Portfolio Selection and Equilibrium

Stock Returns with Quadratic Transaction Costs

by

Itzhak Venezia
School of Business Administration
Hebrew University of Jerusalem

July 1995

The author acknowledges the financial support of the Krueger Center of Finance and the Galanter Foundation.
Abstract

This paper analyzes optimal portfolio selection and equilibrium stock returns when there are transaction costs. These costs are approximated as a quadratic function of the volume of transactions. The transaction costs depend on deviations from the previously held portfolio, and hence current actions affect future decisions and a dynamic setting is required for the analysis.

The dependence between current and past portfolio is investigated, as well as the autocorrelation between successive market returns induced by the transaction costs.

Since the frequency of revisions of the portfolio depend on transaction costs, this paper also sheds some light on the issue of investment horizon.
I. Introduction

The effect of transaction costs on the portfolio choice of investors and on equilibrium returns of risky assets is an important issue which has been dealt with extensively in the literature. However, exact solutions to the problem are hard to come by because of the discontinuities that transaction costs introduce into the mathematical modelling. Consequently, most of the papers studying this issue had to make some strong simplifying assumptions. For example, Amihud and Mendelson [1986] assume away short sales and their representative investor is not risk averse. Constantinides [1986] assumes that there is only one risky asset, and that transaction costs are proportional. It is well known, however, that transaction costs are usually concave (see Sharpe and Alexander [1990]).

To overcome some of these problems we approximate the transaction costs by a quadratic function of the volume of transactions. This assumption of quadratic costs is made frequently in the analysis of production systems and is a usually a good one. The strong part of this assumption is that it requires no fixed costs. However, the quadratic part, which introduces the desired concavity, mitigates this problem.

In our model, investors choose a portfolio so as to maximize a quadratic expected utility. Transaction costs depend on deviations from the existing portfolio. It, hence follows that current decisions affect future costs, because future transaction costs depend on current portfolio. In the analysis we use a simple two-period model to analyze equilibrium returns, and a more general multiperiod model to investigate issues of portfolio selection, diversification, and frequency of trades.

The paper is constructed as follows: In section II we present the optimal portfolio in a two-period model of an investor facing quadratic transaction costs. Based on these results, we present the equilibrium returns and analyze the effect of transaction costs on these
returns. It is assumed throughout this section that there exists just one representative investor in the economy. In section III we consider the case where there are many investors. In particular, this section is designed to examine the effect of existence of different types of investors with different transaction costs on equilibrium returns. In section IV we present the decisions of the investor in a multiperiod setting. We show how the dynamic considerations affect the portfolio choice. In section V we provide numerical examples and in the last section some concluding remarks are given.

II. The Model

Consider an investor who faces $n$ risky assets and one riskless asset, out of which he or she constructs a portfolio. The objective of the investor is to maximize a multiperiod additive utility function. The utility at any period of time is quadratic in wealth and the time preference is captured by some discount factor $\rho$. The vector of mean excess (over risk free) returns at time $t$ is $\mu_t$ and covariance matrix is denoted by $V_t$. The wealth of the investor to be invested in the portfolio at the beginning of the discrete time period $t$ is $W_t$. The proportion to be invested in risky asset $i$ at time $t$ is $X_{it}$, and the vector $X_t$ denotes the $(n \times 1)$ vector of $X_{it}$'s.

Likewise, we denote by $Z_{it}$ the returns ($1 +$ rate of returns) on risky security $i$, and by $Z_t$ the $(n \times 1)$ vector of these returns. The returns on the riskless asset are denoted by $r_{Pt}$, the excess returns, $Z_{rt} - r_{Pt}$ are denoted by $y_{rt}$.

From the above notation and definitions, one obtains that if there were no transaction costs, then the wealth of the investor at the end of period $t$ is given by:
\[ (2.1) \quad W_{t+1} = W_t \left[ X_t' Z_t + \left( 1 - 2 \sum X_{t-i} \right) r_{Pt} \right] \\
= W_t \left[ X_t' (Z_t - r_{Pt} \cdot L) + r_{Pt} \right] \\
= W_t \left[ X_t' y_t - r_{Pt} \right] \]

where \( L \) denote a vector of ones.

If transaction costs are introduced, then \( W_{t+1} \) is the expression given in (2.1) minus the transaction costs. The transaction costs, however, depends on the portfolio held by the investor at the beginning of the period. This is true since transaction costs depends on the changes in the portfolio held through the volume of transactions. This volume, unfortunately cannot be express just in terms of proportions of wealth invested, since wealth changes between periods. This happens either because of returns on the previous portfolio, or because of changes in the amounts desired to be invested by the investor. The model could then be expressed in terms of number of securities and prices of securities to analyze transaction costs. However, as will next be shown, we could express the model as investment in terms of proportions of wealth, at the cost of few simplifying assumptions.

Let \( M_a \) denote the volume of transactions at an asset \( i \), and let \( N_a \) denote the number of shares of asset \( i \) held by the investor at period \( t \). \( M_a \) is hence given by:

\[ (2.2) \quad M_a = (N_a - N_{a,t-1}) P_a, \]

where \( P_a \) denotes the price of security \( i \) at time \( t \). (2.2) can be rewritten as:

\[ (2.3) \quad M_a = N_t P_a - N_{a,t-1} P_a \\
= N_i P_t - N_{i,t-1} P_{i,t-1} Z_{i,t-1} \\
= W_t X_{i,t} - W_{t-1} X_{i,t-1} Z_{i,t-1} \\
= W_t \left( X_{a,t} - \frac{W_{t-1} Z_{i,t-1}}{W_t} X_{i,t-1} \right) \\
= W_t \left( X_{a,t} - m_{a,t} X_{i,t-1} \right) \]
where

\[ m_i = \frac{W_{i-1} Z_{i,t-1}}{W_i} = \frac{Z_{t,t-1}}{W_i / W_{t-1}} \]

i.e., \( m_i \) is the relative return on security \( i \) as compared to the total increase in wealth to be invested (recall that \( W_{i+1} \) may change either due to returns or if the investor decides to change the total amount allocated to investment; this is especially relevant to institutional investors and money managers whose total portfolio depends on purchase/sales participation units by other investors).

To simplify the exposition we shall assume that \( m_i = 1 \) for all \( i \) to obtain:

\[ M_i = W_i (X_i - X_{i-1}), \]

i.e., it will be assumed that volume depends just on changes in the proportions of wealth allocated to any given security.

Given the above assumption, and assuming that transaction costs are given by a quadratic function of volume we express these costs by:

\[ \text{TC} = b' M_i - \frac{1}{2} M_i' C M_i \]

where \( b \) is an \((nx1)\) vector and \( C \) is an \((nxn)\) matrix of cost parameters. The vector \( b \) is the linear part, and \( C \) introduces a quadratic element. If there is interaction between the costs of different securities, then the off diagonal elements of \( C \) will be non-zero. The elements of \( b \), as well as the diagonal elements of \( C \) may all be the same. Transaction costs in this case can still be expressed as:

\[ C = b_0 \cdot M_i - \frac{1}{2} C_0 M_i' M_i \]

where \( b_0 \) and \( C_0 \) are some cost scalars. More generality, however, is obtained through the former formulation which we shall adopt.
Inserting (2.5) into (2.5) one obtains:

\[(2.8) \quad TC = W_t \bar{b}' (X_t - X_{t-1}) - \frac{1}{2} W_t^2 (X_t - X_{t-1})' C (X_t - X_{t-1})\]

Given these transaction costs, \(W_{t+1}\) can be expressed (subtracting (2.8) from (2.1))

\[(2.9) \quad W_{t+1} = W_t (X_t' y_t + r_{yt}) - W_t \bar{b}' (X_t - X_{t-1})
+ \frac{1}{2} W_t^2 (X_t - X_{t-1})' C (X_t - X_{t-1})\]

The expected value of \(W_{t+1}\) is thus given by:

\[(2.10) \quad E [W_{t+1}] = W_t X_t' \mu_t + W_t r_{yt} - W_t \bar{b}' (X_t - X_{t-1})
- 0.5 W_t^2 (X_t - X_{t-1})' C (X_t - X_{t-1})\]

and the variance is given by

\[(2.11) \quad \text{Var} (\tilde{W}_{t+1}) = W_t^2 X_t' V X_t\]

The \(t^{th}\) period utility function is assumed to be given by \(^3\)

\[(2.12) \quad U (W_{t+1}) = E (W_{t+1}) - \frac{a}{2} \text{VAR} (\tilde{W}_{t+1})\]

where \(a\) is a risk aversion measure. In this case \(U(W_t)\) can be written as:

\[(2.13) \quad E [U (W_{t+1})] = W_t \left\{ X_t' \mu_t - \bar{b}' (X_t - X_{t-1})
+ \frac{1}{2} W_t (X_t - X_{t-1})' C (X_t - X_{t-1})
- \frac{a}{2} W_t X_t' V X_t + r_{yt} \right\}\]
Further rearrangement of terms yields:

$$
\mathbb{E} [U (W_{t+1})] = W_t \left\{ X_t' (\mu_t - b - W_t C X_{t-1}) - \frac{1}{2} W_t X_t' (aV - C) X_t \\
+ \left[ b' X_{t-1} + \frac{1}{2} W_t X_{t-1}' C X_{t-1} + r_{t-1} \right] \right\}
$$

(2.14)

The above equation is similar to the usual mean variance equation with the exceptions that $\mu_t$ is replaced by $(\mu_t - b - W_t C X_{t-1})$ and $aV$ is replaced by $(aV - C)$. Note that if costs were just linear, then one would just have to subtract the linear part from $\mu$ and the rest of the analysis would go through the same as in the usual portfolio selection problem.

In our case the optimal portfolio is obtained by differentiating (2.14) with respect to $X_t$ and equating the derivative to zero. One obtains:

$$
X_t' = \frac{1}{W_t} (aV_t - C)^{-1} [\mu_t - b - W_t C X_{t-1}]
$$

(2.15)

One notes from (2.15) that the optimal portfolio at $t$ depends on the portfolio held at $(t-1)$ through the transaction costs. The dependence runs through one of the following two channels: the element $W_t C X_{t-1}$ is introduced into the linear part of the returns, and $(aV)$ is replaced by $(aV - C)$ through the quadratic terms.

Given the above result we next consider how would transaction costs affect equilibrium returns. Given that there is just one investor and assuming that in equilibrium $X_t' = X^m$, where $X^m$ is the market portfolio, one obtains by substituting (2.15) for $X_t'$,

$$
X^m = \frac{1}{W_t} (aV_t - C)^{-1} (\mu_t - b - W_t C X_{t-1})
$$

(2.16)

Rearranging terms we obtain:
\begin{equation}
\mu_t = b + W_t \alpha X_{t-1} + W_t (a V_t - C)X^m
\end{equation}

We observe from (2.17) that the equilibrium returns are similar to those in the usual CAPM with the following differences: the linear part is simply added to $\mu$. This is what is expected as the investor requires in equilibrium that returns on risky asset compensate for these costs. In this model, however, also deviations from previous portfolio affect returns. If $X^m = X_{t-1}$ then there are no deviations and hence no effect of the non-linear costs. We therefore note that the dynamic nature of the equilibrium is due to nonlinearity of the transaction costs.

In the above analysis the equilibrium returns depend not only on correlation with the market and risk aversion but also on deviations from previous equilibrium $C(X^m - X_{t-1})$ where actually $X^m = X_t$. Since $X^m$ depends on $\mu_t$ and $\mu_t$ depends on $X_{t-1}$, auto-correlation results, and it seems quite important. This last result however strongly depends on the assumption of one representative investor. When there are many investors, the deviations $(X^m - X_{t-1})$ aggregate the changes in portfolio of many investors; some transactions may offset each other so that $(X^m - X_{t-1})$ may be small, even when there are many transactions.

In the following section we expand on this and other problems to generalize our model.

III. Equilibrium Returns with Many Types of Investors

In this section, it is assumed that there are many ($J$) types of investors. These investors may differ in their wealth, previous portfolio, risk aversion and also transaction costs. They do share, however, the same beliefs of $\mu$ and $V$. The distinction between
investors and the introduction of many investors serve the following purposes. First, it
overcomes the aggregation problem mentioned above. Second, it is often claimed that some
investors may be more efficient than others, transaction-costs wise. Some of the greater
efficiency may be due to economies of scale (which are captured by the quadratic part of the
transaction costs). However, some of the greater efficiency may be due to other reasons,
specialization, knowhow, or other. It is important therefore to consider the effect of different
transaction costs on the equilibrium returns.

There may also be differences in transaction costs in different securities. Some
securities will be cheaper to trade because of less asymmetry of information, or because of
greater liquidity, etc. In addition, there may be some interaction between investor specific
transaction costs, and from specific transaction costs. Most large firms are traded by large
investors. 4

The formulation of the previous section allows for different transaction costs for
different stocks (independent of size). In this section we introduce also transaction costs
specific to investors.

Technically, to allow for differences in transaction costs between investors, we only
have to add an index \( j \) to the representative investor of Section II. This index is added to the
parameters \( a \) (risk aversion), \( b \) and \( C \) (transaction costs), \( \text{W(wealth)} \) and \( X_{t-1} \) (previous
portfolio). The formula of the optimal portfolio is hence given by:

\[
(3.1) \quad X_t = \left( \frac{1}{W_t} \right) (a' V_t - C')^{-1} [\mu_t - b' - W_t X_{t-1}]
\]

The equilibrium is then defined as the \( \mu_t \) and \( V_t \) that equate aggregate supply to aggregate
demand. This condition can be written as:
\[ N_t^s = N_t^d = \sum_{j=1}^{j} W_t^j X_t^j \]
\[ = \sum_{j=1}^{j} (a^j V_t - C^j)^{-1} [\mu_t - b^j - W_t^j C^j X_t^{j-1}] \]
\[ = A_t^1 \mu_t - B_t + \xi_t, \]

where \( N^s \) and \( N^d \) denote the vector of number of shares demanded and supplied, respectively, and \( A_t, B_t \), and \( \xi_t \) are defined by:

\[ A_t = \sum_{j=1}^{j} (a^j V_t - C^j)^{-1} \]
\[ B_t = \sum_{j=1}^{j} (a^j V_t - C^j)^{-1} b^j \]
\[ \xi_t = \sum_{j=1}^{j} (a^j V_t - C^j)^{-1} W_t^j C^j X_t^{j-1} \]

Assuming that \( N^s \) is exogenously given we obtain:

\[ \mu_t = A_t^{-1} [N_t^s + B_t + \xi_t] \]

This formula is composed of the the element \( A_t^{-1} N_t^s \) which is the usual CAPM (except that \( aV-C \) is used instead of \( aV \)), \( A_t^{-1} B_t \) is the addition due to linear costs, and \( A_t^{-1} \xi_t \) is the addition due to nonlinearities in the transaction costs.
IV. Multiperiod Investment Decisions

In this section we consider the investment decisions of our investor in a multiperiod context. Here, the investor takes into account that the decisions at period $t$, will affect future transaction costs. It is assumed that the utility of the investor is additive over time, where each period's utility is given by (2.12). Given that the investor holds at the beginning of period $t$ the portfolio $X_{t-1}$ and that his/her wealth at that period is $W_t$ (subject to the simplifying assumption that $m_t = 1$), the multiperiod expected utility is given by:

\[
\phi_t(W_t, X_{t-1}) = \max_{X_t} W_t \left\{ X_t' \left( \mu_t' - b_t' - W_t C_t X_{t-1} \right) \right. \\
- \frac{1}{2} W_t X_t' \left( a_t V_t - C_t \right) X_t \\
\left. + \left[ b_t' X_{t-1} + \frac{1}{2} W_t X_{t-1}' C_t X_{t-1} + r_t \right] \right. \\
+ \rho \phi_{t+1}(W_{t+1}, X_t) \right\} \quad t = 1, \ldots, T .
\]

In (4.1) it is assumed that there is a $T$ period horizon (this implies that $\phi_{T+1} = 0$). We have also added an index $t$ to $b$ and $C$ so as to allow for variations in costs over time. There is no necessary relationship between $W_t$ and $W_{t+1}$ (as we mentioned earlier the desired investment at time $t$ may depend on exogenous variables).\(^5\)

The optimal decisions of the investor in this case are obtained by dynamic programming. It is shown below that the form of $\phi_t(W_t, X_{t-1})$ is quadratic in $X_{T-1}$, and then by induction that for all $t \leq T$, $\phi_t(W_t, X_{t-1})$ is quadratic in $X_{t-1}$. It is shown that
\( (4.2) \quad \phi(W_t, X_{t-1}) = W_t \max_{X_t} \left\{ X_t \left( \mu_t - b_t - W_t C_t X_{t-1} \right) \right. \\
- \frac{1}{2} W_t X_t \left( a_t V_t - C_t \right) X_t \\
+ \left[ b_t X_{t-1} + \frac{1}{2} W_t X_{t-1} C_t X_{t-1} + r_{ft} \right] \\
+ \rho \left[ X_t H_{t+1} X_t + h_{t+1} + g_{t+1} \right] \right\} 
\)

This implies that:

\( (4.3) \quad X_t^* = \left( 1/W_t \right) \left( a_t V_t - C_t + \rho H_{t+1} \right)^{-1} \left[ \mu_t - b_t - W_t C_t X_{t-1} + \rho h_{t+1} \right] 
\)

which is a formula similar to \( X_t^* \) in (2.14) except that \( \rho H_{t+1} \) is added to \( (a_t V_t - C_t)^{-1} \) and \( \rho h_{t+1} \) is added to \( (\mu_t - b_t - W_t C_t X_{t-1}) \). These changes reflect the effect of the dynamic considerations, i.e., how decision of \( X_{t-1} \) affects future decisions through their effects on future transaction costs.

The determination of \( h_t, H_t, \) and \( g_t \) is done recursively as follows. If \( t = T+1 \), then these variables vanish. To determine \( h_T, H_T, \) and \( g_T \), we note that when \( t = T \), then the portfolio problem is just a single period one. \( X_t^* \) is then given by (4.4) and it is well defined because \( H_{T+1} \) and \( h_{T+1} \) are known to be zero. Inserting (4.2) into (4.3) and rearranging terms one obtains:

\( (4.4) \quad \phi(W_t, X_{T-1}) = \frac{1}{2} \left( \mu_T - b_T - W_T C_T X_{T-1} + \rho h_{T+1} \right) \)
\( \left( a_T V_T - C_T + \rho H_{T+1} \right)^{-1} \left( \mu_T - b_T - W_T C_T X_{T-1} + \rho h_{T+1} \right) \)
\( + W_T \left[ b_t X_{T-1} + \frac{1}{2} W_T X_{T-1} C_T X_{T-1} + r_{ft} + \rho g_{T+1} \right] \)
which in turn is a quadratic function of \( X_{T-1} \). This becomes explicit after rewriting (4.4) as:

\[
(4.5) \quad \phi(W_t, X_{T-1}) = \frac{1}{2} W_t^2 X'_{T-1} \left[ C_T (a_T V_T - C_T + \rho H_{T+1})^{-1} C_T + C_T \right] X_{T-1} \\
+ W_t X'_{T-1} \left[ b_T C_T (a_T V_T - C_T + \rho H_{T+1})^{-1} (\mu_T - b_T + \rho h_{T+1}) \right] \\
+ W_t (r_{T+1} + \rho g_{T+1}) \\
+ \frac{1}{2} (\mu_T - b_T + \rho h_{T+1})' (a_T V_T - C_T + \rho H_{T+1})^{-1} (\mu_T - b_T + \rho h_{T+1})
\]

The functions \( \phi(W_t, X_{T-1}) \) can then be expressed for \( t < T \) as follows:

\[
(4.6) \quad \phi(W_t, X_{T-1}) = h_t' X_{T-1} + X'_{T-1} H_{T-1} + g_t,
\]

where

\[
h_t = W_t \left[ b_t + C_t (a_t V_t - C_t + \rho H_{t+1})^{-1} (\mu_t - b_t + \rho h_{t+1}) \right]
\]

\[
H_t = \frac{1}{2} W_t \left[ C_t (a_t V_t - C_t + \rho H_{t+1})^{-1} C_t + C_t \right]
\]

\[
g_t = W_t r_{T+1} + \frac{1}{2} (\mu_t - b_t + \rho h_{t+1})' (a_t V_t - C_t + \rho H_{t+1})^{-1} (\mu_t - b_t + \rho h_{t+1}) + \rho W_t \delta_{t+1}
\]

with boundary conditions:

\[
h_{T+1} = 0, H_{T+1} = 0, \text{ and } g_{T+1} = 0.
\]
Bibliography


Endnotes

1. The mean-variance approach is also more suitable to institutional investors and money managers, then the utility of consumption framework of Constantinides [ ] which applies better to small investors. Our approach is based on a similar use of the mean-variance approach, by Roll [1991] and Brennan [1993].

2. See [ ].

3. This is a simplified form of a quadratic utility function. It is equivalent to assuming that the investors minimizes variance subject to a given expected value. This assumption has been made extensively in the literature, see, e.g., Brennan [1993].

4. See, e.g., Schlaifer [ ].

5. This suits best the decisions of professional investors, managers of mutual funds, and brokerage firms.