Optimal Portfolio Selection with Transaction Costs for a CARA Investor with Finite Horizon

Valeri. I. Zakamouline

Institute of Finance and Management Science
Norwegian School of Economics and Business Administration
Helleveien 30, 5045 Bergen, Norway
Valeri.Zakamouline@nhh.no

This revision: December 9, 2002

Abstract

In this paper we study the optimal portfolio selection problem for a CARA investor who faces fixed and/or proportional transaction costs and maximizes expected utility of end-of-period wealth. We use a continuous time model and apply the method of the Markov chain approximation to numerically solve for the optimal trading policy. The numerical solution indicates that, most of the time, the portfolio space is divided into three disjoint regions (Buy, Sell, and No-Transaction), and the optimal policy is described by four boundaries. If a portfolio lies in the Buy region, the optimal strategy is to buy the risky asset until the portfolio reaches the Buy target boundary. Similarly, if a portfolio lies in the Sell region, the optimal strategy is to sell the risky asset until the portfolio reaches the Sell target boundary. However, we discovered that for a CARA investor with a finite horizon there is generally a time interval, close to the terminal date, when the NT region consists of two disjoint sub-regions which, in their turn, divide either the Buy region or the Sell region into two parts. Nevertheless, as in the former case, the target boundaries are unique. The effects on the optimal policy of varying volatility, drift, ARA, and the level of transaction costs are also examined.

1 Introduction

In this paper, we study the optimal portfolio selection problem for a constant absolute risk averse (CARA) investor. The investor faces fixed and/or proportional transaction costs and maximizes expected utility of end-of-period wealth.

This asset allocation problem is a variant of the classical consumption-investment problem in modern finance. In the absence of transaction costs, the closed-form solution was obtained by Merton (1971). The two-asset problem is of particular interest. When the stock price follows a geometric Brownian motion, the solution indicates that it is optimal for the investor to keep a certain amount of his wealth in the risky asset. As time passes, the portfolio is assumed to be adjusted continuously. Moreover, this amount is independent of the investor's total wealth.

The introduction of transaction costs adds considerable complexity to the optimal portfolio selection problem. The problem is simplified if one assumes that the transaction costs are proportional to the amount of the risky asset traded, and there are no transaction costs on trades in the riskless asset. In this case, the problem amounts to a *stochastic singular control* problem. The problem is often further simplified if the investor's horizon is infinite, which gives a stationary portfolio policy. The solution of the optimal portfolio selection problem where each transaction has a fixed cost component is more complicated and is based on the theory of *stochastic impulse controls*.

Even though a utility function that exhibits constant absolute risk aversion is well known, it received little attention in the literature concerning optimal portfolio policy with transaction costs[1]. On the contrary, there are a large number of papers analyzing the optimal portfolio policy for a constant relative risk averse (CRRA) investor. Some examples of the papers studying the consumption-investment problem for a CRRA investor facing proportional transaction costs are Davis and Norman (1990), Dumas and Luciano (1991), Shreve and Soner (1994), and Akian, Menaldi, and Sulem (1996). The same problem for a CRRA investor facing transaction costs which have a fixed fee component was analyzed by Eastham and Hast-

[1]One usually notes that CARA utility function is behaviorally less plausible than CRRA utility function

ings (1988), Schroder (1995), Øksendal and Sulem (1999) and Chancelier, Øksendal, and Sulem (2000).

Closely related models of optimal consumption-investment with a CARA investor and transaction costs which have a fixed fee component have been previously analyzed by Korn (1998) and Liu (2001). Korn (1998) also considers a CARA investor who maximizes expected utility of end-of-period paper wealth (that is, there are no transaction costs on the terminal date). He uses the asymptotic analysis developed by Atkinson and Wilmott (1995) and Whalley and Wilmott (1997) to solve the problem approximately. Liu (2001) studies the intertemporal consumption and investment policy of a CARA investor with an infinite horizon. He solves the problem by reducing it to the solution of a system of ordinary differential equations. He also provides the analysis of the investor's optimal policy for a large set of realistic parameters. In both of these papers, authors need to make some apriory assumptions about the shape of the investor's optimal policy, and then they find a solution to the problem under these particular assumptions.

In this paper, we numerically solve the asset allocation problem for the investor with a finite horizon without making any apriory assumptions of how the optimal strategy looks. That is, our analysis relies solely on numerical calculations. We apply the method of the *Markov chain approximation* (see, for example, Kushner and Dupuis (1992)). Using this method, the solution to the problem is obtained by turning the stochastic differential equations into Markov chains in order to apply the discrete-time dynamic programming algorithm.

It is known that in the presence of proportional and fixed transaction costs the general description of the optimal portfolio policy is as follows: The portfolio space can be divided into three disjoint regions, which can be specified as the Buy, Sell, and No-Transaction (NT) regions, and the optimal policy is described by four boundaries. The Buy and the NT regions are divided by the lower no-transaction boundary, and the Sell and the NT regions are divided by the upper no-transaction boundary. If a portfolio lies in the Buy region, the optimal strategy is to buy the risky asset until the portfolio reaches the Buy target boundary. Similarly, if a portfolio lies in the Sell region, the optimal strategy is to sell the risky asset until the portfolio reaches the Sell target boundary.

However, we discovered that for an investor with CARA utility and a finite horizon there is generally a time interval, close to the terminal date, when the NT region consists of two disjoint sub-regions which, in their turn, divide either the Buy region (when the drift of the risky asset is greater than the risk-free interest rate, $\mu > r$) or the Sell region (when $\mu < r$) into two parts. Nevertheless, as in the former case, the target boundaries are unique.

The rest of the paper is organized as follows. Section 2 presents the continuous-time model. Section 3 is concerned with the construction of a discrete time approximation of the continuous time price processes used in Section 2, and the solution method. For the sake of comparison and completeness, Sections 4, 5, and 6 present the models and numerical solutions for the problems in the absence of transaction costs, in the presence of proportional transaction costs only, and in the presence of fixed transaction costs only, respectively. The numerical results of the model with both fixed and proportional transaction costs are presented in Section 7. Throughout the paper, we discuss some properties of the value functions and provide the numerical analysis of the optimal policy. The effects on the optimal policy of varying volatility, drift, ARA, and the level of transaction costs are also examined. Our findings here agree with those of Liu (2001). Section 8 concludes the paper and discusses some possible extensions.

2 The Continuous Time Model

Originally, we consider a continuous-time economy with one risky and one risk-free asset. Let $(\Omega, \mathcal{F}, P)$ be a probability space with a given filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$. The risk-free asset, which we will refer to as the bank account, pays a constant interest rate of $r \geq 0$, and, consequently, the evolution of the amount invested in the bank, x_t, is given by the ordinary differential equation

$$dx_t = rx_t dt \tag{1}$$

We will refer to the risky asset as the stock, and assume that the price of the stock, S_t, evolves according to a geometric Brownian motion defined by

$$dS_t = \mu S_t dt + \sigma S_t dB_t \tag{2}$$

where μ and σ are constants, and B_t is a one-dimensional $\mathcal{F}_t$-Brownian motion.

The investor holds x_t in the bank account and the amount y_t in the stock at time t. We assume that a purchase or sale of stocks of the amount ξ incurs a transaction costs consisting of a sum of a fixed cost $k \geq 0$ (independent of the size of transaction) plus a cost $\lambda|\xi|$ proportional to the transaction ($\lambda \geq 0$). These costs are drawn from the bank account.

We suppose that at any time the investor can decide to transfer money from the bank account to the stock and conversely. The control of the investor is a pure *impulse control* $v = (\tau_1, \tau_2, \ldots; \xi_1, \xi_2, \ldots)$. Here $0 \leq \tau_1 < \tau_2 < \ldots$ are $\mathcal{F}_t$-stopping times giving the times when the investor decides to change his portfolio, and ξ_j are $\mathcal{F}_{\tau_j}$-measurable random variables giving the sizes of the transactions at these times. If such a control is applied to the system (x_t, y_t), it gets the form

$$\begin{aligned} dx_t &= rx_t dt && \tau_i \leq t < \tau_{i+1} \\ dy_t &= \mu y_t dt + \sigma y_t dB_t && \tau_i \leq t < \tau_{i+1} \\ x_{\tau_{i+1}} &= x_{\tau_{i+1}^-} - k - \xi_{i+1} - \lambda|\xi_{i+1}| \\ y_{\tau_{i+1}} &= y_{\tau_{i+1}^-} + \xi_{i+1} \end{aligned} \tag{3}$$

If the investor has the amount x in the bank account, and the amount y in the stock, his *net wealth* is defined as the holdings on the bank account after selling of all shares of the stock (if the proceeds are positive after transaction costs) or closing of the short position in the stock and is given by

$$X_t(x,y) = \begin{cases} \max\{x_t + y_t(1-\lambda) - k, x_t\} & \text{if } y_t \geq 0, \\ x_t + y_t(1+\lambda) - k & \text{if } y_t < 0. \end{cases} \tag{4}$$

We consider an investor with a finite horizon $[0,T]$ who has utility only of terminal wealth. It is assumed that the investor has a constant absolute risk aversion. In this case his utility function is of the form

$$U(\gamma, W) = -\exp(-\gamma W), \tag{5}$$

where γ is a measure of the investor's absolute risk aversion (ARA), which is independent of the investor's wealth.

The investor's problem is to choose an admissible trading strategy to maximize $E_t[U(\gamma, X_T)]$, i.e. the expected utility of his net terminal wealth, subject to (3). We define the value function at time t as

$$V(t,x,y) = \sup_{v \in \mathcal{A}(x,y)} E_t^{x,y}[U(\gamma, X_T)], \tag{6}$$

where $\mathcal{A}(x,y)$ denotes the set of admissible controls available to the investor who starts at time t with an amount of x in bank and y holdings in the stock. We define the *intervention operator* (or the maximum utility operator) $\mathcal{M}$ by

$$\mathcal{M}V(t,x,y) = \sup_{(x',y') \in \mathcal{S}} V(t,x',y') \tag{7}$$

where x' and y' are the new values of x and y. $\mathcal{M}V(t,x,y)$ represents the value of the strategy that consists in choosing the best transaction. We define the *continuation region* D by

$$D = \Big\{(x,y); V(t,x,y) > \mathcal{M}V(t,x,y)\Big\} \tag{8}$$

The continuation region is the region where it is not optimal to rebalance the investor's portfolio.

Now we intend to characterize the value function and the associated optimal strategy, assuming there exists an optimal strategy for each initial point (t,x,y). Then, if the optimal strategy is to not transact, the utility associated with this strategy is $V(t,x,y)$. On the other hand, selecting the best transaction and then following the optimal strategy gives the utility $\mathcal{M}V(t,x,y)$. Since the first strategy is optimal, its utility is greater or equal to the utility associated with the second strategy. Hence, $V(t,x,y) \geq \mathcal{M}V(t,x,y)$ with equality when it is optimal to make a transaction. Moreover, in the continuation region the application of the dynamic programming principle gives us $\mathcal{L}V(t,x,y) = 0$, where the operator $\mathcal{L}$ is defined by

$$\mathcal{L}V(t,x,y) = \frac{\partial V}{\partial t} + rx\frac{\partial V}{\partial x} + \mu y\frac{\partial V}{\partial y} + \frac{1}{2}\sigma^2 y^2 \frac{\partial^2 V}{\partial y^2}. \tag{9}$$

Theorem 1. *The value function is the unique constrained viscosity solution of the quasi-variational Hamilton-Jacobi-Bellman inequalities (QVHJBI, or just QVI):*

$$\max\Big\{\mathcal{L}V, \quad \mathcal{M}V - V\Big\} = 0 \tag{10}$$

with the boundary condition

$$V(T,x,y) = U(\gamma, X_T).$$

The proof can be made by following along the lines of the proof in Øksendal and Sulem (1999) with the correction for no consumption and finite horizon.

It is easy to see from (3) that the amount x in the bank account at time T is given by

$$x_T = \frac{x_t}{\delta(T,t)} - \sum_{i=0}^{n} \frac{(k + \xi_i + \lambda|\xi_i|)}{\delta(T,\tau_i)}, \tag{11}$$

where $\delta(T,t)$ is the discount factor defined by

$$\delta(T,t) = \exp(-r(T-t)), \tag{12}$$

and $t \leq \tau_1 < \tau_2 < \ldots < \tau_n < T$. Therefore, taking into consideration the investor's utility function defined by (5), we can write

$$V(t,x,y) = \exp(-\gamma\frac{x}{\delta(T,t)})Q(t,y), \tag{13}$$

where $Q(t,y)$ is defined by $Q(t,y) = V(t,0,y)$. It means that the dynamics of y through time is independent of x. This representation suggests transformation of (10) into the following QVI for the value function $Q(t,y)$:

$$\max\left\{\mathcal{D}Q(t,y), \sup_{y'\in\mathcal{A}(y)} \exp\left(\gamma\frac{k-(y-y')+\lambda|y-y'|}{\delta(T,t)}\right)Q(t,y') - Q(t,y)\right\} = 0, \tag{14}$$

where y' is the new value of y, $\mathcal{A}(y)$ denotes the set of admissible controls available to the investor who starts at time t with y holdings in the stock, and the operator $\mathcal{D}$ is defined by

$$\mathcal{L}Q(t,y) = \frac{\partial Q}{\partial t} + \mu y\frac{\partial Q}{\partial y} + \frac{1}{2}\sigma^2 y^2\frac{\partial^2 Q}{\partial y^2}. \tag{15}$$

This is an important simplification of the problem that reduces the dimensionality of the problem. Note that the function $Q(t,y)$ is evaluated in the two-dimensional space $[0,T]\times\mathbb{R}$.

3 A Markov Chain Approximation of the Continuous Time Problem

It is tempting to try to solve the partial differential equation (10) by using the classical finite-difference method, but the PDE has only a formal meaning and is to be interpreted in a symbolic sense. Indeed, we do not know whether the partial derivatives of the value function are well defined, i.e., the value function has a twice continuously differentiable solution. The method of solution of such problems was suggested by Kushner (including Kushner (1977), Kushner (1990), and Kushner and Dupuis (1992)). The basic idea involves a consistent approximation of the problem by a Markov chain, and then solving an appropriate optimization problem for the Markov chain model. Unlike the classical finite-difference method, the smoothness of the solution to the HJB or QVI equations is not needed. The methods of proof of convergence are relatively simple and require the use of only some basic ideas in the theory of weak convergence of a sequence of probability measures of random processes. Some examples of proofs of convergence of the value function of the discrete time models to their continuous time counterparts are: Fitzpatrick and Flemming (1991), Davis, Panas, and Zariphopoulou (1993), and Collings and Haussmann (1998). An alternative approach for proving the convergence of numerical schemes is based on the notion of viscosity solution. For an example of such a proof see Davis et al. (1993).

In practical applications there are two basic approaches to the realization of the Markov chain approximation method. Using the first approach, one constructs a discrete time approximation to the continuous time price processes used in the continuous time model. Then the discrete time program is solved by using the discrete time dynamic programming algorithm. The examples of use of this approach are Hodges and Neuberger (1989) and Davis et al. (1993). Using the second approach, one discretizes a HJB/QVI equation by applying the finite-difference approximation scheme which serves here only as a guide to the construction of a Markov chain. The coefficients of the resulting discrete equation is then used as the transition probabilities. This approach is often denoted as "finite difference" method, but the use of finite differences is just a "device" to get a Markov chain, in itself the approach is not a finite-difference method. The examples of use of this approach are Akian et al. (1996), Øksendal and Sulem (1999) and Chancelier et al. (2000).

The main objective of this section is to present a numerical procedure for computing the optimal trading policy. We will follow the first approach and are concerned with the construction of a discrete time approximation to the continuous time price processes used in the continuous time model presented in the previous section. The reason is to be able to solve the problem numerically, i.e., our discrete time utility maximization problem is a Markov chain approximation to the associated continuous time problem. The discrete time program is then solved by using the backward recursion algorithm.

Consider the partition $0 = t_0 < t_1 < \ldots < t_n = T$ of the time interval $[0, T]$ and assume that $t_i = i\Delta t$ for $i = 0, 1, \ldots, n$ where $\Delta t = \frac{T}{n}$. Let ε be a stochastic variable:

$$\varepsilon = \begin{cases} u & \text{with probability } p, \\ d & \text{with probability } 1 - p. \end{cases}$$

We define the discrete time stochastic process of the stock as:

$$S_{t_{i+1}} = S_{t_i}\varepsilon \tag{16}$$

and the discrete time process of the risk-free asset as:

$$x_{t_{i+1}} = x_{t_i}\rho \tag{17}$$

If we choose $u = e^{\sigma\sqrt{\Delta t}}$, $d = e^{-\sigma\sqrt{\Delta t}}$, $\rho = e^{r\Delta t}$, and $p = \frac{1}{2}\left[1 + \frac{\mu}{\sigma}\sqrt{\Delta t}\right]$, we obtain the binomial model proposed by Cox, Ross, and Rubinstein (1979). An alternative choice is $u = e^{\mu\Delta t + \sigma\sqrt{\Delta t}}$, $d = e^{\mu\Delta t - \sigma\sqrt{\Delta t}}$, $\rho = e^{r\Delta t}$, and $p = \frac{1}{2}$, which was proposed by He (1990). As n goes to infinity, the discrete time processes (16) and (17) converge in distribution to their continuous counterparts (2) and (1).

The following discretization scheme is proposed for the QVI (10):

$$\mathbb{V}^{\Delta t} = \mathcal{O}(\Delta t)\mathbb{V}^{\Delta t}, \tag{18}$$

where $\mathcal{O}(\Delta t)$ is an operator given by

$$\begin{aligned}\mathcal{O}(\Delta t)\mathbb{V}^{\Delta t} = \max\Big\{ &\max_m \mathbb{V}^{\Delta t}(t_i, x-k-(1+\lambda)m\Delta y, y+m\Delta y), \\ &\max_m \mathbb{V}^{\Delta t}(t_i, x-k+(1-\lambda)m\Delta y, y-m\Delta y), \\ &E\{\mathbb{V}^{\Delta t}(t_{i+1}, x\rho, y\varepsilon)\}\Big\},\end{aligned} \tag{19}$$

where m runs through the positive integer numbers, and

$$\begin{aligned}&\mathbb{V}^{\Delta t}(t_i, x-k-(1+\lambda)m\Delta y, y+m\Delta y) \\ &\qquad = E\left\{\mathbb{V}^{\Delta t}(t_{i+1}, (x-k-(1+\lambda)m\Delta y)\rho, (y+m\Delta y)\varepsilon)\right\}\end{aligned} \tag{20}$$

$$\begin{aligned}&\mathbb{V}^{\Delta t}(t_i, x-k+(1-\lambda)m\Delta y, y-m\Delta y) \\ &\qquad = E\left\{\mathbb{V}^{\Delta t}(t_{i+1}, (x-k+(1-\lambda)m\Delta y)\rho, (y-m\Delta y)\varepsilon)\right\},\end{aligned} \tag{21}$$

as at time t_i we do not know yet the value function. Instead, we use the known values at the next time instant, t_{i+1}. Here we have discretized the y-space in a lattice with grid size Δy, and the x-space in a lattice with grid size Δx[2]. This scheme is based on the principle that the investor's policy is the choice of the optimal transaction, that is, to buy, sell, or do nothing for a particular state given the value function for all states in the next time instant.

Theorem 2. *The solution $V^{\Delta t}$ of (18) converges locally uniformly to the unique continuous constrained viscosity solution of (10) as $\Delta t \to 0$*

The proof is based on the notion of viscosity solutions and can be made, we believe, in the same manner as the proof of Theorem 4 in Davis et al. (1993). For the present we leave it as a problem for future research.

Also in the discrete time framework the dynamics of y through time is independent of x. Therefore (13) can be written as follows:

$$\mathbb{V}^{\Delta t}(t, x, y) = \exp(-\gamma \frac{x}{\delta(T,t)})\mathbb{Q}^{\Delta t}(t, y) \tag{22}$$

The discretization scheme for the function $\mathbb{Q}^{\Delta t}(t, y)$ is derived from (18) and

[2] It is supposed that $\lim_{\Delta t \to 0} \Delta y \to 0$, and $\lim_{\Delta t \to 0} \Delta x \to 0$, that is, $\Delta y = c_y \Delta t$, and $\Delta x = c_x \Delta t$ for some constants c_y and c_x

(22) to be

$$\mathbb{Q}^{\Delta t}(t_i, y) = \max\Big\{ \max_m \exp\left(\gamma \frac{k + (1+\lambda)m\Delta y}{\delta(T, t_i)}\right) \mathbb{Q}^{\Delta t}(t_i, y + m\Delta y), \\ \max_m \exp\left(\gamma \frac{k - (1-\lambda)m\Delta y}{\delta(T, t_i)}\right) \mathbb{Q}^{\Delta t}(t_i, y - m\Delta y), \\ E\{\mathbb{Q}^{\Delta t}(t_{i+1}, y\varepsilon)\}\Big\}. \tag{23}$$

4 Optimal Policy Without Transaction Costs

We consider the case without any transaction costs ($\lambda = 0$, $k = 0$) for the sake of comparison. The investor's problem can be rewritten as

$$V(t, x, y) = \sup_{(x,y)} E_t^{x,y}[U(\gamma, x_T + y_T)], \tag{24}$$

subject to the self-financing condition

$$d(x_t + y_t) = (rx_t + \mu y_t)dt + \sigma y_t dB_t. \tag{25}$$

Merton (including Merton (1969), Merton (1971), and Merton (1973)) re-parametrized the problem by introducing new variables $w_t = x_t + y_t$ (the total wealth) and $\pi_t = \frac{y_t}{w_t}$ (the fraction of the total wealth held in the stock). Since transactions are costless and instantaneous we can regard π_t as a sole decision variable. The reformulated stochastic control problem becomes

$$V(w, t) = \sup_{\pi} E_t^{w}[U(\gamma, w_T)], \tag{26}$$

subject to

$$dw_t = [(\mu - r)\pi_t + r]w_t dt + \sigma w_t \pi_t dB_t. \tag{27}$$

The solution of (26) is known to satisfy the HJB equation

$$\frac{\partial V}{\partial t} + \sup_{\pi}\left\{(rw + (\mu - r)\pi)\frac{\partial V}{\partial w} + \frac{1}{2}\sigma^2\pi^2\frac{\partial^2 V}{\partial w^2}\right\} = 0, \tag{28}$$

subject to the boundary condition

$$V(w, T) = U(\gamma, w). \tag{29}$$

The maximum in (28) is achieved at

$$\pi^*(w,t) = -\frac{\mu - r}{\sigma^2}\frac{\frac{\partial V}{\partial w}}{\frac{\partial^2 V}{\partial w^2}}. \tag{30}$$

We focus for now on the special case of the CARA utility function (5). Then a straightforward, but tedious, calculation shows that

$$y_t^* = \frac{e^{-r(T-t)}}{\gamma}\frac{\mu - r}{\sigma^2}. \tag{31}$$

Observe that y_t^* is not stochastic and, in particular, is independent of the investor's wealth. The optimal policy requires continuous trading in the stock. Note that as the investor's risk aversion increases and as the time to the terminal date increases, the amount of money invested in the risky asset decreases.

5 Proportional Transaction Costs

Here we consider the case with proportional transaction costs only ($\lambda > 0$, $k = 0$). In this case the problem can be formulated as a *singular stochastic control* problem (see Davis and Norman (1990) and Shreve and Soner (1994)). In contrast to the no transaction cost case, at any time t the portfolio space is divided into three disjoint regions, which can be specified as the Buy region, the Sell region, and the No-Transaction (NT) region, and two boundaries describe the optimal policy. If a portfolio lies either in the Buy region or in the Sell region, the optimal strategy is to buy/sell the risky asset until the portfolio reaches the closest boundary of the NT region.

The HJB-equation for this singular stochastic control problem is given by

$$\max\left\{\mathcal{L}V(t,x,y), -(1+\lambda)\frac{\partial V}{\partial x} + \frac{\partial V}{\partial y}, (1-\lambda)\frac{\partial V}{\partial x} - \frac{\partial V}{\partial y}\right\} = 0 \tag{32}$$

where $\mathcal{L}V(t,x,y)$ is defined by (9).

The heuristic arguments for this are as follows: If for some initial point (t,x,y) the optimal strategy is to not transact, the utility associated with this strategy is $V(t,x,y)$. The necessary conditions for the optimality of the no transaction strategy is $V_y \le (1+\lambda)V_x$ and $V_y \ge -(1-\lambda)V_x$. That is, it is not possible for the investor to increase his indirect utility function by either buying or selling some amount of the stock at the expense of lowering or increasing, respectively, the holdings in the bank account. These inequalities hold with equalities when it is optimal to rebalance the portfolio. The set of (x, y) points for which $V_y = (1+\lambda)V_x$ defines the Buy region. Similarly, the equation $V_y = (1-\lambda)V_x$ defines the Sell region. Moreover, in the NT region, the application of the dynamic programming principle gives $\mathcal{L}V(t, x, y) = 0$

Let's for the moment write the investor's value function as $V(t, \gamma, x, y, \lambda)$. By this we want to emphasize that the value function depends on the investor's coefficient of absolute risk aversion and the level of proportional transaction costs.

Proposition 1. *For an investor with the exponential utility function with absolute risk aversion γ, an initial endowment (x, y), and proportional costs λ we have*

$$V(t, \gamma, x, y, \lambda) = V(t, \theta\gamma, \frac{x}{\theta}, \frac{y}{\theta}, \lambda), \tag{33}$$

where θ is any positive real number.

Proof. This relationship can be easily established from the form of the exponential utility function. In particular, the portfolio process $\{\frac{x_s}{\theta}, \frac{y_s}{\theta}; s > t\}$ is admissible given the initial portfolio $(\frac{x_t}{\theta}, \frac{y_t}{\theta})$ if and only if $\{x_s, y_s; s > t\}$ is admissible given the initial portfolio (x_t, y_t). Furthermore, $U(\gamma, X_T) = U(\theta\gamma, \frac{X_T}{\theta})$.

Proposition (1) establishes a very convenient property which allows to calculate the optimal policy for a single absolute risk aversion coefficient γ and then to obtain the optimal policy for another γ' by simple re-scaling.

Again, it can be shown that the dynamics of y through time is independent of x and we can introduce a new value function $Q(t, y) = V(t, 0, y)$ as in (13). Then, (32) could be transformed into the following HJB for the value function $Q(t, y)$:

$$\max\left\{\mathcal{D}Q(t, y), \frac{\gamma(1+\lambda)}{\delta(T, t)}Q + \frac{\partial Q}{\partial y}, -\left(\frac{\gamma(1-\lambda)}{\delta(T, t)}Q - \frac{\partial Q}{\partial y}\right)\right\} = 0, \tag{34}$$

where the operator $\mathcal{D}$ is defined by (15).

For fixed values of μ, σ, r, γ, and λ the NT boundaries are functions of the investor's horizon only and do not depend on the investor's holdings in the bank account, so that a possible description of the optimal policy may be given by

$$
\begin{aligned}
y &= y_l(\tau) \\
y &= y_u(\tau),
\end{aligned}
\tag{35}
$$

where $\tau = T - t$ represents the time remaining until the terminal date. The equations describe the lower and the upper no-transaction boundaries respectively. If the function $Q(t, y)$ is known in the NT region, then

$$
Q(t, y) = \begin{cases} \exp\left(-\gamma \frac{(1-\lambda)(y-y_u)}{\delta(T,t)}\right) Q(t, y_u) & \forall y(t) \geq y_u(t), \\ \exp\left(-\gamma \frac{(1+\lambda)(y-y_l)}{\delta(T,t)}\right) Q(t, y_l) & \forall y(t) \leq y_l(t). \end{cases} \tag{36}
$$

This follows from the optimal transaction policy described above. That is, if a portfolio lies in the Buy or Sell region, then the investor performs the minimum transaction required to reach the closest NT boundary.

The two inequalities in (32), which describe the Buy and the Sell region, cannot be directly implemented in a numerical method. The catch is in that these two inequalities were derived under assumption that the value function *is known*. In this case, the two inequalities tell us that at any time t and in any point (x, y) in the domain of the value function, the differential of the value function along the Buy or the Sell line starting from (x, y) must be less or equal to zero. The practical implication of these two inequalities is as follows: it is impossible to increase the value function by either buying or selling some amount of the stock at the expense of lowering or increasing, respectively, the holdings in the bank account. We next show how to implement this practical implication in a numerical method.

Any numerical method, either a finite-difference or a Markov chain approximation, implements a discrete dynamic programming algorithm where the value function at time t_i is found by using the known values at the next time instant t_{i+1}. The obvious numerical procedure to solve the optimal portfolio selection problem with only proportional transaction costs is analogous to that used to solve the optimal portfolio selection problem with both fixed and proportional transaction costs. First, we solve the PDE equation $\mathcal{L}V(t, x, y) = 0$ for the no-transaction problem. Then we need to compare the value function at each point (x, y) with the maximum attainable values from either buying or selling some amount of the stock. Mathematically this procedure is described by the maximum utility operator $\mathcal{M}$ defined by (7).

The subsequent theorem formalizes this intuition.

Theorem 3. *The value function V of the optimal portfolio selection problem in the presence of only proportional transaction costs is the unique constrained viscosity solution of the QVI (10).*

Moreover, we can prove that the two different formulations of the same problem, (32) and (10), yield the same result. It suffices to prove the following theorem.

Theorem 4. *For the optimal portfolio selection problem with proportional transaction costs only,*

$$\begin{aligned} -(1+\lambda)V_x + V_y &\leq 0 \\ (1-\lambda)V_x - V_y &\leq 0 \end{aligned} \tag{37}$$

if and only if

$$\mathcal{M}V - V \leq 0 \tag{38}$$

Proof. The first part. Assume (37) holds. Chose any point (x_0, y_0). Suppose that the maximum along the Buy line starting in (x_0, y_0) is attained at the point $(x_0 - (1+\lambda)\alpha, y_0 + \alpha)$, and that the maximum along the Sell line starting in (x_0, y_0) is attained at the point $(x_0 + (1-\lambda)\beta, y_0 - \beta)$. Then for the maximum along the Buy line we have that

$$\begin{aligned} &V(t, x_0 - (1+\lambda)\alpha, y_0 + \alpha) = V(t, x_0, y_0) \\ &+ \int_0^\alpha \big[-(1+\lambda)V_x(t, x_0 - (1+\lambda)s, y_0 + s) + V_y(t, x_0 - (1+\lambda)s, y_0 + s)\big] ds \\ &\qquad \leq V(t, x_0, y_0). \end{aligned}$$

Similarly, for the maximum along the Sell line we have that

$$\begin{aligned} &V(t, x_0 + (1-\lambda)\beta, y_0 - \beta) = V(t, x_0, y_0) \\ &+ \int_0^\beta \big[(1-\lambda)V_x(t, x_0 + (1-\lambda)s, y_0 + s) - V_y(t, x_0 - (1-\lambda)s, y_0 + s)\big] ds \\ &\qquad \leq V(t, x_0, y_0). \end{aligned}$$

Consequently, $\mathcal{M}V(t, x_0, y_0) - V(t, x_0, y_0) \leq 0$. Since the point (x_0, y_0) was chosen arbitrary, this holds for every point (x, y) in the domain of V.

The second part. Assume (38) holds. Chose any point (x_0, y_0). Then for any point along the Buy line starting in (x_0, y_0) we have that

$$V(t, x_0 - (1+\lambda)h, y_0 + h) \le V(t, x_0, y_0),$$

and for any point along the Sell line starting in (x_0, y_0) we have that

$$V(t, x_0 + (1-\lambda)h, y_0 - h) \le V(t, x_0, y_0),$$

where h is an arbitrary positive real number. Allowing $h \to 0$ we obtain that

$$\lim_{h\to 0} \frac{1}{h}\big[V(t, x_0 - (1+\lambda)h, y_0 + h) - V(t, x_0, y_0)\big] \\ = -(1+\lambda)V_x(t, x_0, y_0) + V_y(t, x_0, y_0) \le 0$$

$$\lim_{h\to 0} \frac{1}{h}\big[V(t, x_0 + (1-\lambda)h, y_0 - h) - V(t, x_0, y_0)\big] \\ = (1-\lambda)V_x(t, x_0, y_0) - V_y(t, x_0, y_0) \le 0$$

Again, since the point (x_0, y_0) was chosen arbitrary, this holds for every point (x, y) in the domain of V. □

Even though the two different formulations of the same problem, (32) and (10), yield the same result, the latter has direct implications for the practical realization of a numerical procedure.

Consequently, to solve the problem numerically, the following discretization scheme is proposed for the QVI (10) for the case with proportional transaction costs only:

$$\mathbb{V}^{\Delta t} = \mathcal{O}(\Delta t)\mathbb{V}^{\Delta t}, \tag{39}$$

where $\mathcal{O}(\Delta t)$ is an operator given by

$$\begin{aligned} \mathcal{O}(\Delta t)\mathbb{V}^{\Delta t} = \max\Big\{ &\max_m \mathbb{V}^{\Delta t}(t_i, x - (1+\lambda)m\Delta y, y + m\Delta y), \\ &\max_m \mathbb{V}^{\Delta t}(t_i, x + (1-\lambda)m\Delta y, y - m\Delta y), \\ &E\{\mathbb{V}^{\Delta t}(t_{i+1}, x\rho, y\varepsilon)\}\Big\}, \end{aligned} \tag{40}$$

where m runs through the positive integer numbers, and

$$\mathbb{V}^{\Delta t}(t_i, x-(1+\lambda)m\Delta y, y+m\Delta y) \\ = E\left\{\mathbb{V}^{\Delta t}(t_{i+1}, (x-(1+\lambda)m\Delta y)\rho, (y+m\Delta y)\varepsilon)\right\} \quad (41)$$

$$\mathbb{V}^{\Delta t}(t_i, x+(1-\lambda)m\Delta y, y-m\Delta y) \\ = E\left\{\mathbb{V}^{\Delta t}(t_{i+1}, (x+(1-\lambda)m\Delta y)\rho, (y-m\Delta y)\varepsilon)\right\}, \quad (42)$$

as at time t_i we do not know yet the value function. Instead, we use the known values at the next time instant, t_{i+1}.

Theorem 5. *The solution $\mathbb{V}^{\Delta t}$ of (39) converges locally uniformly to the unique continuous constrained viscosity solution of (10) as $\Delta t \to 0$*

The proof is similar to the proof of Theorem (2).

Also in the discrete time framework the dynamics of y through time is independent of x. Therefore the alternative discretization scheme for the function $\mathbb{Q}^{\Delta t}(t, y)$ is given by

$$\mathbb{Q}^{\Delta t}(t_i, y) = \max\Big\{ \max_m \exp\left(\gamma \frac{(1+\lambda)m\Delta y}{\delta(T, t_i)}\right) \mathbb{Q}^{\Delta t}(t_i, y+m\Delta y), \\ \max_m \exp\left(\gamma \frac{-(1-\lambda)m\Delta y}{\delta(T, t_i)}\right) \mathbb{Q}^{\Delta t}(t_i, y-m\Delta y), \quad (43) \\ E\{\mathbb{Q}^{\Delta t}(t_{i+1}, y\varepsilon)\}\Big\}.$$

As in the continuous time case, if the value function $\mathbb{Q}^{\Delta t}(t_i, y)$ is known in the NT region, then it can be calculated in the Buy and Sell region by using the discrete space versions of (36). The discrete time program for the value function $\mathbb{V}$ (or $\mathbb{Q}$) in the model with proportional costs only is much easier to compute as compared to its counterpart in the model with additional fixed cost component. The reason is that in this case the optimal transaction policy is simpler.

To illustrate the optimal policy we provide numerical calculations (see Figures (1) and (2)) with the following data: $\mu = 10\%$, $r = 5\%$, $\sigma = 20\%$ (all in annualized terms), $\gamma = 0.001$. Note, that in all our graphical illustrations here and thereafter we always assume that the horizontal axis, say τ-axis, represents the time remaining until the terminal date, i.e., $\tau = T - t$.

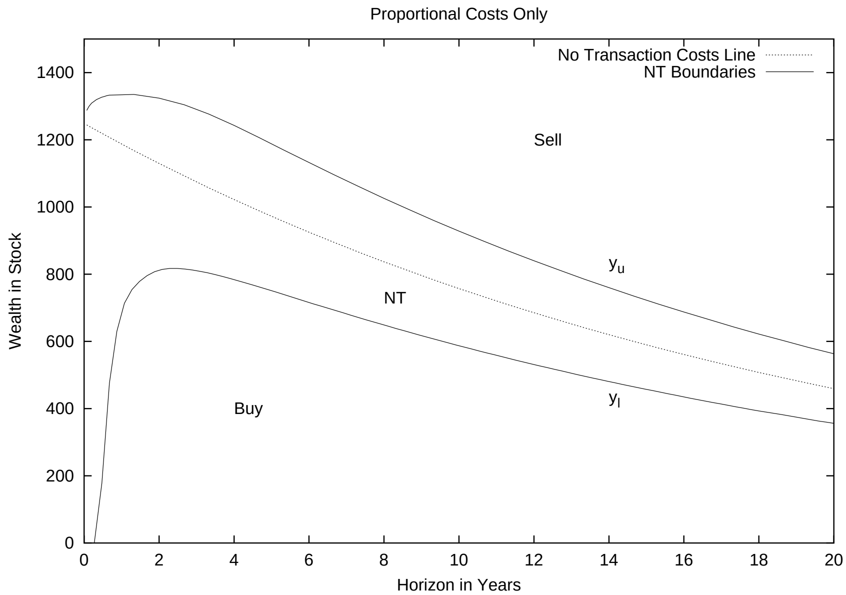

Figure 1: NT boundaries as functions of the investor's horizon for $\lambda = 0.01$

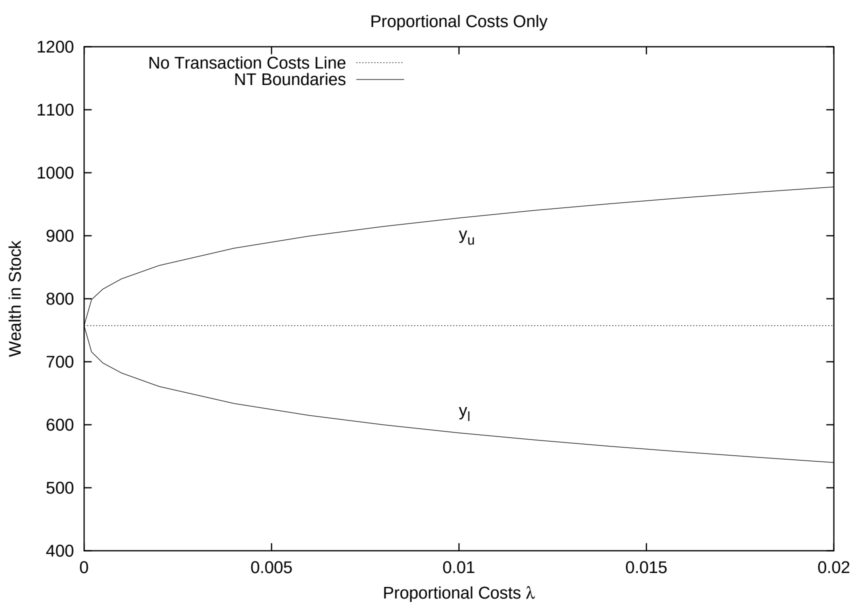

Figure 2: NT boundaries as functions of the proportional transaction costs at 10 year horizon

As it is seen from Figure (1), as the terminal date approaches, the NT region widens and shifts below with respect to the no transaction costs line defined by equation (31). And as the terminal date retreats, the NT region narrows and lies quite symmetrically around the no transaction costs line. Figure (2) shows that the introduction of even very low transaction costs has a significant impact on the optimal portfolio strategy. As the transaction costs increase the lower NT boundary comes down and the upper boundary goes up. The sensitivity of a NT boundary to the level of proportional transaction costs is very high for very low levels of λ. As λ increases, this sensitivity decreases drastically.

It is well known that if the drift of the risky asset is greater than the risk-free rate of return, $\mu > r$, and there are no transaction costs, then an investor will always invest some amount of his wealth in the risky asset no matter how short his investment horizon is (see, for example, equation (31)). And the optimal policy requires continuous rebalancing of the investor's portfolio. The introduction of transaction costs associated with the trade in the risky asset will pull down the investor's actual returns. The investor's optimal policy has to balance the costs against the benefits of trading. In the presence of transaction costs the investor modifies drastically the frequency and size of his trades. It is easy to understand that with transaction costs, even if $\mu > r$, not all investment horizons are feasible; that is, for some "short" investment horizons the net[3] expected returns from the investment in the risky asset are negative. That is why we introduce the following definition:

Definition 1. The *minimal feasible investor's horizon* in the presence of transaction costs is defined as the investment horizon over which the investor is indifferent between the two alternatives: (i) a utility maximization problem where he trades in *both* the riskless and risky assets, and (ii) when he refrains from trading in the risky asset and invests all his wealth in the riskless asset.

In other words, the minimal feasible investor's horizon in the presence of transaction costs is defined as the shortest investment horizon over which the net expected returns from the investment in the risky asset are greater than zero. For the chosen model parameters the minimal feasible investor's

[3]With the correction for transaction costs

horizon is approximately 0.4 year.

6 Fixed Transaction Costs

Here we consider the case with fixed transaction costs only ($\lambda = 0$, $k > 0$). The problem can be formulated in the same manner as in Section 2 with the correction for zero proportional transaction costs.

Again, let's for the moment write the investor's value function as $V(t, \gamma, x, y, k)$. By this we want to emphasize that the value function depends on the investor's coefficient of absolute risk aversion and the fixed transaction fee.

Proposition 2. *For an investor with the exponential utility function with absolute risk aversion γ, an initial endowment (x, y), and a fixed transaction fee k we have*

$$V(t, \gamma, x, y, k) = V(t, \theta\gamma, \frac{x}{\theta}, \frac{y}{\theta}, \frac{k}{\theta}), \tag{44}$$

where θ is any positive real number.

Proof. As in the proof of Proposition (1), this relationships can be easily established from the form of the exponential utility function. In particular, the portfolio process $\{\frac{x_s}{\theta}, \frac{y_s}{\theta}; s > t\}$ is admissible given the initial portfolio $(\frac{x_t}{\theta}, \frac{y_t}{\theta})$ and fixed transaction cost fee $\frac{k}{\theta}$ if and only if $\{x_s, y_s; s > t\}$ is admissible given the initial portfolio (x_t, y_t) and fixed transaction cost fee k. Furthermore, $U(\gamma, X_T) = U(\theta\gamma, \frac{X_T}{\theta})$.

Proposition (2) establishes a property that is not so useful as the corresponding property of the value function in the problem with proportional transaction costs only. The obstacle is that in re-scaling an optimal portfolio strategy from some γ to another γ' one needs to re-scale a fixed transaction fee as well. Thus, the possible scope of application of this property is somewhat limited.

The numerical calculations show that most of the time the portfolio space again can be divided into three disjoint regions, which can be specified as the Buy region, the Sell region, and the NT region. If a portfolio lies either in the Buy region or in the Sell region, the optimal strategy is to buy/sell the risky asset until the portfolio reaches the so-called "target" boundary. However, there is generally a time interval, say $[\tau_1, \tau_2)$, which is usually close to the terminal date, when the NT region consists of two disjoint sub-regions which, in their turn, divide either the Buy region (when $\mu > r$) or the Sell region (when $\mu < r$) into two parts. Nevertheless, as in the former case, the target boundary is unique. All these boundaries are functions of the investor's horizon and do not depend on the investor's holdings in the bank account, so that a possible description of the optimal policy for $\tau \in (0, \tau_1) \cup [\tau_2, \infty)$ may be given by

$$\begin{aligned} y &= y_u(\tau) \\ y &= y^*(\tau) \\ y &= y_l(\tau), \end{aligned} \tag{45}$$

where the first and the third equations describe the upper and the lower NT boundaries respectively, and the second equation describes the target boundary. For $\tau \in [\tau_1, \tau_2)$ a possible description of the optimal policy may be given by

$$\begin{aligned} y &= y_u(\tau) \\ y &= y^*(\tau) \\ y &= y_l(\tau) \\ y &= y_{2u}(\tau) \\ y &= y_{2l}(\tau) = 0. \end{aligned} \tag{46}$$

The first and the third equations describe the upper and the lower boundaries of the main NT sub-region. The second equation describes the target boundary. The last two equation characterize the minor NT sub-region which lies in between $y = y_{2u}(\tau) < k$ and $y = y_{2l}(\tau) = 0$.

To illustrate the optimal policy we provide numerical calculations with the following data: $\mu = 10\%$, $r = 5\%$, $\sigma = 20\%$ (all in annualized terms), and $\gamma = 0.001$. Figure (3) plots the NT and target boundaries as functions of the investor's horizon. Figure (5) presents a schematic sketch of the optimal portfolio policy in the time interval where the NT region consists of two disjoint sub-regions. Figure (4) plots the NT and target boundaries as functions of the fixed transaction fee at 10 year horizon.

Figure (3) together with Figure (4) show that the presence of even very small fixed transaction costs has a tremendous impact on the optimal portfolio policy as compared to the case with no transaction costs. In the same manner as in the previous case with proportional transaction costs only, as

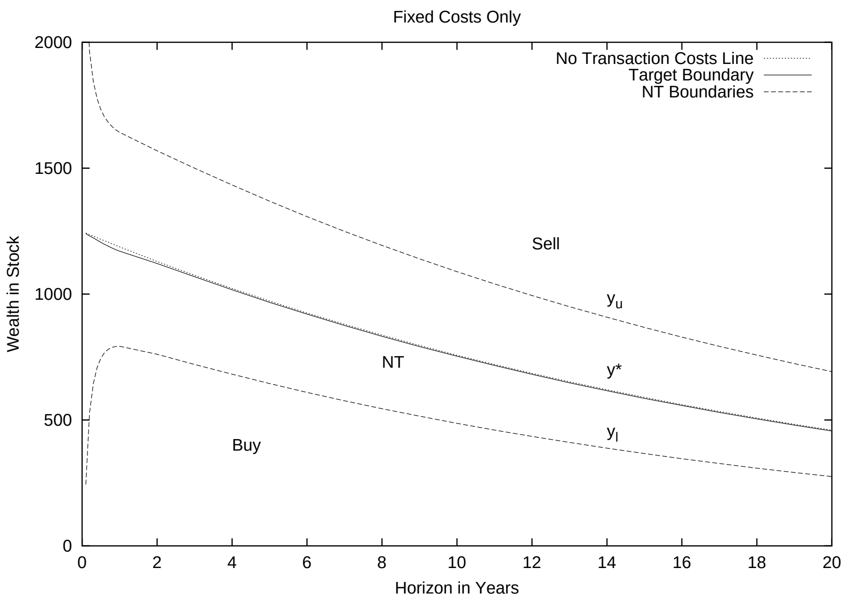

Figure 3: NT and target boundaries as functions of the investor's horizon for $k = 2$

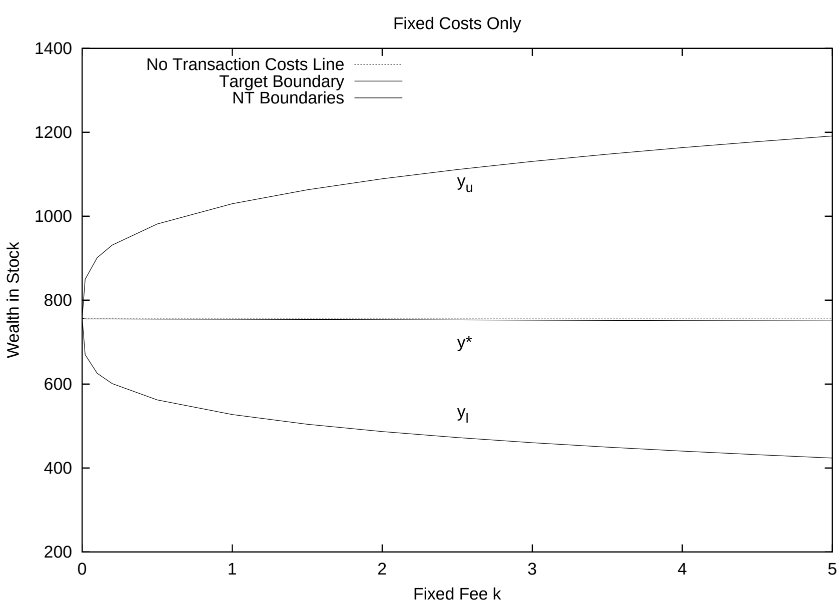

Figure 4: NT and target boundaries as functions of the fixed transaction fee at 10 year horizon

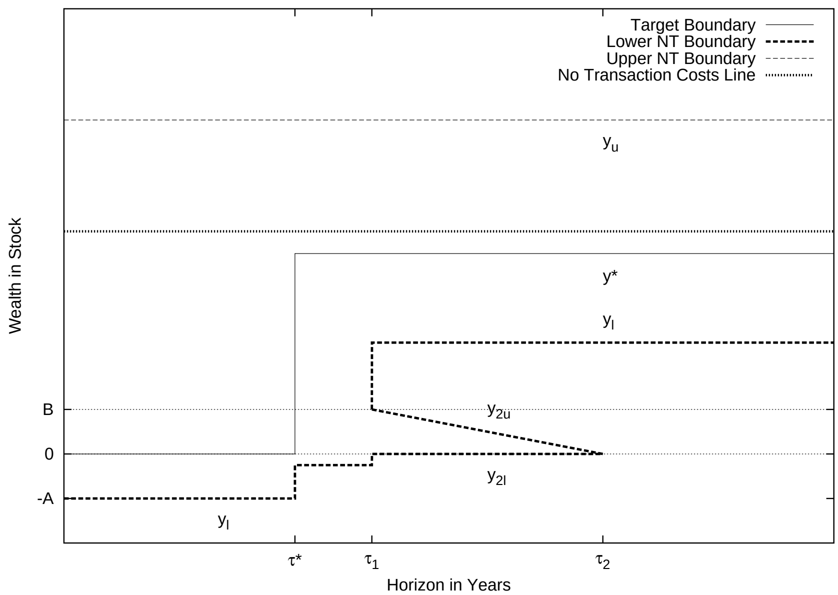

Figure 5: A schematic sketch of the optimal portfolio policy with two NT sub-regions

the transaction costs increase the lower NT boundary comes down and the upper NT boundary goes up. The sensitivity of a NT boundary to the level of fixed transaction costs is extremely high for very low levels of k. As k increases, this sensitivity decreases drastically. On the contrary, the target boundary is almost insensitive to the level of transaction costs. Note also that as the terminal date approaches, the NT region widens, and as the terminal date retreats, the NT region narrows. Independently of horizon the NT region lies more or less symmetrically around the no transaction costs line. However, the distance from the upper NT boundary to the no transaction costs line is noticeable greater than the corresponding distance from the lower NT boundary. The target boundary lies a bit below the no transaction costs line.

We turn now to the study of the optimal portfolio policy in the case when the NT region consists of two disjoint sub-regions (see Figure (5)). This picture is typical for investment horizons that are slightly greater than the minimal feasible investor's horizon and when transaction costs have a fixed fee component. Note that τ_2 is actually the minimal feasible investor's horizon.

In the time interval $(0, \tau^*)$, which is closest to the terminal date, the target boundary is equal to zero. The intuitive explanation for this is as follows. Since the terminal date is very close and in case a portfolio lies outside of the NT region, it is better for the investor to sell all shares of the stock or to close the short position in the stock immediately than to transact to a certain level of holdings in the stock and after a short while to eliminate his holdings in the stock. Note that in the former case the investor pays transaction costs once, but in the latter case at least twice. In the same time interval the lower NT boundary lies slightly below zero. The rationale for this is that for very low negative holdings in the stock it is more reasonable to close the short position in the stock and, thus, pay the transaction costs at the terminal date than to do it right away. That is, the investor expects to gain more by holding the transaction costs on the bank account up to the terminal date, than he possibly loses if the stock price increases.

At time τ^* the target boundary jumps closer to the no-transaction cost line given by equation (31). At the same time the lower NT boundary moves closer to zero.

At time τ_1 the NT region splits into two disjoint sub-regions. The main NT sub-region is located in between $y = y_u(\tau)$ and $y = y_l(\tau)$, and the minor NT sub-region lies in between $y = y_{2u}(\tau)$ and $y = 0$. Again, the rationale for the existence of a second (minor) NT sub-region can be explained in terms of fixed transaction costs. Recall how we define the investor's net wealth (see equation (4)). If the investor's holdings in the stock are positive, he will sell all his shares of the stock at the terminal date only if the proceeds are positive after transaction costs. Putting it another way, the rational investor will not sell his shares of the stock if $y(1 - \lambda) < k$.

Any investment horizon which is shorter than τ_2 is not feasible. That is, the net expected returns[4] from the investment in the risky stock[5] are negative. But what if the investor at time $\tau \in [\tau_1, \tau_2)$ has some holdings in the stock y_τ? Putting it another way, what is the difference between having $0 < y_\tau < y_{2u}(\tau)$ and having $y_{2u}(\tau) < y_\tau < y_l(\tau)$ in the stock?

Suppose for the moment that $y_\tau \to 0^+$. Consider the two alternatives: (*i*) No trade at τ and thereafter up to the terminal date, and (*ii*) buy a

[4]Taking into account the round trip transaction costs

[5]For an investor starting with zero holdings in the stock

certain number of shares of the stock at τ in order to move closer to the optimal level of holdings in the model with no transaction costs. In the former case it is almost sure that on the terminal date the holdings in the stock will not exceed the fixed transaction fee k. That is, $y_{0+} < k$ *a.s.*, and, thus, it is not optimal to sell shares of the stock. Hence, in the first alternative the investor does not pay any transaction costs. In the second alternative the investor pays at least round trip transaction costs equal to $2k$ (we ignore the time value of money). It turns out that the first alternative is better than the second one as the investment horizon is not feasible.

Suppose now that y_τ is sufficiently high but still less than $y_l(\tau)$ so that without any trade at τ and thereafter it will be most likely that $y_{0+} > k$. Consider again the same two alternatives as above. Following the first alternative the investor will sell all his shares of the stock at the terminal date paying k in the transaction costs. In the second alternative, when the investor buys additional shares of the stock at τ, the round trip transaction costs will amount to $2k$, but the difference in the amounts of transaction costs between the two alternatives will be only k. This makes the second alternative better than the first one for $\tau \in [\tau_1, \tau_2)$ and high enough initial holdings in the stock.

As time passes, the net expected returns from holding some amount in the risky asset increase. That is why the upper boundary of the minor NT sub-region, $y_{2u}(\tau)$, decreases in the time interval $[\tau_1, \tau_2)$. The values of A and B in Figure (5) are commensurable with the fixed transaction fee k. Under chosen model parameters the minimal feasible investment horizon is approximately 0.1 year.

7 Proportional and Fixed Transaction Costs

Here we consider the case with proportional and fixed transaction costs ($\lambda > 0$ and $k > 0$). Similar to the previous sections, let's for the moment write the investor's value function as $V(t, \gamma, x, y, \lambda, k)$. By this we want to emphasize that the value function depends on the investor's coefficient of absolute risk aversion, the level of proportional transaction costs, and the fixed transaction fee.

Proposition 3. *For an investor with the exponential utility function with absolute risk aversion γ, an initial endowment (x, y), proportional costs λ,* *and a fixed transaction fee k we have*

$$V(t,\gamma,x,y,\lambda,k) = V(t,\theta\gamma,\frac{x}{\theta},\frac{y}{\theta},\lambda,\frac{k}{\theta}), \tag{47}$$

where θ is any positive real number.

Proof. The proof is similar to the proofs of Propositions (1) and (2).

The numerical calculations show that most of the time the portfolio space again can be divided into three disjoint regions (Buy, Sell, and NT), but four (instead of three as in the previous section) boundaries describe the optimal policy. As before, the Buy and the NT regions are divided by the lower no-transaction boundary, and the Sell and the NT regions are divided by the upper no-transaction boundary. If a portfolio lies in the Buy region, the optimal strategy is to buy the risky asset until the portfolio reaches the Buy target boundary. Similarly, if a portfolio lies in the Sell region, the optimal strategy is to sell the risky asset until the portfolio reaches the Sell target boundary.

However, as in the case with fixed transaction costs only, there is generally a time interval, $[\tau_1,\tau_2)$, when the NT region consists of two disjoint sub-regions which, in their turn, divide either the Buy region (when $\mu > r$) or the Sell region (when $\mu < r$) into two parts. Nevertheless, as in the former case, the target boundaries are unique. All these boundaries are functions of the investor's horizon and do not depend on the investor's holdings on the bank account, so that a possible description of the optimal policy for $\tau \in (0,\tau_1) \cup [\tau_2,\infty)$ may be given by

$$\begin{aligned} y &= y_u(\tau) \\ y &= y_l^*(\tau) \\ y &= y_u^*(\tau) \\ y &= y_l(\tau), \end{aligned} \tag{48}$$

where the first and the forth equations describe the upper and the lower NT boundaries respectively, and the second and the third equations describe the target boundaries. For $\tau \in [\tau_1,\tau_2)$ a possible description of the optimal policy may be given by

$$\begin{aligned} y &= y_u(\tau) \\ y &= y_l^*(\tau) \\ y &= y_u^*(\tau) \\ y &= y_l(\tau) \\ y &= y_{2u}(\tau) \\ y &= y_{2l}(\tau) = 0. \end{aligned} \tag{49}$$

The first and the forth equations describe the upper and the lower boundaries of the main NT sub-region. The second and the third equations describe the target boundaries. The last two equation characterize the minor NT sub-region which lies in between $y = y_{2u}(\tau) < k$ and $y = y_{2l}(\tau) = 0$.

Figure (6) illustrates the optimal strategy for $\mu = 10\%$, $r = 5\%$, $\sigma = 20\%$ (all in annualized terms), $\gamma = 0.001$, $\lambda = 0.01$, $k = 0.5$, and 20 year horizon. Figure (7) presents a schematic sketch of the optimal portfolio policy in the time interval where the NT region consists of two disjoint sub-regions. Figure (8) plots the NT and target boundaries as functions of the proportional transaction costs at 10 year horizon. Figure (9) plots the NT and target boundaries as functions of the fixed transaction fee at 10 year horizon.

Some important properties of the optimal policy are as follows (see Figure (6)). The NT boundaries are found to be wider than those in the model with proportional transaction costs only (as one quite logically expects). As the terminal date approaches, the NT region widens and shifts below. On the contrary, and as the terminal date retreats, the NT region narrows and lies more or less symmetrically around the no transaction costs line.

The schematic sketch of the optimal portfolio policy in the time interval where the NT region consists of two disjoint sub-regions (see Figure (7)) is basically the same as in the previous section. The main difference is in that there exists two target boundaries instead of one, and the Buy target boundary jumps to zero earlier (at time τ_l^*) than the Sell target boundary (at time τ_u^*) as the terminal date approaches. Moreover, τ_l^* practically coincides with τ_1. Under chosen model parameters the minimal feasible investment horizon is approximately 0.5 year, and $\tau_2 - \tau_1$ is about two weeks.

Note from Figure (9) that as the fixed transaction fee tends to zero, y_u and y_u^* converge to y_u in the model with proportional transaction costs only. Respectively, y_l and y_l^* converge to y_l in the model with proportional

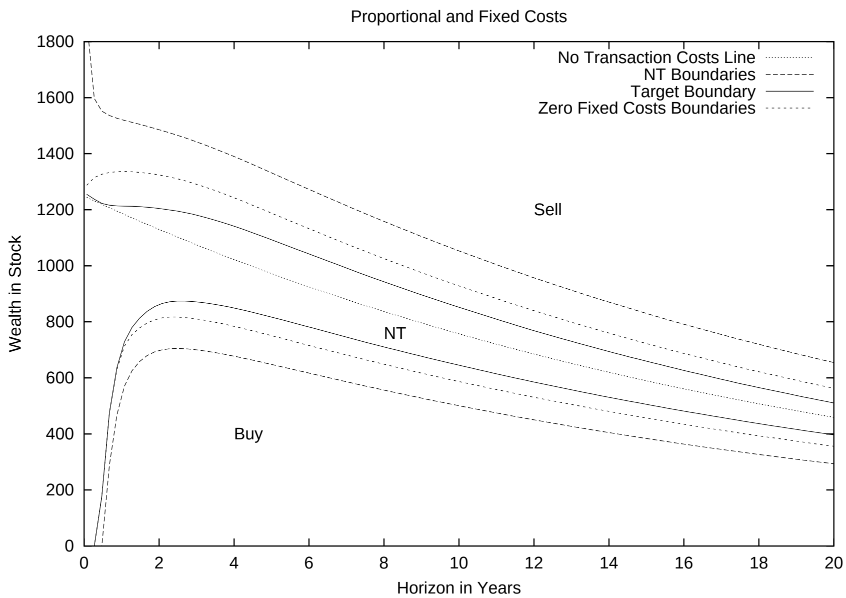

Figure 6: NT and target boundaries as functions of the investor's horizon for $\lambda = 0.01$ and $k = 0.5$

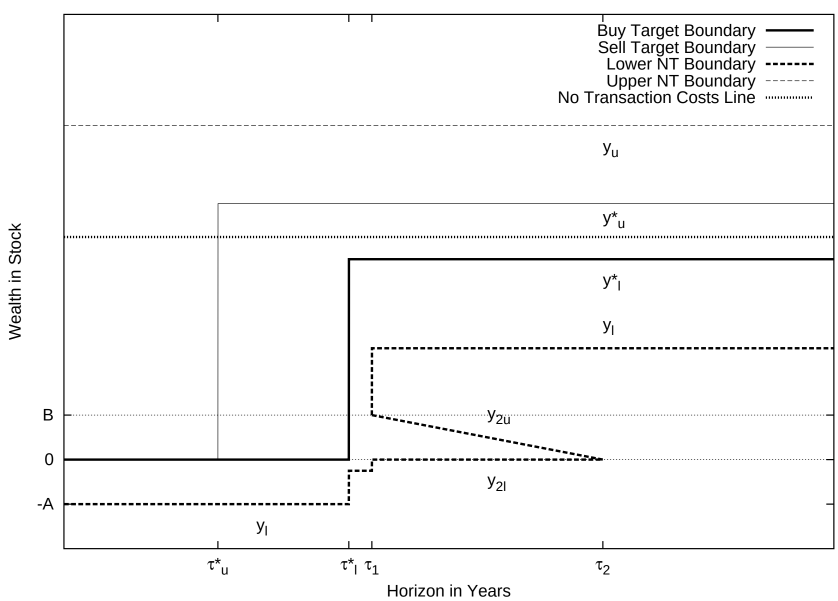

Figure 7: A schematic sketch of the optimal portfolio policy with two NT sub-regions

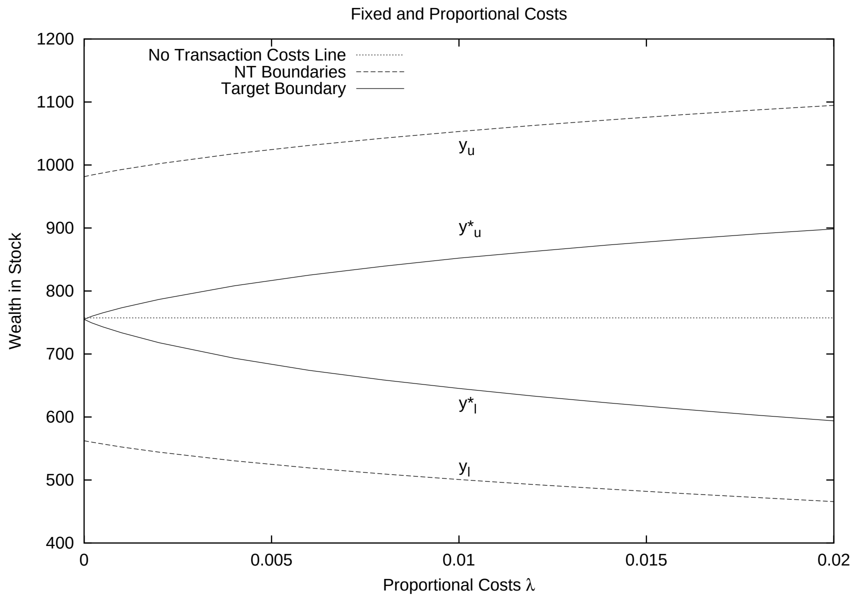

Figure 8: NT and target boundaries as functions of the proportional transaction costs at 10 year horizon and $k = 0.5$

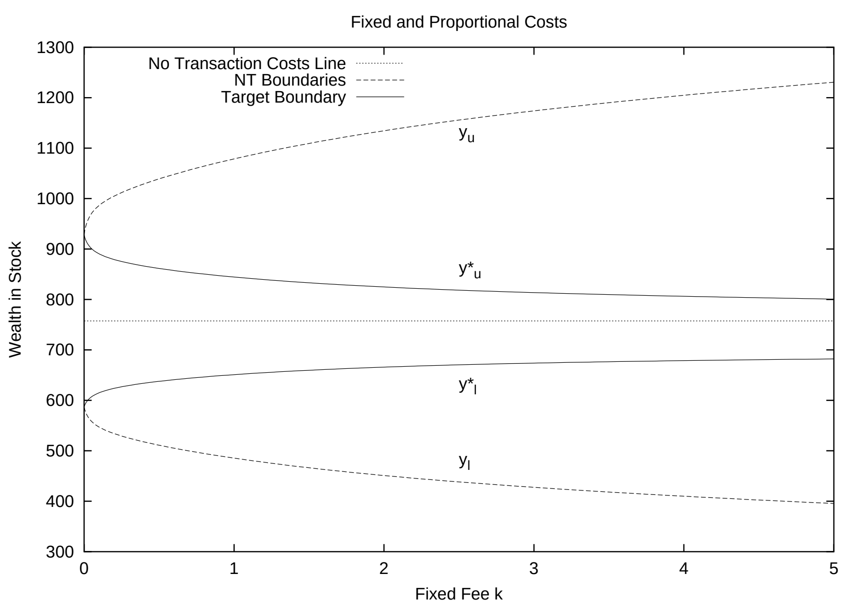

Figure 9: NT and target boundaries as functions of the fixed transaction fee at 10 year horizon and $\lambda = 0.01$

transaction costs only. In addition, as the fixed transaction fee increases, the NT region widens and both y_l^* and y_u^* tends to y^* in the model with fixed transaction costs only. Similarly, as it is seen from Figure (8), as the proportional transaction costs approach zero, y_l^* and y_u^* converge to y^* in the model with fixed transaction costs only. Furthermore, as the proportional costs increase, the NT region widens as well as the distance between the target boundaries.

The careful comparative statics analysis of the behavior of the NT and target boundaries is beyond the scope of this paper. Indeed, every boundary is a function of many parameters and may be written as

$$y = y(\tau, \mu, r, \sigma, \gamma, \lambda, k)$$

To do the comparative statics for every single parameter would take a huge amount of space. Besides, the presence of four boundaries makes this task rather cumbersome. Therefore we only combine comparative statics analysis for some important parameters such as the volatility σ, the drift μ, and the absolute risk aversion coefficient γ. In addition, we have to choose some benchmark to make the comparisons. For this purpose we use the deviation of a boundary from the no transaction cost line given by (31).

Our combined comparative statics analysis is based on the following idea. In the absence of transaction costs, the amount invested in the risky asset is defined by equation (31). One can note that either doubling the volatility σ^2, the absolute risk aversion coefficient γ, or halving the risk premium $\mu - r$ has a similar effect on the optimal policy. Namely, the investor halves the amount invested in the risky asset. But what happens with the optimal policy in the presence of both fixed and proportional transaction costs?

Figures (10) and (11) present some results of the comparison of the NT and target boundaries for different volatilities, drifts, and relative risk aversion coefficients. The benchmark parameters are $\mu = 10\%$, $ARA = 0.001$, and $\sigma = 20\%$. The rest of the parameters are $r = 5\%$, $k = 0.5$, $\lambda = 0.01$.

The analysis shows that either doubling the volatility, the ARA, or halving the risk premium has similar general consequences. The NT region narrows that causes more frequent transactions. At the same time the NT region shifts downwards causing the investor to move out of the risky stock

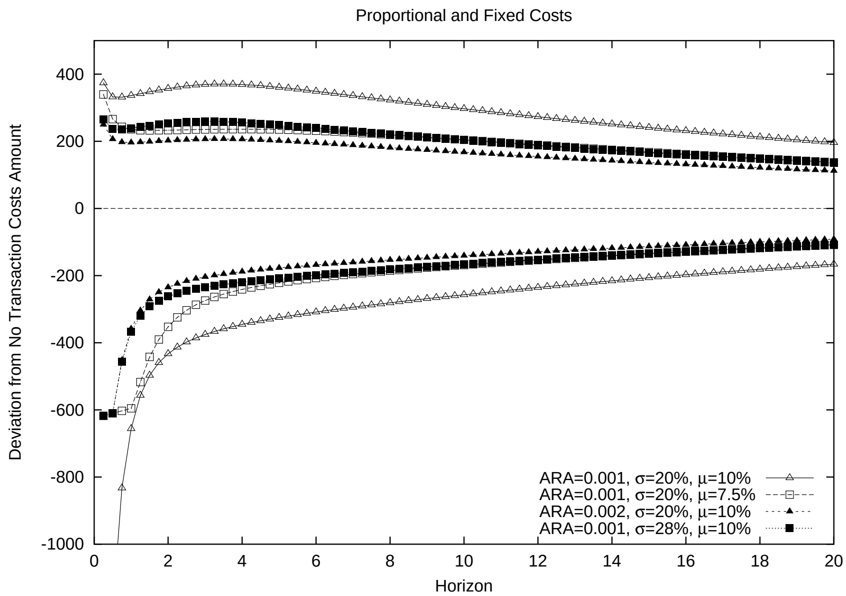

Figure 10: NT boundaries as functions of the investor's horizon

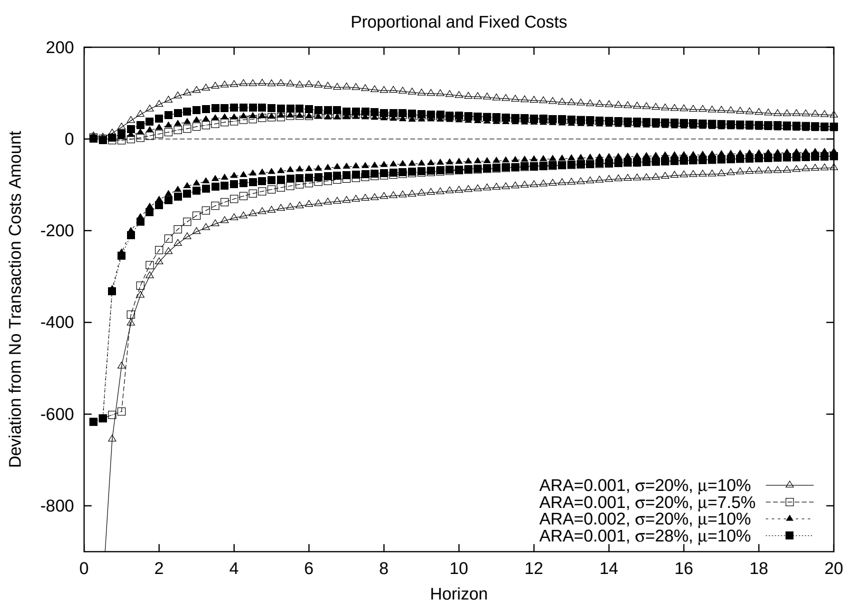

Figure 11: Target boundaries as functions of the investor's horizon

and into the riskless asset. The target boundaries also move closer to the no transaction costs line. In particular, doubling either the volatility or halving the risk premium has almost the same effect on the NT and target boundaries, especially for sufficiently long investment horizons. As compared to these, doubling the ARA produces a narrower NT region and the target boundaries lie closer to the no transaction costs line.

8 Conclusions and Extensions

In this paper we studied the optimal portfolio selection problem for a constant absolute risk averse investor who faces fixed and/or proportional transaction costs and maximizes expected utility of end-of-period wealth. We used a continuous time model and applied the method of the Markov chain approximation to solve numerically for the optimal trading policy without making any apriory assumptions on its shape.

The numerical solution indicates that, most of the time, the portfolio space is divided into three disjoint regions (Buy, Sell, and No-Transaction), and four boundaries describe the optimal policy. If a portfolio lies in the Buy region, the optimal strategy is to buy the risky asset until the portfolio reaches the Buy target boundary. Similarly, if a portfolio lies in the Sell region, the optimal strategy is to sell the risky asset until the portfolio reaches the Sell target boundary. However, we discovered that for a CARA investor with a finite horizon there is generally a time interval, close to the terminal date, when the NT region consists of two disjoint sub-regions which, in their turn, divide either the Buy region or the Sell region into two parts. Nevertheless, as in the former case, the target boundaries are unique.

We examined the effects on the optimal policy from varying volatility, drift, ARA, and the level of transaction costs. Some important properties of the optimal policy are as follows: As the terminal date approaches, the NT region widens and shifts below. On the contrary, as the terminal date retreats, the NT region narrows and lies more or less symmetrically around the no transaction costs line. The presence of even very small transaction costs has a tremendous impact on the optimal portfolio policy as compared to the case with no transaction costs. Namely, the investor drastically reduces the frequency and sizes of his trades. As transaction costs increase, the NT region widens. Our analysis showed that either doubling the volatility, the ARA, or halving the risk premium has similar general consequences: The NT region narrows that causes more frequent transactions. At the same time the NT region shifts downwards causing the investor to move out of the risky stock and into the riskless asset. The target boundaries also move closer to the no transaction costs line.

The approach of this paper may be generalized in a straightforward manner to incorporate intermediate consumption, more general utility functions, and a more general structure of transaction costs. Another interesting extension would be the case of two or more risky assets. Finally, the utility maximization approach and the numerical technique used in this paper may be successfully applied to price options in markets with both fixed and proportional transaction costs[6].

[6]this approach was pioneered by Hodges and Neuberger (1989)

References

Akian, M., Menaldi, J. L., and Sulem, A. (1996). "On an Investment-Consumption Model with Transaction Costs", *SIAM Journal of Control and Optimization*, *34*(1), 329–364.

Atkinson, C. and Wilmott, P. (1995). "Portfolio Management with Transaction Costs: an Asymptotic Analysis of the Morton and Pliska Model", *Mathematical Finance*, *5*, 395–422.

Chancelier, J.-P., Øksendal, B., and Sulem, A. (2000). "Combined Stochastic Control and Optimal Stopping, and Application to Numerical Approximation of Combined Stochastic and Impulse Control", Preprint, Department of Mathematics, University of Oslo.

Collings, P. and Haussmann, U. G. (1998). "Optimal Portfolio Selection with Transaction Costs", In:Proceedings of the Conference on Control of Distributed and Stochastic System. Kluwer.

Cox, J. M., Ross, S. A., and Rubinstein, M. (1979). "Option Pricing: A Simplified Approach", *Journal of Financial Economics*, *7*, 229–263.

Davis, M. H. A. and Norman, A. R. (1990). "Portfolio Selection with Transaction Costs", *Mathematics of Operations Research*, *15*(4), 676–713.

Davis, M. H. A., Panas, V. G., and Zariphopoulou, T. (1993). "European Option Pricing with Transaction Costs", *SIAM Journal of Control and Optimization*, *31*(2), 470–493.

Dumas, B. and Luciano, E. (1991). "An Exact Solution to a Dynamic Portfolio Choice Problem under Transaction Costs", *Journal of Finance*, *XLVI*(2), 577–595.

Eastham, J. and Hastings, K. (1988). "Optimal Impulse Control of Portfolios", *Mathematics of Operations Research*, *13*, 588–605.

Fitzpatrick, B. G. and Flemming, W. H. (1991). "Numerical Methods for an Optimal Investment-Consumption Model", *Mathematics of Operations Research*, *16*, 823–841.

He, H. (1990). "Convergence from Discrete to Continuous Time Contingent Claim Prices", *Review of Financial Studies*, *3*, 523–546.

Hodges, S. D. and Neuberger, A. (1989). "Optimal Replication of Contingent Claims under Transaction Costs", *Review of Futures Markets*, *8*, 222–239.

Korn, R. (1998). "Portfolio Optimization with Strictly Positive Transaction Costs and Impulse Controls", *Finance and Stochastics*, *2*, 85–114.

Kushner, H. J. and Dupuis, P. G. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, New York.

Kushner, H. J. (1977). *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York.

Kushner, H. J. (1990). "Numerical Methods for Stochastic Control Problems in Continuos Time", *SIAM Journal of Control and Optimization*, *28*, 999–1048.

Liu, H. (2001). "Optimal Consumption and Investment with Transaction Costs and Multiple Stocks", Working Paper, The Olin School of Business, Washington University.

Merton, R. C. (1969). "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case", *Review of Economics and Statistics*, *51*, 247–257.

Merton, R. C. (1971). "Optimum Consumption and Portfolio Rules in a Continuous-Time Model", *Journal of Economic Theory*, *3*, 373–413.

Merton, R. C. (1973). "An Intertemporal Capital Asset Pricing Model", *Econometrica*, *41*(5), 867–887.

Øksendal, B. and Sulem, A. (1999). "Optimal Consumption and Portfolio with both Fixed and Proportional Transaction Costs", Preprint, Department of Mathematics, University of Oslo.

Schroder, M. (1995). "Optimal Portfolio Selection with Fixed Transaction Costs: Numerical Solutions", Working Paper, Michigan State University.

Shreve, S. and Soner, H. M. (1994). "Optimal Investment and Consumption with Transaction Costs", *The Annals of Applied Probability*, *4*, 609–692.

Whalley, A. E. and Wilmott, P. (1997). "An Asymptotic Analysis of an Optimal Hedging Model for Option Pricing with Transaction Costs", *Mathematical Finance*, *7*(3), 307–324.