 & \mu-r \\ \mu-r & \sigma^2 \end{bmatrix} \begin{bmatrix} x(t) - \frac{B_M(t)}{A_M(t)} \\ \pi(t) \end{bmatrix} dt \Big\}$$

subject to the same constraints in (9). Observing that the optimal solution of the Merton problem $\pi_M^*(t,x)$ is given by (6), this problem can be written as

$$\begin{cases} \inf_{\xi(\cdot)} E\Big\{ \int_0^T \frac{\lambda}{2}\xi(t)^2 + \frac{1}{2}\sigma^2 A_M(t)\{\pi(t) - \pi_M^*(t,x(t))\}^2 dt \Big\} \\ \text{subject to:} \\ dx(t) = \{x(t)r + \pi(t)(\mu-r)\}dt + \pi(t)\sigma dw(t) \\ d\pi(t) = \{\xi(t) + \pi(t)\mu\}dt + \pi(t)\sigma dw(t) \\ \xi(\cdot) \in \mathscr{A},\ x(0) = x_0,\ \pi(0) = \pi_0, \end{cases} \quad (25)$$

which is exactly (18) with $q(t) = \sigma^2 A_M(t)$. The dynamic programming equation for this problem is

$$\begin{cases} W_t + \inf_{\xi}\Big\{ \frac{1}{2}\lambda\xi^2 + \frac{1}{2}\sigma^2 A_M(t)\{\pi - \pi_M^*(t,x)\}^2 + xrW_x \\ \quad + \pi(\mu-r)W_x + \pi\mu W_\pi + \xi W_\pi \\ \quad + \frac{1}{2}\pi^2\sigma^2(W_{xx} + 2W_{x\pi} + W_{\pi\pi}) \Big\} = 0 \\ W(T,x,\pi) = 0, \end{cases} \quad (26)$$

for which $W(t,x,\pi)$ is the solution.

In summary, we have related two formulations (9) and (25) of portfolio selection problems with market impact costs and the Merton problem (3). The following result summarizes these observations and also relates the optimal trading rate policy of (25) to that of the illiquid problem (9).

Theorem 4.1: Let $V(t,x,\pi)$ and $W(t,x,\pi)$ denote the value functions for the market impact problems (9) and (25), respectively, and $V_M(t,x)$ denote the value function for the Merton problem (3). Then

$$W(t,x,\pi) = V_M(t,x) - V(t,x,\pi).$$

Moreover, both problems (9) and (25) have identical optimal trading policies:

$$\bar{\rho}^*(t,x,\pi) = \xi^*(t,x,\pi) = \frac{1}{\lambda}(\beta_1(t) - \alpha_{12}(t)x - \alpha_{22}(t)\pi). \quad (27)$$

Proof: We have already shown that $W(t,x,\pi) = V(t,x,\pi) - V_M(t,x)$ satisfies (20) and (22)-(24), and can be shown by direct substitution that (20)-(24) solves the dynamic programming equations (26) for the problem (25) and that the optimal trading policy $\xi^*(t,x,\pi)$ is given by (27), which by Proposition 3.1 is also the optimal trading rate for the illiquid problem (9). ∎

V. Asymptotic Expansions

Although Theorem 4.1 establishes an interesting relationship between the Merton problem and two seemingly different formulations of portfolio selection problems that account for market impact costs, the relationship between the optimal trading rate (27) and the solution of the fully liquid Merton problem is not particularly clear. In this section, we utilize multiple time scale perturbation methods to derive an approximation of the system (22)-(24) and the optimal trading rate (27) that is exact as the trading cost parameter λ vanishes. This analysis establishes an intuitive connection between the optimal trading rate and the value function of the illiquid problem and the frictionless Merton problem.

Intuitively, we expect that the value function $W(t,x,\pi)$ of the problem (25) will go to zero as the trading cost λ becomes small. This motivates us to derive asymptotic expansions of $W(t,x,\pi)$ in terms of λ. As a first step, we have the following result on the limiting behavior of $\alpha(t),\beta(t)$ and $\gamma(t)$. We adopt definitions of $O(\varepsilon)$ and $o(\varepsilon)$

as follows: a real function $f(t,\varepsilon)$ is $O(\varepsilon)$ over an interval $[t_1,t_2]$ if there exist positive constants k and ε^* such that

$$|f(t,\varepsilon)| \le k\varepsilon \quad \forall \varepsilon \in [0,\varepsilon^*],\ \forall t \in [t_1,t_2],$$

and $f(t,\varepsilon)$ is $o(\varepsilon)$ as ε approaches ε_0 if $\lim_{\varepsilon\to\varepsilon_0} \frac{|f(t,\varepsilon)|}{\varepsilon} = 0$.

Lemma 5.1: $\alpha(t),\beta(t)$ and $\gamma(t)$ are $O(\sqrt{\lambda})$ over the interval $[0,T]$.

Lemma 5.1 allows us to write

$$\begin{cases} \alpha(t) = \sqrt{\lambda}a(t) + o(\sqrt{\lambda}), \\ \beta(t) = \sqrt{\lambda}b(t) + o(\sqrt{\lambda}), \\ \gamma(t) = \sqrt{\lambda}c(t) + o(\sqrt{\lambda}). \end{cases}$$

where $a(t),b(t)$ and $c(t)$ are $O(1)$ for all $t \in [0,T]$. The coefficients $a(t),b(t)$ and $c(t)$ can be calculated using multiple time scale perturbation techniques [12]. The result is summarized in the following proposition.

Proposition 5.1: As λ approaches zero, the functions $\alpha(t),\beta(t),\gamma(t)$ satisfy

$$\alpha(t) = \sqrt{\lambda}\,\sigma\sqrt{A_M(t)}\tanh\Big(\sigma\sqrt{A_M(t)}\frac{T-t}{\sqrt{\lambda}}\Big) \times \begin{bmatrix} \frac{(\mu-r)^2}{\sigma^4} & \frac{(\mu-r)}{\sigma^2} \\ \frac{(\mu-r)}{\sigma^2} & 1 \end{bmatrix} + o(\sqrt{\lambda}) \tag{28}$$

$$\beta(t) = \sqrt{\lambda}\,\frac{\sigma B_M(t)}{\sqrt{A_M(t)}}\tanh\Big(\sigma\sqrt{A_M(t)}\frac{T-t}{\sqrt{\lambda}}\Big) \times \begin{bmatrix} \frac{(\mu-r)^2}{\sigma^4} & \frac{(\mu-r)}{\sigma^2} \end{bmatrix} + o(\sqrt{\lambda}) \tag{29}$$

$$\gamma(t) = \sqrt{\lambda}\,\frac{\sigma B_M^2(t)}{2A_M^{3/2}(t)}\tanh\Big(\sigma\sqrt{A_M(t)}\frac{T-t}{\sqrt{\lambda}}\Big)\frac{(\mu-r)^2}{\sigma^4} + o(\sqrt{\lambda}). \tag{30}$$

Proof: First we assume that the $\alpha(t),\beta(t)$ and $\gamma(t)$ depend on two different time scales which are defined by

$$u \triangleq T-t \quad\text{and}\quad v \triangleq \frac{T-t}{\sqrt{\lambda}}, \tag{31}$$

and write

$$\begin{cases} \alpha(u,v) = \sqrt{\lambda}a(u,v) + o(\sqrt{\lambda}), \\ \beta(u,v) = \sqrt{\lambda}b(u,v) + o(\sqrt{\lambda}), \\ \gamma(u,v) = \sqrt{\lambda}c(u,v) + o(\sqrt{\lambda}). \end{cases}$$

Under this re-parametrization, the time derivative becomes

$$\frac{d}{dt}(\cdot) = -\frac{\partial}{\partial u}(\cdot) - \frac{1}{\sqrt{\lambda}}\frac{\partial}{\partial v}(\cdot)$$

and the system (22)-(24) is transformed into a system of partial differential equations

$$\begin{aligned} \sqrt{\lambda}\frac{\partial}{\partial u}a(u,v) + \frac{\partial}{\partial v}a(u,v) &= -\,a(u,v)'SS'a(u,v) \\ &\quad + \begin{bmatrix} (\frac{\mu-r}{\sigma^2})^2 & \mu-r \\ \frac{\mu-r}{\sigma^2} & 1 \end{bmatrix} A_M(u) + o(\sqrt{\lambda}), \\ \sqrt{\lambda}\frac{\partial}{\partial u}b(u,v) + \frac{\partial}{\partial v}b(u,v) &= -\,a(u,v)SS'b(u,v) \\ &\quad + \begin{bmatrix} (\frac{\mu-r}{\sigma})^2 & \mu-r \end{bmatrix} B_M(u) + o(\sqrt{\lambda}), \\ \sqrt{\lambda}\frac{\partial}{\partial u}c(u,v) + \frac{\partial}{\partial v}c(u,v) &= \frac{1}{2}(S'b(u,v))^2 \\ &\quad - \frac{1}{2}(\frac{\mu-r}{\sigma})^2\frac{B_M(u)^2}{A_M(u)} + o(\sqrt{\lambda}). \end{aligned}$$

The system corresponding to $O(1)$ terms is

$$\begin{aligned} \frac{\partial}{\partial v}a(u,v) &= -a(u,v)'SS'a(u,v) \\ &\quad + \begin{bmatrix} (\frac{(\mu-r)}{\sigma^2})^2 & (\mu-r) \\ \frac{(\mu-r)}{\sigma^2} & 1 \end{bmatrix} A_M(u), \\ \frac{\partial}{\partial v}b(u,v) &= -a(u,v)SS'b(u,v) + \begin{bmatrix} (\frac{\mu-r}{\sigma})^2 & \mu-r \end{bmatrix} B_M(u), \\ \frac{\partial}{\partial v}c(u,v) &= \frac{1}{2}(S'b(u,v))^2 - \frac{1}{2}(\frac{\mu-r}{\sigma})^2\frac{B_M(u)^2}{A_M(u)}. \end{aligned}$$

which is explicitly solved by

$$\begin{aligned} a(u,v) &= \sigma\sqrt{A_M(u)}\tanh\Big(\sigma\sqrt{A_M(u)}v\Big)\begin{bmatrix} \frac{(\mu-r)^2}{\sigma^4} & \frac{(\mu-r)}{\sigma^2} \\ \frac{(\mu-r)}{\sigma^2} & 1 \end{bmatrix} \\ b(u,v) &= \frac{\sigma B_M(u)}{\sqrt{A_M(u)}}\tanh\Big(\sigma\sqrt{A_M(u)}v\Big)\begin{bmatrix} \frac{(\mu-r)^2}{\sigma^4} & \frac{(\mu-r)}{\sigma^2} \end{bmatrix} \\ c(u,v) &= -\frac{\sigma B_M^2(u)}{2A_M^{3/2}(u)}\tanh\Big(\sigma\sqrt{A_M(u)}v\Big)\frac{(\mu-r)^2}{\sigma^4}. \end{aligned}$$

The result in Proposition 5.1 is obtained by replacing u and v by their definitions in (31). ■

Observe that the approximation in Proposition 5.1 introduces a new *liquidity time scale* $\frac{T-t}{\sqrt{\lambda}}$. Intuitively, the amount of trading that can be done on the time horizon $T-t$ depends on the liquidity/market impact costs. For example, one week of trading could be plenty of time if liquidity costs are small (i.e. $\frac{T-t}{\sqrt{\lambda}}$ is "large") or barely enough time if trading costs are large (i.e. $\frac{T-t}{\sqrt{\lambda}}$ is "small"), which is precisely the effect that the liquidity time scale is trying to capture.

The following theorem gives asymptotic expansions of the value functions $V(t,x,\pi)$ and $W(t,x,\pi)$ as well as the optimal trading policy for the problems (9) and (25).

Theorem 5.1: With small λ, the value function of the market impact problem (9) satisfies

$$V(t,x,\pi) = V_M(t,x) - \sqrt{\lambda}\frac{\sigma\sqrt{A_M(t)}}{2}\tanh\Big(\sigma\sqrt{A_M(t)}\frac{T-t}{\sqrt{\lambda}}\Big) \times \Big\{\pi_M^*(t,x) - \pi\Big\}^2 + o(\sqrt{\lambda}). \tag{32}$$

The optimal trading policy is

$$\bar{\rho}^*(t,x,\pi) = \frac{\sigma\sqrt{A_M(t)}}{\sqrt{\lambda}}\tanh\left(\sigma\sqrt{A_M(t)}\frac{T-t}{\sqrt{\lambda}}\right)\left\{\pi_M^*(t,x)-\pi\right\} + o(1/\sqrt{\lambda}). \qquad (33)$$

Let $x(t)$ and $\pi(t)$ be wealth and risky asset holding processes of an investor who adopts the trading rate (33). The differece between the holding $\pi(t)$ and that of the Merton investor is

$$E\left\{\int_0^T \left[\pi_M^*\Big(t,x(t)\Big)-\pi(t)\right]^2 dt\right\} = \sqrt{\lambda}\frac{K}{2}(\pi_M^*(0,x(0))-\pi(0))^2 + o(\sqrt{\lambda}) \qquad (34)$$

where the coefficient

$$K = \frac{1}{2\sigma\sqrt{A_M(0)}}\tanh(\sigma\sqrt{A_M(0)}\frac{T}{\sqrt{\lambda}})$$

Proof: (32) and (33) follow immediately from the characterization (20)-(24) of $W(t,x,\pi)$, Theorem 4.1 on the relationship between the value functions of the Merton problem and the two illiquidity problems, and the expansions of $\alpha(t)$, $\beta(t)$ and $\gamma(t)$ in Proposition 5.1. The proof of (34) is in [9]. ■

Theorem 5.1 establishes a close relationship between the portfolio selection problem in an illiquid market and the Merton problem. Several properties of the value function and the optimal trading rate are worth noting:

- As the market impact costs λ decreases the value function V increases, and when λ goes to zero, V converges to the value function of the Merton problem.
- As λ decreases, the optimal holding $\pi(t)$ converges to that of the optimal Merton investor if the initial wealths of both investors are equal.
- As λ decreases, the optimal trading rate $\bar{\rho}^*$ increases and becomes arbitrarily large as λ goes to zero. Intuitively, the investor trades infinitely quickly when trading costs vanish, which allows him/her to track the optimal Merton portfolio with vanishing error.
- The optimal trading rate $\bar{\rho}^*$ increases in the difference between the current risky asset holding and the Merton optimal holding $\pi_M^* - \pi$; investor trades to close the gap between his/her current position π and the (ideal) position π_M^* he/she would adopt if there were no liquidity costs.
- The value function V decreases in the distance $|\pi_M^* - \pi|$ between his/her desired position π_M^* and his/her actual holding π. A large distance means that an investor has to trade more aggressively to reach the desired holding which incurs more costs.
- The liquidity time scale $(T-t)/\sqrt{\lambda}$ affects the optimal trading rate $\bar{\rho}^*$ and the value function only when the remaining time is comparable to $\sqrt{\lambda}$, since the hyperbolic tangent term will be significant if $(T-t)/\sqrt{\lambda}$ is small. When the liquidity time is small an investor will reduce his/her trading rate since market impact costs dominate potential improvements in utility from rebalalancing. In contrast to the solution of the classical Merton problem (Proposition 2.1) the investor needs to consider both the the liquidity as well as the original time scale when there are market impact costs.
- A larger volatility σ makes the optimal trade rate $\bar{\rho}^*$ increase. When the volatility is large, an investor trades more aggressively to close the gap between his/her position and the ideal position π_M^*.
- A larger risk-aversion parameter causes an investor to trade faster.

VI. Examples

In this section we test and compare our portfolio selection model using simulated stock markets. The markets consist of one risk-free asset and one risky asset: the risk-free return is $r = 2\%$ per year, the risky asset expected return is $\mu = 6\%$ per year, the risky asset volatility is $\sigma = 20\%$ per year, and he risky asset initial price is one dollar. An investor risk aversion factor is $\eta = 3.5\times 10^{-7}$ which is chosen according to the investor initial wealth \$1,000,000.

In the simulation, we discretize the problem by allowing an investor to adjust their trading rates once a day, or their risky asset holding if he/she is a Merton investor. We make a minor change in order to capture the real effect of liquidity costs by subtracting these costs directly from investor wealth every time he/she executing a trading order instead of subtracting from the final wealth. Hence the discrete-time version of the wealth dynamics in the simulation are

$$\Delta x(t) = \{x(t)r + \pi(t)(\mu - r)\}\Delta t + \pi(t)\sigma\Delta w(t) - \lambda\bar{\rho}(t)^2\Delta t.$$

Example 1

In this example, the market is assumed to be illiquid. We compare performances of two types of investor, a Merton investor who adopts Merton policy (6) and an illiquid investor who adopts our trading policy (13). Initially, each investor has \$1,000,000 in total wealth and holds the optimal Merton portfolio in the risky asset.

In Fig. 1, we simulate various illiquid situations using the value of λ in the range of $10^{-11} - 10^{-6}$ and plot the average of the trading rate and the average of the final wealth. As we see in the upper figure, a Merton investor (dash line), who trades with the same rate regardless the value of λ, trades faster than an illiquid investor (solid line), who reduces his/her trading rate according to an increase of λ. As a result, the Merton investor incurs significant amount of liquidity costs while the illiquid investor does not as shown in the lower figure.

The plot of the average final wealth in Fig. 1 also provides a sensible way to calibrate the value of liquidity coefficient λ. Specifically, we may choose λ according to the difference between the average final wealth of the Merton investor and the illiquid investor. For comparison, we choose the following value of λ throughout the rest of the examples:

- A mildly illiquid market $\lambda = 4.5\times 10^{-10}$ (a Merton investor loses 5% compared to a liquid market)
- A moderately illiquid market $\lambda = 1.2\times 10^{-9}$ (a Merton investor loses 15%)

- A highly illiquid market $\lambda = 2.3 \times 10^{-9}$ (a Merton investor loses 30%).

Fig. 2 shows sample paths of the total wealth, the risky asset holding, the simulated stock price, and the trading rate from a single simulation when the market is moderately illiquid. From the plot of the risky asset holding, an illiquid investor tries to decrease the gap between the Merton optimal holding and his/her current holding throughout the trading horizon, however, the randomness of the stock price and the costs of liquidity prevent him/her from doing that, and hence the crossing pattern is created as shown in the figure.

Example 2

In this example, we compare performances of investors in four different markets: a perfectly liquid market, a mildly illiquid market $\lambda = 4.5 \times 10^{-10}$, a moderately illiquid market $\lambda = 1.2 \times 10^{-9}$, and a highly illiquid market $\lambda = 2.3 \times 10^{-9}$. The investor in each market behaves optimally according to that market situation. Initially, each investor has \$1,000,000 in the risk-free asset without any holding in the risky asset.

In Fig. 3 and 4, we plot the cross-sectional average of portfolio wealth $x(t)$ and risky asset holding $\pi(t)$. An investor in a liquid market can purchase the optimal holding in risky asset according to the Merton optimal policy at the beginning of trade without liquidity costs and hold it throughout the trading horizon. On the other hand, investors in illiquid markets can only accumulate the risky asset over time to attenuate the liquidity costs. The length of time needed to build up the risky asset holding depends on the level of illiquidity in each market. As a result, the average final wealth is highest in the liquid market and decreases as λ increases.

VII. Conclusion

We study a portfolio selection problem in an illiquid market by extending the well-known Merton portfolio selection problem. The key part of our model is that we capture illiquidity effects using the penalty term which is quadratic in the trading rate. In the case of quadratic utility function, our

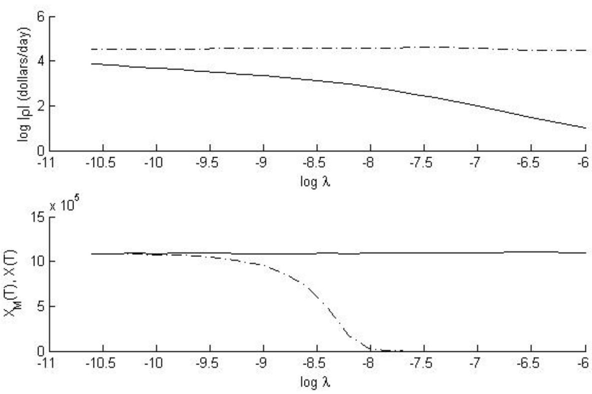

Fig. 1. The upper figure compares the average of the absolute value of trading rate obtained by our model (solid line) and by Merton model (dash line). Similarly, the lower one compares the final wealth.

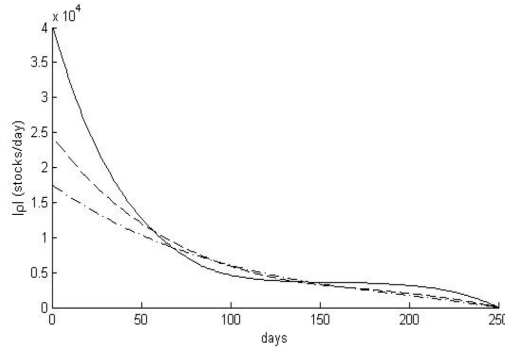

Fig. 4. The cross-sectional average of the trading rate using the same setting as in Fig. 3.

problem become a linear-quadratic control problem which is easy to solve. Then we derive the approximation of the optimal trading policy in term of the optimal solution of perfectly liquid market case. The result clearly shows the relationship between our solution and the solution to the Merton problem, and how the state variables affect the optimal trading policy. Our analysis also gives rise to the notion of a *liquidity time scale*.

Though this paper focuses on quadratic utility and assumes a single risky asset, our results can be extended to the multiple asset case quite easily (and will be reported elsewhere). We also believe that non-quadratic utilities can be handled by applying similar ideas to the associated dynamic programming equations. Other interesting extensions are also possible, including models with stochastic market impact costs that jump whenever asset prices jump.

References

[1] R. C. Merton, “Optimum consumption and portfolio rules in a continuous-time model,” *Journal of Economic Theory*, vol. 3, pp. 373–413, 1971.

[2] R. Almgren and N. Chriss, “Optimal execution of portfolio transaction,” *Journal of Risk*, vol. 3, pp. 5–39, 2000.

[3] D. Bertsimas and A. Lo, “Optimal control of execution costs,” *Journal of Financial Markets*, vol. 1, pp. 1–50, 1998.

[4] N. Garleanu and L. H. Pedersen, “Liquidity and risk management,” *The American Economic Review Papers and Proceedings*, vol. 97, pp. 193–197, 2007.

[5] H. He and H. Mamaysky, “Dynamic trading with price impact,” *Journal of Economic Dynamics and Control*, 2005.

[6] F. A. Longstaff, “Optimal portfolio and the valuation of illiquid securities,” *Review of Financial Studies*, vol. 14, no. 2, pp. 407–431, 2001.

[7] L. C. G. Rogers and S. Singh, “The cost of illiquidity and its effects on hedging,” *Mathematical Finance*, vol. 20, no. 4, pp. 597–615, 2010.

[8] A. Schied, T. Schoneborn, and M. Tehranchi, “Optimal basket liquidation for cara investors is deterministic,” *Preprint*, 2010.

[9] A. Lim and P. Wimonkittiwat, “Dynamic portfolio selection with market impact costs,” *Working paper*, 2011.

[10] N. Garleanu and L. H. Pedersen, “Dynamic trading with predictable returns and transaction costs,” *Preprint*, 2009.

[11] R. Grinold, “A dynamic model of portfolio management,” *Journal of Investment Management*, vol. 4, pp. 5–22, 2006.

[12] P. Kokotovic, H. Khalil, and J. O’Reilly, *Singular Perturbation Methods in Control: Analysis and Design*. Academic Press, 1986.

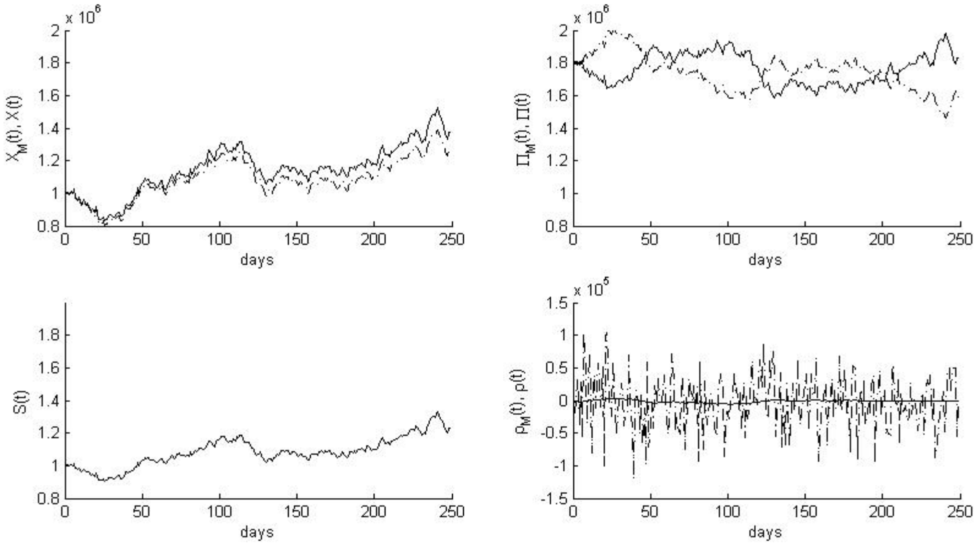

Fig. 2. This figure shows sample paths of the total wealth (top left), the risky asset holding (top right), the simulated stock price (bottom left), and the trading rate (bottom right) of a Merton investor (dash line) and an illiquid investor (solid line) in a moderately illiquid market ($\lambda = 1.2 \times 10^{-9}$).

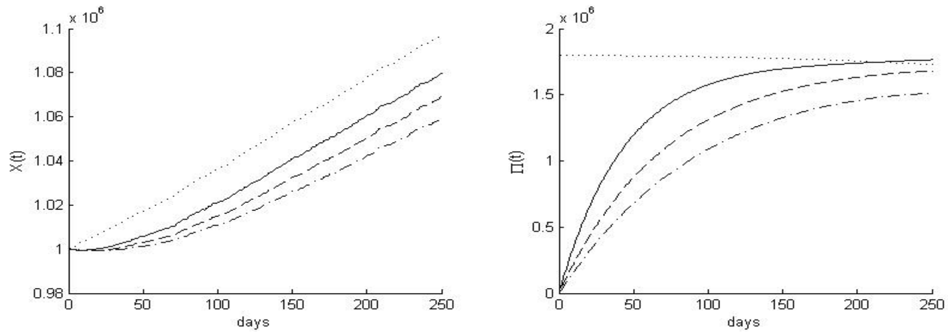

Fig. 3. This figure shows the cross-sectional average of wealth $x(t)$ (Left) and the risky asset holding $\pi(t)$ (Right) from four different markets: 1.) Perfectly liquid market (dotted line) 2.) Mildly illiquid market $\lambda = 4.5 \times 10^{-10}$ (solid line) 3.) Moderately illiquid market $\lambda = 1.2 \times 10^{-9}$ (dash line) and 4.) Highly illiquid market $\lambda = 2.3 \times 10^{-9}$(dash-dotted line).

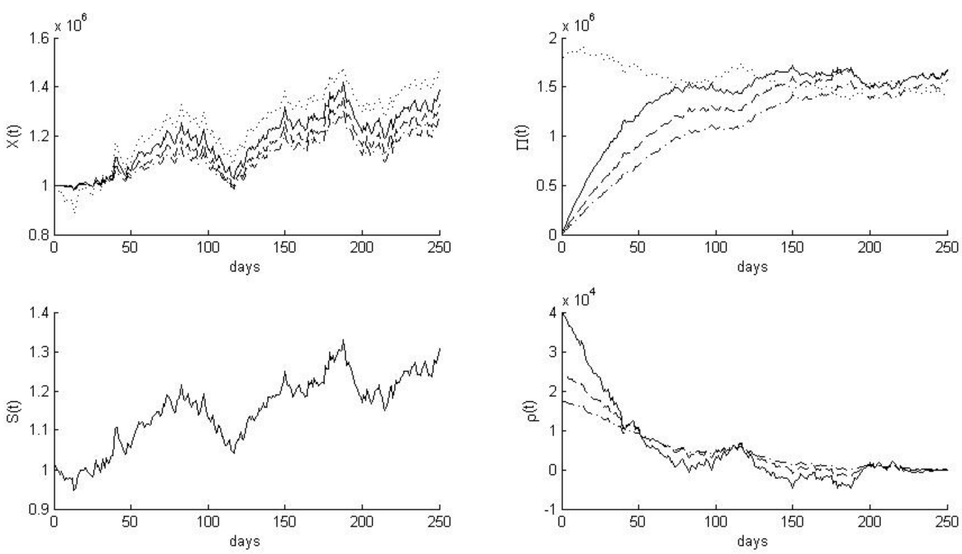

Fig. 5. This figure shows the sample path of the total wealth (top left), the risky asset holding (top right), the simulated stock price (bottom left), and the trading rate (bottom right) from four different markets: 1.) Perfectly liquid market (dotted line) 2.) Mildly illiquid market $\lambda = 4.5 \times 10^{-10}$ (solid line) 3.) Moderately illiquid market $\lambda = 1.2 \times 10^{-9}$ (dash line) and 4.) Highly illiquid market $\lambda = 2.3 \times 10^{-9}$(dash-dotted line).