TWO-STAGE DYNAMIC MULTI-PERIOD PORTFOLIO OPTIMIZATIONS

Shu-Ing Liu
Professor, Department of Finance, Shih Hsin University, No. 111, Mu-Cha Road, Sec 1, Taipei, Taiwan, 11604

ABSTRACT
This paper extends the traditional Markowitz’s mean-variance optimization to a two-stage dynamic multi-period portfolio optimization. The underlying assets time series data are supposed to follow a discrete time triangular cointegrated vector model, and in addition random quadratic transaction costs are taken into consideration. A two-stage dynamic multi-period approach is proposed, and the optimal solution under the discussed model is analytically derived. Also, comparisons between a standard one-stage static approach and some two-stage approaches will be numerically examined under constructed data. The results indicate that the proposed two-stage dynamic method performs quite efficiently, and that higher net returns per risk can be expected.

Keywords: Cointegration, Dynamic asset allocation, Markowitz optimization, Portfolio weight vector.

1. INTRODUCTION
The modern asset allocation theory was originated from the mean-variance portfolio model introduced by Markowitz [1] in 1952. The original Markowitz model simply dealt with a static single period asset allocation problem, calculating the tradeoff between risks and returns. A comprehensive theoretical framework was introduced, adapting the diversification concepts, that investors selected the ideal portfolio which gave the best return for a given risk. The proposals initiated the research foundation on current portfolio selection discussions. By applying the Bayesian analysis technique, Black and Litterman [2] introduced the concept that investor view has an input into the asset allocation problem. They believed that equilibrium exists in the financial market, thus the existing capitalization weights could serve as the basis for establishing an optimal allocation. In the investment industry, the model has significantly been applied by Goldman Sachs since 1990 though it is not so popular in academic research. Morton and Pliska [3] considered some correlated geometric Brownian motion stocks and a riskless bond as the underlying assets with numerical results showing that transaction costs were significantly influential in designing the optimal trading strategy. The impact of joining with transaction costs was discussed by Lobo, Fazel and Boyd [4], and Borkovec, Domowitz, Kiernan and Seibin [5], and the empirical results indicated that model with transaction costs lead to improvement in realized returns, and better alignment of return with risk. Gârleanu and Pedersen [6] analytically investigated an optimal dynamic portfolio policy with quadratic transaction costs and predictable returns under different mean-reversion speeds. Numerical results revealed that the proposed optimal dynamic strategy significantly performed better than the optimal static strategy.

An extension of the mean-variance formula in multi-period portfolio selection with analytical optimal solution was discussed by Li and Ng [7]. The explicit solution could provide investors with the optimal strategy to follow in a dynamic investment environment. A look back straddle approach was applied by Darius et al. [8] for evaluating the return characteristics of a trend-following strategy via a multi-period dynamic portfolio model. Empirical results showed that the advantages of the discussed strategy for investors were at the top end of the multi-period efficient frontier. Moreover, in order to obtain a more accurate estimation of the covariance matrix of underlying assets, the shrinkage technique popularly used in Bayesian analysis, was applied by Ledoit and Wolf [9]. The proposed shrinkage estimator could select portfolios with significantly lower out-of-sample variance demonstrated by some numerical examples.

It is popularly known that the cointegration approach, introduced by Granger [10], could represent the relationship of the long-term mobile trend between some non-stationary time series. Empirical results significantly showed that for most financial economic data, the log-prices of the underlying assets were non-stationary, and the cointegration structure helped to model non-stationary data without taking differencing transformation. A naive two-stage mean-variance portfolio selection approach, discussed by Rudoy and Rohrs [11], assumed that the underlying assets follow a discrete-time cointegrated vector autoregressive model. Some promising numerical results were obtained by comparing this with other trading strategies.

Extending Rudoy and Rohrs’ work [11] a two-stage multi-period portfolio selection with random portfolio weights was constructed by Liu [12]. The optimal solutions were derived in closed forms and equivalence comparisons between different formulations were shown analytically. In this paper, attention will be focused on the optimal mean return under a prescribed risk portfolio selection approach. A more efficient and informative two-stage multi-period
optimal solution, an extension of Liu’s work [12], will be analytically discussed and numerically investigated. This paper is organized as follows. In Section 2, the discussed discrete-time cointegrated vector autoregressive model is briefly reviewed. Under the proposed cointegrated model, multi-period portfolio selection formulation with quadratic transaction cost, and discussion of a traditional one-stage optimal portfolio weight, are included in Section 3. In Section 4 an innovative extension of Liu’s work [12] is proposed with a n-stage multi-period portfolio selection model that includes a random portfolio vector and the optimal portfolio allocation presented in closed form. Numerical comparisons between the one-stage and the two-stage approaches, one proposed by Liu [12] and the other discussed in this paper, under some special cases are demonstrated in Section 5. Finally, the conclusions are given in Section 6.

2. DISCRETE TIME SERIES MODEL FOR ASSET RETURNS

Granger [10] suggested that in a vector time series when all of the components are stationary after taking the first difference, there may exist stationary linear combinations. This caused the study of a financial cointegrated time series model which is a common approach to eliminate illogical correlation and still keep the long-term equilibrium between individual time series data. Suppose a capital market with \( k \) assets is considered and the investor will allocate his wealth among the \( k \) assets, and re-allocate at the beginning of each of the following \( n \) consecutive periods.

Let \( \mathbf{y}_t \) be a \( k \times 1 \) random vector time series, representing the log-prices of each asset. The portfolio selection problem for time series data \( \{ \mathbf{y}_t \} \) with model structure discussed by Liu’s [12] approach is adopted. A brief review is stated as follows: Suppose that each entry of \( \mathbf{y}_t \) has a unit root and a triangle cointegrated model with rank \( k_i \) is defined as follows:

\[
\mathbf{y}_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \delta_1 + \Pi^T y_{2,t-1} + a_{1t} \\ \delta_2 + y_{2,t-1} + a_{2t} \end{pmatrix},
\]

for \( t = 1, 2, \ldots, n \). Both \( y_{1t}, \delta_1 \) and \( a_{1t} \) are \( k_1 \times 1 \) vectors, with \( k_1 + k_2 = k \) and \( \Pi \) is a \( k_2 \times k_1 \) matrix; and the random vector, \( a_t^r = (a_{1t}^T, a_{2t}^T) \), is supposed to satisfy a \( \text{VAR}(p) \) model. After algebraic computations, the \( a \)'s could be assumed to be multivariate normally distributed, with mean 0 and some variance-covariance matrix, \( \Sigma_1 \), denoted by \( \text{MN}(0, \Sigma_1) \). The elements of \( \Sigma_1 \) are function of parameters involved in the \( \text{VAR}(p) \) model.

Denote \( \mathbf{I}_j \) as a \( j \times j \) identity matrix, \( \mathbf{0}_j \) as a \( j \times j \) zero matrix, \( \mathbf{0}_s \) as a \( k_2 \times k_1 \) zero matrix, \( \mathbf{P}_0 = \begin{pmatrix} \mathbf{I}_k & \Pi^T \\ \mathbf{0}_s & \mathbf{I}_{k_1} \end{pmatrix} \),

\[
\mathbf{P}_1 = \begin{pmatrix} \mathbf{0}_k & \Pi^T \\ \mathbf{0}_s & \mathbf{I}_{k_1} \end{pmatrix}, \quad \delta = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}, \quad \beta = \mathbf{P}_0 \delta, \quad \text{and } \epsilon_t = \mathbf{P}_0 \alpha_t, \] then model (1) is re-written as:

\[
\mathbf{y}_t = \beta + \mathbf{P}_1 \mathbf{y}_{t-1} + \epsilon_t.
\]

Here \( \begin{pmatrix} \epsilon_1^T, \epsilon_2^T, \ldots, \epsilon_n^T \end{pmatrix} \sim \text{MN}(0, \Sigma_{(n)}) \), with \( \Sigma_{(n)} = (\mathbf{I}_n \otimes \mathbf{P}_0)^T \Sigma_1 (\mathbf{I}_n \otimes \mathbf{P}_0^T) \), and \( \otimes \) denotes the Kronecker product. For detailed derivations please refer to Liu [12].

Let \( x_0 \) denote the investor initial wealth, and \( \mathbf{y}_0 \) is the natural logarithm of \( x_0 \); and \( \mathbf{P}_2 = \mathbf{P}_1 - \mathbf{I}_k \) and \( \epsilon_t = \mathbf{P}_2 \epsilon_{t-1} + \epsilon_t \), then for \( t \geq 2 \)

\[
r_t = \mathbf{y}_t - \mathbf{y}_{t-1} = \beta + \mathbf{P}_2 \mathbf{y}_{t-1} + \epsilon_t = \beta + \mathbf{P}_2 \left( \beta + \mathbf{P}_1 \mathbf{y}_{t-2} + \epsilon_{t-1} \right) + \epsilon_t = \mathbf{P}_1 \beta + \mathbf{P}_2 \epsilon_{t-1} + \epsilon_t = \mathbf{P}_1 \delta + \epsilon_t.
\]

Therefore,

\[
r_t = \mathbf{y}_t - \mathbf{y}_{t-1} = \begin{cases} \beta + \mathbf{P}_2 \mathbf{y}_0 + \epsilon_1, & \text{for } t = 1, \\ \mathbf{P}_1 \delta + \epsilon_t, & \text{for } t \geq 2. \end{cases}
\]

Moreover, define \( \mathbf{r}_{(n)} = \begin{pmatrix} r_1^T, r_2^T, \ldots, r_n^T \end{pmatrix} \) and \( \mathbf{e}_{(n)} = \begin{pmatrix} \epsilon_1^T, \epsilon_2^T, \epsilon_3^T, \ldots, \epsilon_n^T \end{pmatrix} \), then given an initial log-asset value, \( \mathbf{y}_0 \), the overall \( n \)-period joint return resulted as:

\[
r_{(n)} = \alpha_{(n)} + \mathbf{e}_{(n)},
\]
where 
\[ \alpha_i = E(r_{i(a)} \mid y_0) = (\alpha_i^T, \alpha_i^T, \cdots, \alpha_i^T) \],
\[ \alpha_i = \left\{ \begin{array}{ll} \beta + P_i y_0, & i = 1, \\ P_i \delta, & i \geq 2, \end{array} \right. \]
\[ \text{Var}(r_{i(a)} \mid y_0) = \Phi_{i(a)} = P_{i(a)} \Sigma_{i(a)} P_{i(a)}^T, \]
\[ \text{With } P_{i(a)} = \begin{pmatrix} I_k & 0 & 0 & 0 \\ 0 & P_2 & I_k & 0 \\ 0 & 0 & \ddots & \ddots \\ 0 & 0 & \cdots & P_2 \end{pmatrix}. \] Also, \( \varepsilon_{i(a)} \sim MN(0, \Phi_{i(a)}) \) and as \( \delta_2 = 0 \), this implies that \( P_i \delta = 0 \).

3. MULTI-PERIOD MEAN-VARIANCE FORMULATIONS

3.1 The framework

The investor plans to allocate his wealth among some \( k \) assets by deciding the portfolio weights at time \( t \) denoted by \( w_t \). Positive entries of \( w_t \) denote purchasing positions, while negative entries of \( w_t \) denote shorting positions. In the following discussion no restrictions on the sign of entries of \( w_t \) are attached. For convenience, when asset prices are mentioned, they mean log-asset prices, instead of the actual asset prices. At the beginning time, suppose the investor has initial asset value \( y_0 \), which is known in advance. The goal is to design an optimal \( n \)-period asset allocation by deciding portfolio weights denoted by \( w_{(a)} = (w_t^1, w_t^2, \cdots, w_t^n)^T \).

Let the one-period transaction costs associated with trading \( w_i - w_{i-1} \) shares, may change the price by the amount, say, \( \frac{1}{2} \Theta_i (w_i - w_{i-1}) \). The coefficient, one-half, is set for computational convenience, and \( \Theta_i \) is a symmetric matrix, measuring the level of trading costs. Then, the one-period convex transaction costs are defined as \( \frac{1}{2} \Theta_i (w_i - w_{i-1})^T (w_i - w_{i-1}) \), and the entire \( n \)-period transaction costs are thus
\[ \frac{1}{2} \sum_{i=1}^{n} (w_i - w_{i-1})^T \Theta_i (w_i - w_{i-1}) = \frac{1}{2} w_{(a)}^T A_{(a)} w_{(a)}. \]

For convenience, \( w_0 \) is defined as 0, and \( A_{(a)} \) is a \( nk \times nk \) symmetric positive-definite matrix expressed as:
\[ A_{(a)} = \begin{pmatrix} \Theta_1 + \Theta_2 & -\Theta_2 & \cdots & -\Theta_2 \\ -\Theta_2 & \Theta_2 + \Theta_3 & \cdots & -\Theta_3 \\ \vdots & \vdots & \ddots & \vdots \\ -\Theta_{n-1} & -\Theta_{n-2} & \cdots & \Theta_n \end{pmatrix}. \]

In this paper the discussed mean-variance algorithm will be focused on the one that maximizes the overall \( n \)-period net expected returns, under a prescribed overall risk level, under the given initial asset value \( y_0 \). That is
\[ \max_w \left[ E(w_{(a)}^T r_{(a)} \mid y_0) - \frac{1}{2} E(w_{(a)}^T A_{(a)} w_{(a)} \mid y_0) \right], \text{ subject to } \text{Var}(w_{(a)}^T r_{(a)} \mid y_0) = \sigma_0^2. \]

The future portfolio weights will be formulated at the beginning time: Two approaches will be considered, depending upon the structure of the portfolio weight vector, \( w_{(a)} \): One, treated \( w_{(a)} \) as deterministic, is called a standard one-stage method hereafter. The future portfolio weights are explicitly determined at the beginning time; the other is designed as a random vector constructed by a two-stage procedure, called a two-stage method. The future portfolio weights conditionally determined at the beginning time, are updated period by period. These two
approaches will be summarized in the following subsection.

3.2 Standard One-Stage Multi-period Approach

Firstly, we discuss the case where the portfolio weight \( w_{(n)} \) is deterministic, a standard one-stage method: By using the Lagrange multiplier method, the optimal weight is straightly obtained from maximizing the following objective function:

\[
J_s(w_{(n)} | y_0) = E(w_{(n)}^{T} r_{(n)} | y_0) - \frac{1}{2} w_{(n)}^{T} \Lambda_{(n)} w_{(n)} - \frac{\lambda_s}{2} \left[ \text{Var}(w_{(n)}^{T} r_{(n)} | y_0) - \sigma_0^2 \right].
\]

Here \( \lambda_s \) is the Lagrange multiplier. Substituting the model structure stated in Section 2, into the above mentioned formula, the objective function is re-written as:

\[
J_s(w_{(n)} | y_0) = w_{(n)}^{T} \alpha_{(n)} - \frac{1}{2} w_{(n)}^{T} (\Lambda_{(n)} + \lambda_s \Phi_{(n)}) w_{(n)} + \frac{\lambda_s}{2} \sigma_0^2.
\]

Thus the optimal allocations under the standard one-stage algorithm, will be obtained by solving the equation, \( \alpha_{(n)} - (\Lambda_{(n)} + \lambda_s \Phi_{(n)}) \hat{w}_{(n)} = 0 \). The optimal solution resulted

\[
\hat{w}_{(n)} = \Psi_{(n)} \alpha_{(n)}, \quad \text{where} \quad \Psi_{(n)} = (\Lambda_{(n)} + \lambda_s \Phi_{(n)}).
\]

The optimal Lagrange multiplier \( \lambda_s \) can be solved from the following equation:

\[
\alpha_{(n)}^{T} \left( \Lambda_{(n)} + \lambda_s \Phi_{(n)} \right)^{-1} \Phi_{(n)} \left( \Lambda_{(n)} + \lambda_s \Phi_{(n)} \right)^{-1} \alpha_{(n)} = \sigma_0^2. \quad (6)
\]

Then the optimal returns under the constant weight approach is

\[
J_s(\hat{w}_{(n)} | y_0) = \frac{1}{2} \alpha_{(n)}^{T} \left( \Lambda_{(n)} + \lambda_s \Phi_{(n)} \right)^{-1} \alpha_{(n)} + \frac{\lambda_s}{2} \sigma_0^2.
\]

And the overall net expected return per risk is defined as

\[
\text{Ratio}_{n,s} = \frac{E(\hat{w}_{(n)}^{T} r_{(n)} | y_0) - \frac{1}{2} \hat{w}_{(n)}^{T} \Lambda_{(n)} \hat{w}_{(n)}}{\sqrt{\text{Var}(\hat{w}_{(n)}^{T} r_{(n)} | y_0)}} = \frac{\alpha_{(n)}^{T} \left( \Lambda_{(n)} + \lambda_s \Phi_{(n)} \right)^{-1} \alpha_{(n)} + \frac{\lambda_s}{2} \sigma_0^2}{\sigma_0}.
\]

4. TWO-STAGE DYNAMIC MULTI-PERIOD APPROACH

In this section, an innovative optimal portfolio problem, with a random portfolio vector and random transaction costs are investigated. At each time epoch \( t \), the current portfolio weight is determined conditional on the information up to a given time epoch \( t-1 \): At the first stage, starting from the beginning time epoch, a sequence of optimal portfolio weight vector is period by period consecutively constructed. Following this an overall optimal portfolio weight vector will be suitably combined together at the second stage to constitute a random portfolio weights vector. This approach is denoted as a two-stage method. A detailed construction is discussed as follows:

4.1 Background of the Proposed Two-stage Approach

From model (3), the \( n \)-period returns \( r_{(n)} \) follow a multivariate normal distribution, say, \( r_{(n)} \sim MN(\alpha_{(n)}, \Phi_{(n)}) \), where \( \alpha_{(n)} \) are \( \Phi_{(n)} \) defined by equation (4) and (5) respectively. In this subsection, instead of treating the portfolio weight \( w_{(n)} \) as a deterministic vector, it is regarded as a random vector. By the conditional property of a multivariate normal distribution, it turns out that, for \( i \geq 1 \),

\[
r_{(i)} | r_{(i-1)}, y_0 \sim MN(\xi_{i,i-1}, \Psi_{i,i-1}), \quad \text{and} \quad w_{(i)}^{T} r_{(i)} | r_{(i)}, y_0 \sim MN(w_{(i)}^{T} \xi_{i,i}, w_{(i)}^{T} \Psi_{i,i} w_{(i)}), \quad \text{where} \quad \xi_{i,i} = \alpha_{(i)} + \Phi_{i,i} \xi_{i-1,i} \Phi_{i,i}^{-1}, \quad \Psi_{i,i} = \Phi_{i-1,i} \Psi_{i-1,i} \Phi_{i-1,i}^{-1} - \Phi_{i,i} \Phi_{i-1,i} \Phi_{i-1,i}^{-1}, \quad \Phi_{i,j} = \text{Cov}(r_{(i)}, r_{(j)}), \quad \text{and} \quad \Phi_{0,j} = \text{Cov}(r_{(1)}, r_{(j)}).\]

And the initial conditional distribution is:

\[
r_{(1)} | y_0 \sim MN(\xi_{1,0}, \Psi_{1,0}), \quad \text{and} \quad w_{(1)}^{T} r_{(1)} \sim MN(w_{(1)}^{T} \xi_{1,0}, w_{(1)}^{T} \Psi_{1,0} w_{(1)}), \quad \text{where} \quad \xi_{1,0} = \alpha_{1} \quad \text{and} \quad \Psi_{1,0} = \Sigma_{(1)}.
\]

At the first stage, the current portfolio weight is constructed depending upon returns of all the preceding periods.
Temporarily define the objective function at the \((i+1)\)-th period, for giving the previous returns \(r_{(i)}\), as

\[
h_{i+1}(w_{i+1} \mid r_{(i)}, y_0) = E(w_{i+1}^T r_{i+1} \mid r_{(i)}, y_0) - \frac{1}{2} \text{Var}(w_{i+1}^T r_{i+1} \mid r_{(i)}, y_0) = w_{i+1}^T \xi_{i+1} - \frac{1}{2} w_{i+1}^T \Psi_{i+1} w_{i+1}, \quad \text{for } i \geq 1.
\]

Then the preliminary optimal solution of \(w_{i+1}\) is

\[
\widetilde{w}_{i+1} = \Psi_{i+1}^{-1} \xi_{i+1} = a_{i+1} + A_{i+1}^T (r_{(i)} - \alpha_{(i)}), \quad \text{for } i = 1, 2, \ldots, n - 1,
\]

where \(a_{i+1} = \Psi_{i+1}^{-1} \alpha_{i+1}, A_{i+1} = \Phi_{i+1}^{-1} \Phi_{i+1} \Psi_{i+1}^{-1}, \) and \(\Psi_{i+1} = \Phi_{i+1} \Psi_{i+1} - \Phi_{i+1} \Phi_{i+1} \Phi_{i+1}^{-1} \Phi_{i+1} \Psi_{i+1}^{-1}.\) It is worthy to note that before \(r_{(i)}\) is observed, \(\widetilde{w}_{i+1}\) is random; while after \(r_{(i)}\) is observed, \(\widetilde{w}_{i+1}\) is deterministic.

After the portfolio weight vectors for each period are temporarily obtained, denoted by \(\hat{w}_2, \hat{w}_3, \ldots, \hat{w}_n\), the final portfolio weight vector for the whole future period are constructed by a rolling scheme: Starting from the beginning period, the portfolio weight denoted by \(\hat{w}_1\) is assumed to be deterministic; Then the portfolio weights for the remaining periods are \(\rho_1 \hat{w}_2, \rho_1 \hat{w}_3, \ldots, \rho_1 \hat{w}_n\), consecutively. Thus at the beginning time, \(\hat{w}_1\)'s, for \(i \geq 2\), are random, instead of deterministic. Now, according to the strategy, innovative optimal portfolio weights are obtained by finding suitable \(\hat{w}_1, \rho_1\), which maximize the following objective function:

\[
E(w_{i+1}^T r_1 + \rho_2 \hat{w}_2^T r_2 + \cdots + \rho_n \hat{w}_n^T r_n \mid y_0) - \frac{1}{2} \sum_{i=1}^{n} E[(\rho_i \hat{w}_i - \rho_{i+1} \hat{w}_{i+1})^T \Theta (\rho_i \hat{w}_i - \rho_{i+1} \hat{w}_{i+1}) \mid y_0]
\]

\[-\frac{\lambda}{2} \text{Var}(w_{i+1}^T r_1 + \rho_2 \hat{w}_2^T r_2 + \cdots + \rho_n \hat{w}_n^T r_n \mid y_0) - \sigma_0^2 \],

where \(\hat{w}_i = w_i, \hat{w}_0 = 0, \rho_1 = 1, \rho_0 = 0\) and \(\lambda\) is the Lagrange multiplier.

In order to obtain the optimal solution analytically, it is convenient to re-represent the objective function. Define

\[
w_{(n),b} = \begin{pmatrix} w_1 \\ \rho_2 \hat{w}_2 \\ \vdots \\ \rho_n \hat{w}_n \end{pmatrix}, \quad u_{(n)} = \begin{pmatrix} r_1 \\ \hat{w}_2^T r_2 \\ \vdots \\ \hat{w}_n^T r_n \end{pmatrix}, \quad \gamma_{(n)} = \begin{pmatrix} 1_k \\ 0 \\ \vdots \\ \hat{w}_n \end{pmatrix}, \quad \text{and } U_{(n)} = \begin{pmatrix} I_k \\ \hat{w}_2 \\ \vdots \\ \hat{w}_n \end{pmatrix},
\]

then \(w_{(n),b} = U_{(n)} u_{(n)}, \ w_{(n),b}^T r_{(n)} = u_{(n)}^T \gamma_{(n)}, \ w_{(n),b}^T A_{(n)} w_{(n),b} = U_{(n)}^T U_{(n)} A_{(n)} U_{(n)} u_{(n)}.\) Let \(\xi_{(n)} = E(\gamma_{(n)} \mid y_0).\)

\[
\Omega_{(n)} = \text{Var}(\gamma_{(n)} \mid y_0) \quad \text{and } C_{(n)} = E(U_{(n)}^T A_{(n)} U_{(n)} \mid y_0); \text{ in detail:}
\]

\[
\xi_{(n)} = E(\gamma_{(n)} \mid y_0) = \begin{bmatrix} E(r_1 \mid y_0)^T, \ E(\hat{w}_2^T r_2 \mid y_0), \cdots, E(\hat{w}_n^T r_n \mid y_0) \end{bmatrix}^T,
\]

\[
\Omega_{(n)} = \begin{bmatrix} \text{Cov}(r_1, r_1 \mid y_0) & \text{Cov}(r_1, \hat{w}_2^T r_2 \mid y_0) & \cdots & \text{Cov}(r_1, \hat{w}_n^T r_n \mid y_0) \\ \text{Cov}(\hat{w}_2^T r_2, r_1 \mid y_0) & \text{Cov}(\hat{w}_2^T r_2, \hat{w}_2^T r_2 \mid y_0) & \cdots & \text{Cov}(\hat{w}_2^T r_2, \hat{w}_n^T r_n \mid y_0) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(\hat{w}_n^T r_n, r_1 \mid y_0) & \text{Cov}(\hat{w}_n^T r_n, \hat{w}_2^T r_2 \mid y_0) & \cdots & \text{Cov}(\hat{w}_n^T r_n, \hat{w}_n^T r_n \mid y_0) \end{bmatrix},
\]

and
4.2 Two-stage Optimal Portfolio Solution

The objective function is re-written by notations stated in subsection 4.1, as:

\[ h_b(u_n) | y_0 = E(u_n^T y_n | y_0) - \frac{1}{2} E(u_n^T U_n^T \lambda_n^T U_n u_n | y_0) - \frac{\lambda_k}{2} [Var(u_n^T y_n | y_0) - \sigma_0^2] \]

\[ = u_n^T \tilde{\xi}_n - \frac{1}{2} u_n^T \Omega_n u_n - \frac{\lambda_k}{2} (u_n^T \Omega_n u_n - \sigma_0^2). \]

The optimal solution is \( \tilde{u}_n = (C_n + \tilde{\lambda}_n \Omega_n)^{-1} \tilde{\xi}_n \), where \( \tilde{\lambda}_n \) is determined by the equation

\[ \tilde{\xi}_n (C_n + \tilde{\lambda}_n \Omega_n)^{-1} \Omega_n (C_n + \tilde{\lambda}_n \Omega_n)^{-1} \tilde{\xi}_n = \sigma_0^2. \]  

Finally, the optimal return under the two-stage approach is summarized as:

\[ h_b(\tilde{u}_n) | y_0 = \frac{1}{2} \tilde{\xi}_n^T (C_n + \tilde{\lambda}_n \Omega_n)^{-1} \tilde{\xi}_n + \frac{\lambda_k}{2} \sigma_0^2. \]

And the overall net expected return per risk is defined as:

\[ Ratio_{n,b} = \frac{E(\tilde{u}_n^T y_n | y_0) - \frac{1}{2} E(\tilde{u}_n^T U_n^T \lambda_n^T U_n \tilde{u}_n | y_0)}{\sqrt{Var(\tilde{u}_n^T y_n | y_0)}} \]

\[ = \frac{\tilde{\xi}_n^T (C_n + \tilde{\lambda}_n \Omega_n)^{-1} \left( \frac{1}{2} C_n + \tilde{\lambda}_n \Omega_n \right)^{-1} \left( C_n + \tilde{\lambda}_n \Omega_n \right)^{-1} \tilde{\xi}_n}{\sigma_0}. \]

As the optimal \( \tilde{u}_n \) is determined, the optimal portfolio weight vector \( \tilde{W}_{(n)_*} \) is obtained, say \( \tilde{w}_{(n)_b} = U_n \tilde{u}_n = \tilde{w}_1 \tilde{p}_1 \tilde{w}_2 \tilde{p}_2 \tilde{w}_3 \cdots \tilde{p}_r \tilde{w}_r \). The two-stage approach decides the portfolio weight period by period: Starting from the beginning period the optimal portfolio weight vector \( \tilde{w}_1 \) is first decided, then once the first period return \( r_1 \) is observed, the optimal weight vector for the second period is determined by \( \tilde{p}_2 \tilde{w}_2 \).

Analogically, the last optimal weight vector is decided by \( \tilde{p}_r \tilde{w}_r \), which depends upon all the previous \((n-1)\) returns \( r_{(n-1)} \). Finally, an adaptive approach of the optimal portfolio weight vector is thus consecutively established. Before \( r_{(i)} \) is observed, \( \tilde{w}_{i+1} \) is random; whereas after \( r_{(i)} \) is observed, \( \tilde{w}_{i+1} \) is deterministic. Therefore, at the beginning time, the portfolio weights for each future period are random, except for the first period weights.

It is worthy to note that the discussed dynamic weight, \( \tilde{w}_{i+1} \) depending upon the initial wealth \( y_0 \) and all the previous returns \( r_j \), for \( j \leq i \), instead of only the recently observed return \( r_i \), discussed by Liu [12]. To complete the proposed allocation strategy, values of the Lagrange multiplier coefficients, \( \tilde{\lambda}_n \), and \( \tilde{\xi}_n \), \( \Omega_n \) and \( C_n \) should be explicitly expressed. The computing procedures are similar to those stated in the Appendix of Liu [12]. Some results are sketched in the Appendix A.

**Lemma 1.** Computation of \( \tilde{\xi}_n \):

\[ E(\tilde{w}_{i+1} r_s | y_0) = A_{(i+1)} \alpha_i + tr(A_{(i+1)} \phi_{(i+1)(i)}), \text{ for } i = 1, 2, \cdots, n - 1. \]
**Lemma 2.** Computation of $\Omega_{(a)}$:

Embedding matrix $A_j$ into a suitable $nk \times nk$ matrix, say $A^*_j$, for $j \geq 2$, where $A^*_j = \tilde{A}^*_j + (\tilde{A}^*_j)^T$, and

$$(\tilde{A}^*_j)_{s,l} = \begin{cases} (A^*_j)_{s,l-(j-1)k} & \text{if } 1 \leq s \leq (j-1)k \& (j-1)k+1 \leq l \leq jk, \\ 0, & \text{otherwise}. \end{cases}$$

The re-constructed matrix $A^*_j$ is obviously a symmetric matrix, and satisfying the following properties:

1. $(\mathbf{r}_{(j-1)} - \mathbf{\alpha}_{(j-1)})^T A_j (\mathbf{r}_j - \mathbf{\alpha}_j) = \frac{1}{2} (r - \mathbf{\alpha})^T A^*_j (r - \mathbf{\alpha}).$

2. $\text{Cov}(\tilde{w}^T r, \tilde{w}^T r_j | y_0) = a^T_i \Phi_{i,i} a_j + a^T_j \Phi_{j,j} a_j + a^T_j \Phi_{j,j} a_j + a^T_i \Phi_{i,i} A^*_j A_j.$

3. $\text{Cov}(\tilde{w}^T r_j | y_0) = \text{Cov}(\tilde{w}^T r_j | y_0) = \text{Cov}(\alpha_i, \alpha_j).$

4. $\text{Cov}(\tilde{w}^T r_j | y_0) = (\Phi_{i,i} + \Phi_{j,j} \Phi_{j,j}) A_i \alpha_j.$

**Lemma 3.** Computation of $C_{(a)}$:

Embedding matrix $A \Theta A^*_j$ into a suitable $nk \times nk$ matrix, say $C^*_j$, for $i \geq 2$ and $J \geq 2$, where

$C^*_j = (\tilde{c}^*_j + (\tilde{c}^*_j)^T$, and

$$(\tilde{c}^*_j)_{s,l} = \begin{cases} (A \Theta A^*_j)_{s,l-(j-1)k} & \text{if } 1 \leq s \leq (i-1)k \& (i-1)k+1 \leq l \leq (j-1)k, \\ 0, & \text{otherwise}. \end{cases}$$

Similarly, the re-constructed matrix $C^*_j$ is a symmetric matrix, and satisfying the following properties:

1. $(\mathbf{r}_{(i-1)} - \mathbf{\alpha}_{(i-1)})^T A \Theta A^*_j (\mathbf{r}_{(j-1)} - \mathbf{\alpha}_{(j-1)}) = \frac{1}{2} (r - \mathbf{\alpha})^T C^*_j (r - \mathbf{\alpha}).$

2. $E(\tilde{w}_2 | y_0) = a_2.$

3. $E(\tilde{w}^T \Theta \tilde{w}_j | y_0) = a^T_i \Theta a_j + \frac{1}{2} \text{tr}(C^*_j \Phi_{(a)}).$

The one-stage approach deterministically decides the $n$-period optimal portfolio weights at the beginning; while the two-stage approach constructs the portfolio weight vector by period: At the beginning of the $(i + 1)$-th period, before previous information $r_{(i)}$ was completely observed, $\tilde{w}_{i+1}$ is random; whereas after $r_{(i)}$ was observed, then $\tilde{w}_{i+1}$ is deterministic. Comparisons between the two approaches, the standard one-stage and the proposed two-stage will be discussed numerically in the next section. Before this is done however, one special case can first be analytically investigated below.

**Lemma 4:** Suppose in model (1), $y_0 = 0$, $\delta = 0$, and without transaction costs, are assumed, then the expected return to risk ratio for the proposed two-stage, is always greater than that under the one-stage approach.

Proof: Since $y_0 = 0$ and $\delta = 0$, this implies that $\mathbf{\alpha}_{(a)} = 0$; moreover, by no transaction costs assumption, the two objective functions are reduced as follows:

$$f_1(\tilde{w}_{(a)} | y_0) = \frac{1}{2} \tilde{w}_{(a)}^T (\tilde{\lambda}_2 \mathbf{\Omega}_{(a)}) \tilde{w}_{(a)} + \frac{1}{2} \tilde{\sigma}_{0}^2 = \tilde{\lambda}_2 \tilde{\sigma}_{0}^2 = \sqrt{\tilde{\sigma}_{(a)}^T \mathbf{\Omega}_{(a)} \tilde{\sigma}_{(a)}} \tilde{\sigma}_{0} = 0$$

and

$$h_1(\tilde{w}_{(a)} | y_0) = \frac{1}{2} \tilde{w}_{(a)}^T (\tilde{\lambda}_2 \mathbf{\Omega}_{(a)}) \tilde{w}_{(a)} + \frac{1}{2} \tilde{\sigma}_{0}^2 = \tilde{\lambda}_2 \tilde{\sigma}_{0}^2 = \sqrt{\tilde{w}_{(a)}^T \mathbf{\Omega}_{(a)} \tilde{w}_{(a)}} \tilde{\sigma}_{0} = 0 > 0$$

59
5. NUMERICAL ILLUSTRATIONS

In this section, performing comparisons in terms of the overall net expected return per risk, among three approaches the one-stage standard method and two two-stage dynamic methods, are numerically demonstrated: One of the two two-stage methods is the one discussed by Liu [12] and the other is introduced in Section 4. In the following numerical investigations, the former is referred as the two-stage method A, and the latter denoted as the two-stage method B. A brief sketch of the development of the two-stage method A is given in Appendix B. The major differences between the two-stage methods rely on the preliminary estimations of portfolio weight, \( w_{i+1} \), at the \((i+1)\)-stage; the defining \( w_{i+1} \) under method A is conditioned only on information \( r_i \), while that under method B is conditioned on information \( r_{i(1)} \), extracting more information.

In detail, the preliminary \((i+1)\)-th period objective function for the two-stage method A giving the previous returns \( r_i \) is defined as

\[
g_{i+1}(w_{i+1} | r_i, y_0) = E(w_{i+1}^T r_{i+1} | r_i, y_0) - \frac{1}{2} \text{Var}(w_{i+1}^T r_{i+1} | r_i, y_0). \tag{8}
\]

While the preliminary estimate of \( w_i \) for the two-stage method A is to maximize the stated objective function (8). Since \( r_{i(1)} = (r_i^T, r_{i+1}^T, \ldots, r_n^T)^T \), \( r_i \) is an element of \( r_{i(1)} \), it seems that method B contains more information than method A; therefore we may expect that method B should be more efficient than method A.

In order to provide the scenario of a real financial market, the required parameters were set up by using historical market data collected from the Taiwan Economic Journal Data Bank. Data were screened from the 19 Taiwan Industrial Index, with only five of them included, denoted as \( k=5 \) in the following discussion: The five discussed industrial indexes are, Building Construction, Financial and Insurance, Steel and Iron, Electronic and Electrical, and Biotech.

5.1 Parameters Setup

To establish a suitable cointegration model like model (1), we data vector as the logarithm of the five Taiwan industry index. \( k_1 = 2 \) and \( k_2 = 3 \), and values of parameter are set as follows:

\[
y_0 = (1.2,1.5,0.8,0.5,0.9)^T, \quad \delta_1 = \begin{pmatrix} 0.62 \\ 1.50 \end{pmatrix}, \quad \delta_2 = \begin{pmatrix} 0.01 \\ 0.05 \\ 0.02 \end{pmatrix}, \quad \text{and } \Pi = \begin{pmatrix} 0.03 & -0.34 \\ -0.52 & -0.63 \\ 1.37 & 1.64 \end{pmatrix}.
\]

Furthermore, after numerical algebra, parameters in model (2) are obtained as:

\[
\beta = \begin{pmatrix} 0.622 \\ 1.498 \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 0 & 0.34 & -0.52 & 1.37 \\ 0 & 0 & 0.63 & 1.64 \end{pmatrix}, \quad P_{(n)} = \begin{pmatrix} I_5 & 0 & 0 \\ P_2 & I_5 & 0 \\ 0 & P_2 & I_5 \end{pmatrix}, \quad \text{and } P_2 = P_1 - I_5,
\]

For simplicity, \( \Sigma = I_n \otimes \Sigma \) is assumed for some symmetric positive definite matrix \( \Sigma \). Since \( \Sigma_{(n)} = (I_n \otimes P_0) \Sigma (I_n \otimes P_0)^T \), therefore \( \Sigma_{(n)} = I_n \otimes W^* \). Here \( W^* \) is a \( 5 \times 5 \) symmetric positive definite matrix defined as:

\[
W^* = \begin{pmatrix} 0.25 & 0.86 & 0.20 & 0.27 & 0.38 \\ 0.86 & 0.82 & 0.24 & 0.32 & 0.48 \\ 0.20 & 0.24 & 0.30 & 0.18 & 0.20 \\ 0.27 & 0.32 & 0.18 & 0.40 & 0.30 \\ 0.38 & 0.48 & 0.20 & 0.30 & 0.48 \end{pmatrix}.
\]

After \( \Sigma_{(n)} \) is defined, then \( \Phi_{(n)} \) is obtained from the relationship, \( \Phi_{(n)} = P_{(n)} \Sigma_{(n)} P_{(n)}^T \). In order to investigate the
impact of the covariance matrix $\Sigma_{(n)}$, different versions of $W^*$ will be introduced. In the following discussions, for simplicity, we set
\[ \Sigma_{(n)} = I_n \otimes W, \] where $W = W^* + \theta \times I_k$.

The impacts of different values of “$\theta$” will be numerically investigated.

The transaction cost in the Taiwanese stock market, includes a transaction fee of 0.1425% for buying, selling and short selling stocks, a 0.3% of selling trading tax, and also a 0.08% for short selling fee. For simplicity, a fixed ratio is used as a transaction cost for buying, selling and short selling, expressed as,
\[ \text{Transaction fee} + (\text{securities transaction tax} + \text{borrowing cost}) / 2 = 0.3325\% \]

For simplicity, the transaction matrix $\Theta_{(i)}$ is set to be equal, say $\Theta_{(i)} = \Theta$, where $\Theta$ is a diagonal matrix with the diagonal entry taking value 0.3325%. Thus, the overall transaction cost matrix $\Lambda_{(n)}$ is simplified as below:
\[
\Lambda_{(n)} = \begin{pmatrix}
2\Theta & -\Theta \\
-\Theta & 2\Theta & -\Theta \\
& \ddots & \ddots \\
0 & -\Theta & 2\Theta & -\Theta \\
-\Theta & \Theta & \Theta & \Theta
\end{pmatrix}.
\]

### 5.2 Numerical Results

Comparisons are focused on the optimal overall net expected return per risk, for convenience, key expression for each method is summarized as follows: For the standard one-stage method, the net expected return per risk is:
\[
Ratio_{n,S} = \frac{\alpha_n^T \left( \Lambda_{(n)} + \hat{\theta}_2 \Phi_{(n)} \right)^{1/2} \left( \hat{\theta}_2 \Phi_{(n)} + \frac{1}{2} \Lambda_{(n)} \right) \left( \Lambda_{(n)} + \hat{\theta}_2 \Phi_{(n)} \right)^{1/2} \alpha_n}{\sigma_0},
\]

where $\hat{\theta}_2$ is determined by $
\alpha_n^T \left( \Lambda_{(n)} + \hat{\theta}_2 \Phi_{(n)} \right)^{1/2} \Phi_{(n)} \left( \Lambda_{(n)} + \hat{\theta}_2 \Phi_{(n)} \right)^{1/2} \alpha_n = \sigma_0^2$.

For the two-stage method A, proposed by Liu [12], the net expected return per risk is:
\[
Ratio_{n,A} = \frac{\sigma_0}{\zeta_{(n)}},
\]

where $\hat{\omega}$ is decided by the equation
\[
\zeta_{(n)}^T \left( D_{(n)} + \hat{\omega} \Gamma_{(n)} \right)^{1/2} \left( \hat{\omega} \Gamma_{(n)} + \frac{1}{2} D_{(n)} + \hat{\omega} \Gamma_{(n)} \right)^{1/2} \zeta_{(n)} = \sigma_0^2.
\]

Here $\zeta_{(n)} = E(\kappa_{(n)} \mid y_0)$, $\Gamma_{(n)} = Var(\kappa_{(n)} \mid y_0)$, $\kappa_{(n)} = \left( \hat{w}_1^T r_1, \hat{w}_2^T r_2, \cdots, \hat{w}_n^T r_n \right)^T$, and $\hat{w}_{i+1}$ is obtained by maximizing $g_{i+1}(w_{i+1} \mid r_i, y_0)$ defined by equation (8). $D_{(n)} = E(V_{(n)}^T \Lambda_{(n)} V_{(n)} \mid y_0)$, where $V_{(n)}$ has a structure similar to $U_{(n)}$, just replacing $\hat{w}_i$ by $\hat{w}_i$. For more detailed expressions about $\zeta_{(n)}$, $D_{(n)}$ and $\Gamma_{(n)}$ refer to Appendix B.

For the two-stage method B, the net expected return per risk is:
\[
Ratio_{n,B} = \frac{\sigma_0}{\zeta_{(n)}},
\]

where $\hat{\omega}$ is decided by the equation
\[
\zeta_{(n)}^T \left( C_{(n)} + \hat{\omega}_2 \Omega_{(n)} \right)^{1/2} \left( \hat{\omega}_2 \Omega_{(n)} + \frac{1}{2} C_{(n)} + \hat{\omega}_2 \Omega_{(n)} \right)^{1/2} \zeta_{(n)} = \sigma_0^2.
\]

For convenience, we set the two-stage method B as the benchmark for comparisons. Define the two ratios,
\[
Ratio_1 = \frac{Ratio_{n,B}}{Ratio_{n,S}} \quad \text{and} \quad Ratio_2 = \frac{Ratio_{n,B}}{Ratio_{n,A}};
\]

the former ratio explores the efficiency between the two-stage method B and the one-stage method; while the latter ratio provides impacts on the two two-stage methods, which are constructed under different conditional information.
particular, as the time period \( n \) gets longer. When values of \( \theta \) become larger, volatilities enhance, and the net expected returns for either the one-stage method or the two-stage method B decrease; however, the former declines faster than the latter. Referring to the two-stage method A, the results seem not as stable as comparing to that of the two-stage method B. In the meanwhile, both two-stage methods could obtain larger net expected returns than the one-stage method, as data volatilities increase.

As the time horizon gets longer, the two-stage method B outperforms the one–stage method: Since the two-stage method B obtains more previous information therefore net expected returns significantly increase. In comparisons between the two-stage, method B and method A, the numerical results showed that method B is always more efficient that method A. At each current time, method B determines the portfolio weights based on all the previous information, as opposed to method A, which only utilizes information from one period ahead. Thus more information may introduce higher net expected returns. In particular, as the discussed period gets longer, the efficiency becomes more significant. However, when the time horizon is shorter, as volatilities increase, the superiority of method B vanishes gradually and the performances of methods A and B tend to be close. In summary, the two-stage dynamic approaches are more efficient than the static methods, as data volatilities increase.

Referring to the two-stage standard method. As the time horizon gets longer, the two-stage standard method B/one-stage method A, the numerical results showed that net expected returns for either the one-stage standard method or the two-stage standard method decrease; however, the former declines faster than the latter. Referring to the two-stage standard method A, the results seem not as stable as comparing to that of the two-stage standard method B. In the meanwhile, both two-stage standard methods could obtain larger net expected returns than the one-stage standard method, as data volatilities increase.

Table 1. Comparisons of Net Expected Return per Risk

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \theta )</th>
<th>One-stage standard</th>
<th>Two-stage method A</th>
<th>Two-stage method B</th>
<th>Ratio_1</th>
<th>Ratio_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.0</td>
<td>0.9808</td>
<td>1.0753</td>
<td>1.0753</td>
<td>1.0963</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.8231</td>
<td>1.0157</td>
<td>1.0157</td>
<td>1.2340</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.7464</td>
<td>0.9807</td>
<td>0.9807</td>
<td>1.3139</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>0.5029</td>
<td>0.8744</td>
<td>0.8744</td>
<td>1.7387</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>0.3887</td>
<td>0.8372</td>
<td>0.8372</td>
<td>2.1540</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>0.0</td>
<td>1.1961</td>
<td>1.1856</td>
<td>1.3609</td>
<td>1.1378</td>
<td>1.1479</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.0085</td>
<td>1.1852</td>
<td>1.3166</td>
<td>1.3055</td>
<td>1.1109</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.9121</td>
<td>1.1750</td>
<td>1.2853</td>
<td>1.4092</td>
<td>1.0939</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>0.6121</td>
<td>1.1190</td>
<td>1.1860</td>
<td>1.9376</td>
<td>1.0599</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>0.4769</td>
<td>1.0928</td>
<td>1.1528</td>
<td>2.4137</td>
<td>1.0549</td>
</tr>
<tr>
<td>4</td>
<td>0.0</td>
<td>1.3791</td>
<td>1.3213</td>
<td>1.5935</td>
<td>1.1555</td>
<td>1.2060</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.1590</td>
<td>1.3503</td>
<td>1.5614</td>
<td>1.3472</td>
<td>1.1563</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.0516</td>
<td>1.3539</td>
<td>1.5292</td>
<td>1.4542</td>
<td>1.1295</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>0.7058</td>
<td>1.3201</td>
<td>1.4307</td>
<td>2.0271</td>
<td>1.0838</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>0.5501</td>
<td>1.3033</td>
<td>1.3982</td>
<td>2.5417</td>
<td>1.0728</td>
</tr>
<tr>
<td>6</td>
<td>0.0</td>
<td>1.6820</td>
<td>1.5490</td>
<td>1.9825</td>
<td>1.1787</td>
<td>1.2799</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.4156</td>
<td>1.6287</td>
<td>1.9585</td>
<td>1.3835</td>
<td>1.2025</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.2856</td>
<td>1.6533</td>
<td>1.9255</td>
<td>1.4977</td>
<td>1.1646</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>0.8649</td>
<td>1.6534</td>
<td>1.8231</td>
<td>2.1079</td>
<td>1.1027</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>0.6712</td>
<td>1.6419</td>
<td>1.7902</td>
<td>2.6667</td>
<td>1.0903</td>
</tr>
<tr>
<td>8</td>
<td>0.0</td>
<td>1.9366</td>
<td>1.7467</td>
<td>2.3052</td>
<td>1.1903</td>
<td>1.3198</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.6327</td>
<td>1.8655</td>
<td>2.2889</td>
<td>1.4019</td>
<td>1.2270</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.4814</td>
<td>1.9032</td>
<td>2.2561</td>
<td>1.5230</td>
<td>1.1854</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>0.9972</td>
<td>1.9285</td>
<td>2.1460</td>
<td>2.1520</td>
<td>1.1128</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>0.7760</td>
<td>1.9241</td>
<td>2.1105</td>
<td>2.7197</td>
<td>1.0969</td>
</tr>
<tr>
<td>10</td>
<td>0.0</td>
<td>2.1615</td>
<td>1.9261</td>
<td>2.5877</td>
<td>1.1972</td>
<td>1.3435</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.8203</td>
<td>2.0736</td>
<td>2.5755</td>
<td>1.4149</td>
<td>1.2420</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.6526</td>
<td>2.1269</td>
<td>2.5422</td>
<td>1.5383</td>
<td>1.1958</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>1.1140</td>
<td>2.1715</td>
<td>2.4264</td>
<td>2.1781</td>
<td>1.1174</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>0.8658</td>
<td>2.1680</td>
<td>2.3894</td>
<td>2.7596</td>
<td>1.1021</td>
</tr>
<tr>
<td>20</td>
<td>0.0</td>
<td>3.0285</td>
<td>2.6413</td>
<td>4.5480</td>
<td>1.5017</td>
<td>1.7219</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>2.5545</td>
<td>2.9054</td>
<td>4.5471</td>
<td>1.7800</td>
<td>1.5650</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>2.3189</td>
<td>3.0076</td>
<td>4.4524</td>
<td>1.9208</td>
<td>1.4804</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>1.5657</td>
<td>3.2248</td>
<td>4.1004</td>
<td>2.6189</td>
<td>1.2715</td>
</tr>
<tr>
<td></td>
<td>10.0</td>
<td>1.2188</td>
<td>3.2415</td>
<td>3.9971</td>
<td>3.2794</td>
<td>1.2331</td>
</tr>
</tbody>
</table>

Note:
1. \( \Sigma_{(n)} = I_k \otimes W_k \), \( W = W^* + \theta \times I_k \).
2. Ratio_1=two-stage method B/one-stage standard method.
3. Ratio_2=two-stage method B/ two-stage method A.
6. CONCLUSIONS
This paper provides an innovative two-stage dynamic multi-period portfolio selection approach with random portfolio weights and random quadratic transaction costs. The underlying assets time series data are supposed to follow a discrete-time triangle cointegrated model with vector autoregressive noise structure. The utilized mean-variance algorithm is focused on the one that maximizes the overall multi-period net expected returns, under a prescribed overall risk level, and a given initial asset value. Based on the proposed two-stage approach, the optimal portfolio allocation is analytically derived, expressed in a closed form. Moreover, some numerical results show that the proposed algorithm is tremendously more efficient than the static approach in particular, as the discussed time horizon becomes longer or data volatilities increase, higher net expected returns are indicated. Moreover, the performances between the two two-stage dynamic approaches exhibit that the one proposed in this article, which applies more previous information outperforms the other approach, which utilizes less information.

ACKNOWLEDGEMENTS
The author wishes to thank Miss. Meng-Yu Chan for her assistance in carrying out some data collection and parameter estimations.

APPENDIX A

Proofs of Lemma 1-3
The conditional property of a multivariate normal distribution: for \( i \geq 1 \),
\[
r_{i+1} | r_0, y_0 \sim MN(\xi_{i+1}, \Psi_{i+1}), \text{ and } w_{i+1} | r_0, y_0 \sim MN(w_{i+1}, \xi_{i+1}, \Psi_{i+1})
\]
where
\[
\xi_{i+1} = \alpha_{i+1} + \Phi_{i+1}(\xi_{i+1} - \alpha_{i+1}),\Psi_{i+1} = \Phi_{i+1} \Psi_{i+1} \Phi_{i+1}^T, \Phi_{i+1} = Cov(r_{i+1}, r_i), \Phi_{i+1} = Cov(r_{i+1}, r_j),
\]
\[
\Phi_{i+1}^{-1} = Cov(r_i, r_j), \text{ and } \Phi_{i+1}^{-1} = Cov(r_i, r_j).
\]
Therefore, the preliminary optimal weight is,
\[
\tilde{w}_{i+1} = \Psi_{i+1}^{-1} \xi_{i+1} = a_{i+1} + A_{i+1}^T r_i - \alpha_{i+1}, \text{ here } a_{i+1} = \Psi_{i+1}^{-1} \alpha_{i+1}, A_{i+1} = \Phi_{i+1}^{-1} \Phi_{i+1} + \Psi_{i+1}^{-1}
\]
Before developing the proofs of Lemma 1-3, for convenience, some preliminary results are stated without proofs as follows:

Property A1: Let \( Y \) be a random vector and normally distributed, with mean \( \mu \) and covariance matrix \( V \), and \( q_i = Y^T D Y \), here \( D \) is a symmetric matrix, then
\[
(1) \quad E(q_i) = tr(D V) + \mu^T D \mu
\]
\[
(2) \quad Var(q_i) = 2 tr(D V D^T V) + 4 \mu^T D V D \mu
\]
\[
(3) \quad Cov(q_i, q_j) = 2 tr(D V D^T V) + 4 \mu^T D V D \mu
\]
\[
(4) \quad Cov(Y, q_i) = 2 V D \mu
\]
Two more properties are briefly shown as follows, for more detailed derivations please refer to Liu [12]:

Property A2: \( Cov(r_i - \alpha_i, r_j) = A_{i,j}A_{j,i} + \Phi_{i,j} \alpha_{i,j} \)
\[
Pf: \quad Cov(r_i - \alpha_i, r_j) = A_{i,j}A_{j,i} + \Phi_{i,j} \alpha_{i,j}
\]
Property A3: \( Cov(r_{i+1} - \alpha_{i+1}, r_j) = A_{i,j}A_{j,i} + \Phi_{i,j} \alpha_{i,j} \)
\[
Pf: \quad Cov(r_{i+1} - \alpha_{i+1}, r_j) = A_{i,j}A_{j,i} + \Phi_{i,j} \alpha_{i,j}
\]
\[= \text{Cov}[(\ell_{i+1} - \alpha_{i+1})^T A_i \alpha, (\ell_{i+1} - \alpha_{i+1})^T A_j \alpha \mid y_0] + \frac{1}{4} \text{Cov}[(r - \alpha)^T A_i(\alpha - r) \mid y_0]
\]
\[= \alpha_i^T A_i^\dagger \Phi_{\ell_{i+1},j, j-1} A_j \alpha_j + \frac{1}{2} \text{tr}(A_i^\dagger \Phi \alpha \alpha^\dagger A_j^\dagger \Phi)
\]

**Proof of Lemma 1:**
\[E(\widehat{\nu}_i^T r_{i+1} \mid y_0) = E\left[a_i^T r_{i+1} + (r_{i+1} - \alpha_{i+1})^T A_i r_{i+1} \mid y_0\right]
\]
\[= a_i^T \alpha_{i+1} + E\left[(r_{i+1} - \alpha_{i+1})^T A_i r_{i+1} \mid y_0\right]
\]
\[= a_i^T \alpha_{i+1} + \text{tr}(A_i \Phi_{\ell_{i+1}, j, j-1} A_j), \text{ for } i \geq 1.
\]

**Proof of Lemma 2:**
\[\text{Cov}(\widehat{\nu}_i^T r, \widehat{\nu}_j^T r \mid y_0) = \text{Cov}\left[a_i^T r + (r_{i+1} - \alpha_{i+1})^T A_i r_{i+1}, a_j^T r + (r_{j+1} - \alpha_{j+1})^T A_j r_{j+1} \mid y_0\right]
\]
\[= a_i^T \Phi_{\ell_{i+1}, j} A_i \alpha_i + a_j^T \Phi_{\ell_{i+1}, j} A_j \alpha_j + a_i^T A_i^\dagger \Phi_{\ell_{i+1}, j} A_j \alpha_j + a_j^T A_j^\dagger \Phi_{\ell_{i+1}, j} A_i \alpha_i + \frac{1}{2} \text{tr}(A_i^\dagger \Phi A_j^\dagger \Phi).
\]

**Proof of Lemma 3:**
\[(3.3) E(\widehat{\nu}_i^T \Theta_i \bar{w}_j^T r \mid y_0) = E\left[a_i^T r + (r_{i+1} - \alpha_{i+1})^T A_i r_{i+1} + A_j^T (r_{j+1} - \alpha_{j+1}) \mid y_0\right]
\]
\[= a_i^T \Theta_i \alpha_i + E\left[(r_{i+1} - \alpha_{i+1})^T A_j \Theta_i A_j^T (r_{j+1} - \alpha_{j+1}) \mid y_0\right]
\]
\[= a_i^T \Theta_i \alpha_i + \frac{1}{2} E\left[(r - \alpha)^T C_{i,j}^* (r - \alpha) \mid y_0\right] = a_i^T \Theta_i \alpha_i + \frac{1}{2} \text{tr}(C_{i,j}^* \Phi_{\ell_{i+1}, j})
\]

**APPENDIX B**

Reviews of the Two-stage Method A

The following results are summarized from Liu [12]. Firstly, the preliminary objective function for the (i+1)-th period, giving previous return \(r_i\), was defined as

\[g_{i+1}(w_{i+1} \mid r_i, y_0) = E\left[w_{i+1}^T r_{i+1} \mid r_i, y_0\right] - \frac{1}{2} \text{Var}(w_{i+1}^T r_{i+1} \mid r_i, y_0) = w_{i+1}^T \tau_{i+1} r_i - \frac{1}{2} w_{i+1}^T \Delta_{i+1} w_{i+1},
\]

where \(\tau_{i+1} = \alpha_{i+1} + \Phi_{\ell_{i+1}, j} \Phi_{j, j-1}^T (r_i - \alpha_j)\), \(\Delta_{i+1} = \Phi_{\ell_{i+1}, j}^T - \Phi_{\ell_{i+1}, j} \Phi_{j, j-1}^T \Phi_{j, j-1}\), and \(\Phi_{j, j-1} = \text{Cov}(r_j, r_{j-1})\). Thus the preliminary optimal solution of \(w_{i+1}\), denoted as \(\hat{w}_{i+1}\), is derived as:

\[\hat{w}_{i+1} = \Delta_{i+1}^{-1} \tau_{i+1} = b_{i+1} + B_{i+1}^T (r_i - \alpha_j),\]


Next, the overall optimal portfolio weights are obtained, by finding suitable \(w_1\) and \(\rho\), which maximize the following objective function:

\[E\left[w_1^T r + \rho_2 \hat{w}_2^T r + \ldots + \rho_n \hat{w}_n^T r \mid y_0\right] - \frac{1}{2} \sum_{i=1}^n E\left[(\rho_i \hat{w}_i - \rho_{i-1} \hat{w}_{i-1})^T \Theta_i (\rho_i \hat{w}_i - \rho_{i-1} \hat{w}_{i-1}) \mid y_0\right]
\]

\[-\frac{1}{2} \sum_{i=1}^n \text{Var}[\rho_i \hat{w}_i + \rho_{i-1} \hat{w}_{i-1}] - \sigma_w^2 \]

where \(\hat{w}_i = w_i\), \(\hat{w}_0 = 0\), \(\rho_1 = 1\), \(\rho_0 = 0\) and \(\lambda_w\) is the Lagrange multiplier.

To re-organize the objective function, the following notations are introduced:
Then \( w_{(n),a} = V_{(a)} V_{(n)} \), \( w_{(n),a}^T r_n = w_1^T r_1 + \rho_2 w_2^T r_2 + \cdots + \rho_n w_n^T r_n = v_{(n)}^T \kappa_{(n)} \), and
\[
w_{(n),a}^T \Lambda_{(n)} w_{(n),a} = v_{(n)}^T V_{(n)} \Lambda_{(n)} V_{(n)} v_{(n)}, \text{ and the overall objective function is re-written as:}
\]
\[
g_a(v_{(a)} | y_0) = v_{(a)}^T \kappa_{(n)} - \frac{1}{2} v_{(a)}^T D_{(a)} v_{(n)} - \frac{\hat{\lambda}_a}{2} \left( v_{(a)}^T \Gamma_{(a)} v_{(a)} - \sigma_a^2 \right),
\]
here \( \kappa_{(n)} = E(\kappa_{(n)} | y_0) \), \( \Gamma_{(a)} = Var(\kappa_{(a)} | y_0) \) and \( D_{(a)} = E(V_{(a)}^T \Lambda_{(a)} V_{(a)} | y_0) \).

The optimal solution is \( \hat{v}_{(a)} = (D_{(a)} + \hat{\lambda}_a \Gamma_{(a)})^T \kappa_{(a)} \), where \( \hat{\lambda}_a \) is decided by solving the equation
\[
\kappa_{(n)}^T (D_{(a)} + \hat{\lambda}_a \Gamma_{(a)})^T \kappa_{(a)} = \sigma_a^2.
\]
Finally, the optimal net returns under the two-stage method A is:
\[
g_a(\hat{v}_{(a)} | y_0) = \frac{1}{2} \kappa_{(n)}^T (D_{(a)} + \hat{\lambda}_a \Gamma_{(a)})^T \kappa_{(a)} + \frac{\hat{\lambda}_a}{2} \sigma_a^2.
\]
And the overall net expected return per risk is defined as:
\[
\text{Ratio}_{a,a} = \frac{\hat{\lambda}_a}{\sigma_a}.
\]
The detailed representations of \( \kappa_{(n)} \), \( \Gamma_{(a)} \), and \( D_{(a)} \) are summarized as follows:

(B1) \( \kappa_{(n)}^T = E(\kappa_{(n)} | y_0)^T = \left[ E(\kappa_1 | y_0)^T, E(\kappa_2^T r_2 | y_0), \cdots, E(\kappa_n^T r_n | y_0) \right] \), where
\[
E(\kappa_i | y_0) = \alpha_i \text{ and } E(\kappa_i^T r_i | y_0) = b_i^T \alpha_i + tr(B_i \Phi_i), \text{ for } i = 1, 2, \cdots, n - 1.
\]

(B2) \( \Gamma_{(a)} = \left[ \begin{array}{ccc} Cov(\kappa_1, \kappa_1 | y_0) & \cdots & Cov(\kappa_1, \kappa_n r_n | y_0) \\ \vdots & \ddots & \vdots \\ Cov(\kappa_n r_n, \kappa_1 | y_0) & \cdots & Cov(\kappa_n r_n, \kappa_n r_n | y_0) \end{array} \right] \), where
\[
\text{Cov}(\kappa_i, \kappa_j^T r_j | y_0) = (\Phi_{i,j} + \Phi_{j,i} - \Phi_{i,j} \Phi_{j,i}), \text{ for } i = 1, 2, \cdots, n - 1; \text{ and }
\]
\[
\text{Cov}(\kappa_i^T r_i, \kappa_j | y_0) = b_i^{*T} \Phi_i b_j + b_j^{*T} \Phi_i b_j + b_i^{*T} \Phi_{i,j} + b_j^{*T} \Phi_{j,i} B_i \alpha_i + \alpha_i^{*T} B_i \Phi_i B_i \Phi_i \alpha_i + \frac{1}{2} tr(B_i \Phi_i B_i) \phi_i^2, \text{ for } 2 \leq i, j \leq n.
\]
Here, for \( j \geq 2 \), \( B_j^* = \tilde{B}_j^* + \frac{1}{2} tr(\Phi_i B_i \Phi_i) \), and
\[
\tilde{B}_j^* = \begin{cases} B_j, & \text{if } (j - 2)k + 1 \leq s \leq (j - 1)k \text{ and } (j - 1)k + 1 \leq l \leq jk, \\ 0, & \text{otherwise}. \end{cases}
\]
\[
\begin{bmatrix}
2\Theta & -\Theta E(\hat{\omega}^T | y_0) \\
-E(\hat{\omega}^T | y_0)\Theta & 2E(\hat{\omega}^T_2 \Theta \hat{\omega}_2 | y_0) - E(\hat{\omega}^T_2 \Theta \hat{\omega}_1 | y_0) \\
& \vdots \\
& 0 & -E(\hat{\omega}^T_{n-1} \Theta \hat{\omega}_{n-2} | y_0) & 2E(\hat{\omega}^T_{n-1} \Theta \hat{\omega}_{n-1} | y_0) - E(\hat{\omega}^T_{n-1} \Theta \hat{\omega}_n | y_0) \\
& & & -E(\hat{\omega}^T_n \Theta \hat{\omega}_{n-1} | y_0) & E(\hat{\omega}^T_n \Theta \hat{\omega}_n | y_0)
\end{bmatrix}
\]

(B3) \[D_{(n)} = \begin{cases}
2\Theta & -\Theta E(\hat{\omega}^T | y_0) \\
-E(\hat{\omega}^T | y_0)\Theta & 2E(\hat{\omega}^T_2 \Theta \hat{\omega}_2 | y_0) - E(\hat{\omega}^T_2 \Theta \hat{\omega}_1 | y_0) \\
& \vdots \\
& 0 & -E(\hat{\omega}^T_{n-1} \Theta \hat{\omega}_{n-2} | y_0) & 2E(\hat{\omega}^T_{n-1} \Theta \hat{\omega}_{n-1} | y_0) - E(\hat{\omega}^T_{n-1} \Theta \hat{\omega}_n | y_0) \\
& & & -E(\hat{\omega}^T_n \Theta \hat{\omega}_{n-1} | y_0) & E(\hat{\omega}^T_n \Theta \hat{\omega}_n | y_0)
\end{cases}\]

here \(E(\hat{\omega}^T | y_0) = b_2\); for \(i \geq 2 \) & \( j \geq 2 \), \(E(\hat{\omega}^T_i \Theta \hat{\omega}_j | y_0) = b^T_i \Theta b_j + \frac{1}{2} \text{tr}(D_{(i,j)} \Phi)\), with

\[
D_{(i,j)} = \hat{D}_{(i,j)} + \hat{D}^T_{(i,j)}, \quad \text{and} \quad \left(\hat{D}_{(i,j)}\right)_{i,j} = \begin{cases}
(B_2 \Omega B_2^T)_{i-(i-2),j-(i-2)} & \text{if} \quad (i-2)k + 1 \leq s \leq (i-1)k \\
& \text{&} \quad (j-2)k + 1 \leq l \leq (j-1)k,
0, & \text{otherwise.}
\end{cases}
\]

REFERENCES