Optimizing Financial Portfolios from the Perspective of Mining Temporal Structures of Stock Returns

Kai-Chun Chiu and Lei Xu

Department of Computer Science and Engineering
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, P.R. China
`{kcchiu,lxu}@cse.cuhk.edu.hk`

Abstract. In this paper, we aim at introducing how one of the recently developed statistical learning techniques, temporal factor analysis (TFA), which is originally devoted for further study of the arbitrage pricing theory (APT), could be exploited in financial data mining to determine weights in portfolio optimization problem. Furthermore, we study several variants of the APT-based Sharpe ratio maximization technique that utilize the concepts of portfolio downside risk and upside volatility tailored-made for the need of risk-averse as well as aggressive profit-seeking investors.

1 Introduction

A typical task in modern financial data mining is dynamic portfolio optimization. Ever since the inception of Markowitz's modern portfolio theory in 1952, the finance community was overwhelmed by the possibility of first constructing and subsequently investing in the so-called mean-variance efficient portfolios. Prior to that, however, mean-variance portfolio optimization was hardly a practice. Instead, investment practitioners focused on spotting securities with high expected returns. Meanwhile, theoretical research on investments concentrated on modelling expected returns [1], with empirical research focused on testing such equilibrium models, or documenting patterns in stock returns that appear to be inconsistent with these models.

Even though portfolio optimization has been shown to be theoretically feasible, its realization is often constrained to the static stage. That is, the optimal portfolio weights determined based on historical data could not change within the same investment horizon. The situation could be contrasted with dynamic portfolio optimization by which the optimal portfolio weights were to be determined over time according to the most up-to-date price movement information in the market.

Unarguably, dynamic portfolio management was more appealing than static portfolio management due to today's shortened execution time for each transaction and given the tremendous volume of transactions occurred every day. Nonetheless, dynamic portfolio optimization typically increases the computational load. Recall that the most natural technique for solving dynamic portfolio optimization problems is stochastic dynamic programming. However, this approach is often compromised by several factors such as the curse of dimensionality when too many state variables are involved [2] . In general, practical considerations such as taxes, transactions costs increase the number of state variables in the objective function.

Just like other basic problems in financial analysis, the optimal portfolio selection problem was resilient to the onslaught of researchers from the disciplines of artificial neural networks and machine learning [3–6]. Using the neural networks approach requires optimal parameters to be adaptively learned over time. Consequently, we use the term adaptive portfolio management to refer to solving the traditional portfolio optimization problem with the help of artificial neural networks. On the other hand, since trading based on training a trading system on labelled data usually yields suboptimal results [7], adaptive portfolio management via maximizing the well-known Sharpe ratio [8, 9] was studied in [7, 3]. However, such approaches are still far from perfect because they either explicitly treat the weights as constants or depend directly on the security returns, which may not be an optimal solution.

Recently, a new technique called temporal factor analysis (TFA) was proposed in [10] with an aim at providing an alternative way for implementing the classical financial APT model. In this paper, we investigate how the APT-based gaussian TFA model can be used for adaptive portfolio management. Our analysis consists of mainly two aspects. First, we examine this APT-based model from the perspective of mining temporal structure of stock returns and make numerous comparisons with other closely related techniques. Second, we consider some variants of the approach, especially for risk-averse and aggressive profit-oriented investors.

The rest of the paper is organized in the following way. Sections 2 and 3 briefly review the APT and the gaussian TFA models respectively. Section 4 illustrates, with algorithms, how to implement the APT-based gaussian TFA learning for adaptive portfolio management. Three variants of the APT-based Sharpe ratio maximization technique are studied in section 5. Section 6 concludes the paper.

2   Review on Arbitrage Pricing Theory

The APT begins with the assumption that the $n \times 1$ vector of asset returns, R_t, is generated by a linear stochastic process with k factors [11–13]:

$$\mathbf{R_t} = \bar{\mathbf{R}} + \mathbf{Af_t} + \mathbf{e_t} \tag{1}$$

where f_t is the $k \times 1$ vector of realizations of k common factors, A is the $n \times k$ matrix of factor weights or loadings, and e_t is a $n \times 1$ vector of asset-specific risks. It is assumed that f_t and e_t have zero expected values so that $\bar{R}$ is the $n \times 1$ vector of mean returns.

3   Overview of Temporal Factor Analysis

Suppose the relationship between a state $\mathbf{y_t} \in \mathbb{R}^k$ and an observation $\mathbf{x_t} \in \mathbb{R}^d$ is described by the first-order state-space equations as follows [10, 14]:

$$\mathbf{y_t} = \mathbf{By_{t-1}} + \varepsilon_{\mathbf{t}}, \tag{2}$$

$$\mathbf{x_t} = \mathbf{Ay_t} + \mathbf{e_t}, \qquad t = 1, 2, \ldots, N. \tag{3}$$

where $\varepsilon_\mathbf{t}$ and $\mathbf{e_t}$ are mutually independent zero-mean white noises with $E(\varepsilon_\mathbf{i}\varepsilon_\mathbf{j}^T) = \Sigma_\varepsilon \delta_{ij}$, $E(\mathbf{e_i e_j}^T) = \Sigma_\mathbf{e}\delta_{ij}$, $E(\varepsilon_\mathbf{i}\mathbf{e_j}^T) = 0$, Σ_ε and $\Sigma_\mathbf{e}$ are diagonal matrices, and δ_{ij} is the Kronecker delta function:

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

We call ε_t driving noise upon the fact that it drives the source process over time. Similarly, e_t is called measurement noise because it happens to be there during measurement. The above model is generally referred to as the TFA model.

In the context of APT analysis, (1) can be obtained from equation (3) by substituting $(\tilde{R}_t - \bar{R})$ for $\mathbf{x_t}$ and $\mathbf{f_t}$ for $\mathbf{y_t}$. The only difference between the APT model and the TFA model is the added equation (2) for modelling temporal relation of each factor. The added equation represents the factor series $\mathbf{y} = \{\mathbf{y_t}\}_{t=1}^T$ in a multi-channel auto-regressive process, driven by an i.i.d. noise series $\{\varepsilon_\mathbf{t}\}_{t=1}^T$ that are independent of both $\mathbf{y_{t-1}}$ and $\mathbf{e_t}$. This can also be viewed as one way of mining the temporal structures of stock returns. Specifically, it is assumed that $\varepsilon_\mathbf{t}$ is gaussian distributed. Moreover, TFA is defined such that the k sources $y_t^{(1)}, y_t^{(2)}, \ldots, y_t^{(k)}$ in this state-space model are statistically independent. The objective of TFA is to estimate the sequence of y_t's with unknown model parameters $\Theta = \{\mathbf{A}, \mathbf{B}, \Sigma_\varepsilon, \Sigma_\mathbf{e}\}$ through available observations. In implementation, we adopt an adaptive algorithm in [14].

4 Gaussian TFA for Adaptive Portfolio Management

When the APT-based gaussian TFA model is adopted for portfolio management, portfolio weights adjustment can be made under the control of independent hidden factors that affect the portfolio. In the sequel, we illustrate how this can be achieved under the following four scenarios:

	Transaction Cost	Short Sale Permission
Scenario I	No	No
Scenario II	Yes	No
Scenario III	No	Yes
Scenario IV	Yes	Yes

4.1 Scenario I: NO Transaction Cost and Short Sale NOT Permitted

The assumptions underlying this scenario are no transaction cost and short sale not permitted. Consequently, we consider the return of a typical portfolio which is given by [14]

$$R_t = (1 - \alpha_t)r^f + \alpha_t \sum_{j=1}^m \beta_t^{(j)} x_t^{(j)}, \text{ subject to } \begin{cases} \alpha_t > 0, \\ 0 \leq \beta_t \leq 1, \\ \sum_{j=1}^m \beta_t^{(j)} = 1. \end{cases} \tag{5}$$

where r^f denotes the risk-free rate of return, x_t denotes returns of risky securities, α_t the proportion of total capital to be invested in risky securities and $\beta_t^{(j)}$ the proportion of α_t to be invested in the jth risky asset.

Instead of focussing on the mean-variance efficient frontier, we seek to optimize the portfolio Sharpe ratio [15] $S_p = M(R_T)/\sqrt{V(R_T)}$ by [14]:

$$\max_{\psi,\phi} S_p = \frac{M(R_T)}{\sqrt{V(R_T)}} \text{ subject to } \begin{cases} \alpha_t = \exp\left(\zeta_t\right), \\ \zeta_t = g(\mathbf{y_t}, \psi), \\ \beta_t^{(j)} = \exp\left(\xi_t^{(j)}\right)/\sum_{r=1}^m \exp\left(\xi_t^{(r)}\right), \\ \xi_t = f(\mathbf{y_t}, \phi). \end{cases} \tag{6}$$

where $M(R_T) = \frac{1}{T}\sum_{t=1}^T R_t$ is expected return and $V(R_T) = \frac{1}{T}\sum_{t=1}^T [R_t - M(R_T)]^2$ is a measure of risk or volatility, $\{\mathbf{y_t}\}_{t=1}^N$ is the time series of independent hidden factors that drives the observed return series $\{\mathbf{x_t}\}_{t=1}^N$, $g(\mathbf{y_t}, \psi)$ and $f(\mathbf{y_t}, \phi)$ are some nonlinear functions that map $\mathbf{y_t}$ to respectively ζ_t and ξ_t which in turn adjusts the portfolio weights α_t and $\beta_t^{(j)}$ respectively.

Maximizing the portfolio Sharpe ratio in effect balances the tradeoff between maximizing the expected return and at the same time minimizing the risk. In implementation, we can simply use the gradient ascent approach. The time series $\{\mathbf{y_t}\}_{t=1}^N$ can be estimated via the gaussian TFA algorithm in [14]. Although the functions $g(\mathbf{y_t}, \psi)$ and $f(\mathbf{y_t}, \phi)$ are not known *a priori*, it may be approximated via the adaptive extended normalized radial basis function (ENRBF) algorithm in [16].

Like radial basis function (RBF) network, ENRBF is one of the popular models adopted for function approximation. The general form of RBF is

$$f_k(\mathbf{x}) = \sum_{j=1}^k w_j \varphi([\mathbf{x} - \mu_{\mathbf{j}}]^T \mathbf{\Sigma_j}^{-1}[\mathbf{x} - \mu_{\mathbf{j}}]) \tag{7}$$

ENRBF is an improved modification of RBF by replacing w_j with a linear vector function $\mathbf{W_j^T}\mathbf{x} + \mathbf{c_j}$ and dividing the term $\varphi([\mathbf{x} - \mu_{\mathbf{j}}]^T \mathbf{\Sigma_j}^{-1}[\mathbf{x} - \mu_{\mathbf{j}}])$ over the aggregate of all terms to arrive at

$$f_k(\mathbf{x}) = \frac{\sum_{j=1}^k (\mathbf{W_j}^T\mathbf{x} + \mathbf{c_j})\varphi([\mathbf{x} - \mu_{\mathbf{j}}]^T \mathbf{\Sigma_j}^{-1}[\mathbf{x} - \mu_{\mathbf{j}}])}{\sum_{j=1}^k \varphi([\mathbf{x} - \mu_{\mathbf{j}}]^T \mathbf{\Sigma_j}^{-1}[\mathbf{x} - \mu_{\mathbf{j}}])} \tag{8}$$

where $\mathbf{W_j}$ is a parameter matrix.

Basically, each $\mathbf{W_j^T}\mathbf{x} + \mathbf{c_j}$ represents a local linear segment. The ENRBF network approximates a globally nonlinear function by smoothly joining all piecewise linear segments. The set of parameters to be estimated is $\Theta = \{\mu_{\mathbf{j}}, \mathbf{\Sigma_j}, \mathbf{W_j}, \mathbf{c_j}\}_{j=1}^k$.

Specifically, $g(\mathbf{y_t}, \psi)$ and $f(\mathbf{y_t}, \phi)$ can be modelled by the ENRBF shown below.

$$g(\mathbf{y_t}, \psi) = \sum_{p=1}^k (\mathbf{W_p^T}\mathbf{y_t} + c_p)\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k) \tag{9}$$

$$f(\mathbf{y_t}, \phi) = \sum_{p=1}^{\hat{k}} (\mathbf{\hat{W}_p^T}\mathbf{y_t} + \mathbf{\hat{c}_p})\varphi(\hat{\mu}_{\mathbf{p}}, \mathbf{\hat{\Sigma}_p}, \hat{k}) \tag{10}$$

where $\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k) = \frac{\exp\left(-0.5(\mathbf{y_t}-\mu_{\mathbf{p}})^T\mathbf{\Sigma_p}^{-1}(\mathbf{y_t}-\mu_{\mathbf{p}})\right)}{\sum_{r=1}^{k}\exp\left(-0.5(\mathbf{y_t}-\mu_{\mathbf{r}})^T\mathbf{\Sigma_p}^{-1}(\mathbf{y_t}-\mu_{\mathbf{r}})\right)}$.

The set of parameters in (9) and (10) to be estimated is Θ where $\Theta = \psi \cup \phi, \psi = \{\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, \mathbf{W_p}, c_p\}_{p=1}^{k}$ and $\phi = \{\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{\mathbf{W}}_{\mathbf{p}}, \hat{\mathbf{c}}_{\mathbf{p}}\}_{p=1}^{\hat{k}}$. In general, for each $\theta \in \Theta$, updating takes place adaptively in the following form:

$$\theta^{\mathrm{new}} = \theta^{\mathrm{old}} + \eta_0 \nabla_\theta S_p \tag{11}$$

where η_0 is the learning step size, $\nabla_\theta S_p$ denotes the gradient with respect to θ in the ascent direction of S_p. Typically, the adaptive algorithm [17] shown in Table 1 can be adopted for implementation.

Table 1. An adaptive algorithm for implementation of the APT-based portfolio management

---

Updating rules for the parameter set ψ

$\mu_{\mathbf{p}}^{\mathrm{new}} = \mu_{\mathbf{p}}^{\mathrm{old}} + \eta(\nabla_{\zeta_T} S_p)\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k)\tau(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, \mathbf{W_p}, c_p, k)(\mathbf{y_T} - \mu_{\mathbf{p}})$

$\mathbf{\Sigma_p}^{\mathrm{new}} = \mathbf{\Sigma_p}^{\mathrm{old}} + \eta(\nabla_{\zeta_T} S_p)\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k)\tau(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, \mathbf{W_p}, c_p, k)\kappa(\mu_{\mathbf{p}}, \mathbf{\Sigma_p})$

$\mathbf{W}_p^{\mathrm{new}} = \mathbf{W}_p^{\mathrm{old}} + \eta(\nabla_{\zeta_T} S_p)\mathbf{y_T}\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k)$

$c_p^{\mathrm{new}} = c_p^{\mathrm{old}} + \eta(\nabla_{\zeta_T} S_p)\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k)$

Updating rules for the parameter set ϕ

$\hat{\mu}_{\mathbf{p}}^{\mathrm{new}} = \hat{\mu}_{\mathbf{p}}^{\mathrm{old}} + \hat{\eta}(\nabla_{\xi_T^{(j)}} S_p)(\mathbf{y_T} - \hat{\mu}_{\mathbf{p}})\varphi(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{k})\chi(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{\mathbf{W}}_{\mathbf{p},\mathbf{q}}, \hat{c}_{p,q}, \hat{k})$

$\hat{\mathbf{\Sigma}}_{\mathbf{p}}^{\mathrm{new}} = \hat{\mathbf{\Sigma}}_{\mathbf{p}}^{\mathrm{old}} + \hat{\eta}(\nabla_{\xi_T^{(j)}} S_p)\kappa(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}})\varphi(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{k})\chi(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{\mathbf{W}}_{\mathbf{p},\mathbf{q}}, \hat{c}_{p,q}, \hat{k})$

$\hat{\mathbf{W}}_{\mathbf{p},\mathbf{q}}^{\mathrm{new}} = \hat{\mathbf{W}}_{\mathbf{p},\mathbf{q}}^{\mathrm{old}} + \hat{\eta}(\nabla_{\xi_T^{(j)}} S_p)\mathbf{y_T}\varphi(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{k})$

$\hat{c}_{p,r}^{\mathrm{new}} = \hat{c}_{p,r}^{\mathrm{old}} + \hat{\eta}(\nabla_{\xi_T^{(j)}} S_p)\varphi(\hat{\mu}_{\mathbf{p}}, \hat{\mathbf{\Sigma}}_{\mathbf{p}}, \hat{k})$

where η and $\hat{\eta}$ are learning rates,

$M(R_T) = \frac{1}{T}\sum_{t=1}^{T} R_t, \quad V(R_T) = \frac{1}{T}\sum_{t=1}^{T}[R_t - M(R_T)]^2$

$\nabla_{\zeta_T} S_p = \frac{\left[V(R_T)-M(R_T)\left(R_T-M(R_T)-\frac{1}{T}\sum_{t=1}^{T}(R_t-M(R_t))\right)\right]}{T\sqrt{[V(R_T)]^3}}\left(\frac{\sum_{r=1}^{m}\exp(\xi_T^{(r)})x_T^{(r)}}{\sum_{r=1}^{m}\exp(\xi_T^{(r)})} - r^f\right)\exp(\zeta_T)$,

$\nabla_{\xi_T^{(j)}} S_p = \frac{\left[V(R_T)-M(R_T)\left(R_T-M(R_T)-\frac{1}{T}\sum_{t=1}^{T}(R_t-M(R_t))\right)\right]\exp(\zeta_T)x_T^{(j)}\left(\sum_{r=1}^{m}\exp(\xi_T^{(r)})-\exp(\xi_T^{(j)})\right)\exp(\xi_T^{(j)})}{T\sqrt{[V(R_T)]^3}\left(\sum_{r=1}^{m}\exp(\xi_T^{(r)})\right)^2}$,

$\varphi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, k) = \frac{\exp(-0.5(\mathbf{y_T}-\mu_{\mathbf{p}})^T\mathbf{\Sigma_p}^{-1}(\mathbf{y_T}-\mu_{\mathbf{p}}))}{\sum_{r=1}^{k}\exp(-0.5(\mathbf{y_T}-\mu_{\mathbf{r}})^T\mathbf{\Sigma_r}^{-1}(\mathbf{y_T}-\mu_{\mathbf{r}}))}$,

$\kappa(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}) = \mathbf{\Sigma_p}^{-1}(\mathbf{y_T}-\mu_{\mathbf{p}})(\mathbf{y_T}-\mu_{\mathbf{p}})^T\mathbf{\Sigma_p}^{-1} - 0.5\mathrm{diag}[\mathbf{\Sigma_p}^{-1}(\mathbf{y_T}-\mu_{\mathbf{p}})(\mathbf{y_T}-\mu_{\mathbf{p}})^T\mathbf{\Sigma_p}^{-1}]$,

$\tau(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, \mathbf{W_p}, c_p, k) = \frac{(\mathbf{W_p}^T\mathbf{y_T}+c_p)-\sum_{r=1}^{k}(\mathbf{W_r}^T\mathbf{y_T}+c_r)\varphi(\mu_{\mathbf{r}}, \mathbf{\Sigma_r}, k)}{\sum_{r=1}^{k}\exp(-0.5(\mathbf{y_T}-\mu_{\mathbf{r}})^T\mathbf{\Sigma_r}^{-1}(\mathbf{y_T}-\mu_{\mathbf{r}}))}$,

$\chi(\mu_{\mathbf{p}}, \mathbf{\Sigma_p}, \mathbf{W_{p,q}}, c_{p,q}, k) = \frac{(\mathbf{W_{p,q}}^T\mathbf{y_T}+c_{p,q})-\sum_{r=1}^{k}(\mathbf{W_{r}}^T\mathbf{y_T}+c_r)\varphi(\mu_{\mathbf{r}}, \mathbf{\Sigma_r}, k)}{\sum_{r=1}^{k}\exp(-0.5(\mathbf{y_T}-\mu_{\mathbf{r}})^T\mathbf{\Sigma_r}^{-1}(\mathbf{y_T}-\mu_{\mathbf{r}}))}$,

$\mathbf{W_{p,q}}$ denotes the p-th row of the q-th matrix,

$\mathrm{diag}[M]$ denotes a diagonal matrix that takes the diagonal part of a matrix M,

$\zeta_T = g(\mathbf{y_T}, \psi)$ as defined in (6) and $\xi_T^{(j)}$ is the j-th output of $f(\mathbf{y_T}, \phi)$ as defined in (7).

4.2 Simulation

Data Considerations All simulations in this paper are based on the past average fixed deposit interest rate, stock and index data of Hong Kong. Daily closing prices of the 1-week bank aver-

age interest rate, 3 major stock indices as well as 86 actively trading stocks covering the period from January 1, 1998 to December 31, 1999 are used. The number of trading days throughout this period is 522. The three major stock indices are respectively Hang Seng Index (HSI), Hang Seng China-Affiliated Corporations Index (HSCCI) and Hang Seng China Enterprises Index (HSCEI). Of the 86 equities, 30 of them are HSI constituents, 32 are HSCCI constituents and the remaining 24 are HSCEI constituents. The index data are directly used for adaptive portfolio management while the stock prices are used by gaussian TFA for recovering independent hidden factors y_t.

Methodology We consider the task of managing a portfolio which consists of four securities, the average fixed deposit interest rate and the three major stock indices in Hong Kong. The fixed deposit interest rate is used as the proxy for the risk-free rate of return r^f. The first 400 samples are used for training and the last 121 samples for testing. Both training and test are carried out in an adaptive fashion. The APT-based algorithm in Table 1 is adopted that uses hidden independent factors for controlling portfolio weights. We refer to this approach APT-based portfolio management. For each $\mathbf{y}_t$ under test, we can adaptively get $\zeta_t = g(\mathbf{y_t}, \psi)$ and $\xi_t = f(\mathbf{y_t}, \phi_t)$ and then the portfolio weights $\alpha_t = \exp(\zeta_t)$ and $\beta_t^{(j)} = \exp(\xi_t^{(j)})/\sum_{r=1}^{m} \exp(\xi_t^{(r)})$. Finally, returns can be computed via (5). For the sake of comparison, we also implement a traditional approach that directly uses stock returns $\mathbf{x_t}$ instead of hidden factors $\mathbf{y}_t$ [15]. We refer to this approach return-based portfolio management. Daily risk-return statistics of the portfolios are given in Table 2.

Table 2. Daily risk-return statistics of constituents of portfolios

Component Name	Mean Return	Risk
Average interest rate	0.0148%	0.00%
HSI	0.18%	1.48%
HSCCI	0.03%	2.51%
HSCEI	-0.20%	2.55%

Simulation Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 1(a) while daily risk-return statistics of the portfolios are given in Table 3(a).

4.3   Scenario II: HAS Transaction Cost but Short Sale NOT Permitted

Scenario II differs from Scenario I in taking into account the effect of transaction cost. Since any change on $\beta_t^{(j)}$ leads to a transaction that incurs a cost on return c_t given by

$$c_t = -\alpha_t \sum_{j=1}^{m} r_c |\beta_t^{(j)} - \beta_{t-1}^{(j)}| p_t^{(j)}/p_{t-1}^{(j)} = -\alpha_t \sum_{j=1}^{m} r_c |\beta_t^{(j)} - \beta_{t-1}^{(j)}|(1 + x_t^{(j)}) \qquad (12)$$

where r_c is a constant denoting the rate of transaction cost. Consequently, we consider the portfolio return adjusted for transaction cost given by [14]

$$R_t = (1 - \alpha_t)r^f + \alpha_t \sum_{j=1}^{m} \left[\beta_t^{(j)} x_t^{(j)} - r_c |\beta_t^{(j)} - \beta_{t-1}^{(j)}|(1 + x_t^{(j)}) \right], \tag{13}$$

$$\text{subject to} \begin{cases} \alpha_t > 0, \\ 0 \leq \beta_t \leq 1, \\ \sum_{j=1}^{m} \beta_t^{(j)} = 1. \end{cases}$$

The APT-based algorithm in Table 1 could still be adopted in this case except that the two terms $\nabla_{\zeta_T} S_p$ and $\nabla_{\xi_T^{(j)}} S_p$ become respectively

$$\nabla_{\zeta_T} S_p = \frac{\left[V(R_T) - M(R_T)\left(R_T - M(R_T) - \frac{1}{T}\sum_{t=1}^{T}(R_t - M(R_t)) \right) \right]}{T\sqrt{[V(R_T)]^3}}$$
$$\cdot \left(\sum_{j=1}^{m} \left[\frac{\exp(\xi_T^{(j)}) x_T^{(j)}}{\sum_{r=1}^{m} \exp(\xi_T^{(r)})} - r_c \left| \frac{\exp(\xi_T^{(j)})}{\sum_{r=1}^{m} \exp(\xi_T^{(r)})} - \frac{\exp(\xi_{T-1}^{(j)})}{\sum_{r=1}^{m} \exp(\xi_{T-1}^{(r)})} \right| (1 + x_T^{(j)}) \right] - r^f \right)$$
$$\cdot \exp(\zeta_T),$$

$$\nabla_{\xi_T^{(j)}} S_p = \left[V(R_T) - M(R_T)\left(R_T - M(R_T) - \frac{1}{T}\sum_{t=1}^{T}(R_t - M(R_t)) \right) \right]$$
$$\cdot \exp(\zeta_T)\left[x_T^{(j)} - r_c \text{sign}\left(\exp(\xi_T^{(j)}) - \exp(\xi_T^{(j-1)}) \right) \right]$$
$$\cdot \left[\sum_{r=1}^{m} \exp(\xi_T^{(r)}) - \exp(\xi_T^{(j)}) \right] \exp(\xi_T^{(j)}) / \left[T\sqrt{[V(R_T)]^3} \left(\sum_{r=1}^{m} \exp(\xi_T^{(r)}) \right)^2 \right]$$

Simulation For the purpose of simulation we fix the rate of transaction cost at $r_c = 0.1\%$. Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 1(b) while daily risk-return statistics of the portfolios are given in Table 3(b).

4.4 Scenario III: NO Transaction Cost but Short Sale IS Permitted

Scenario III differs from Scenario I in that short sale is now permitted. By removing the non-negative constraints on α_t and β_t in (5), we get

$$R_t = (1 - \alpha_t)r^f + \alpha_t \sum_{j=1}^{m} \beta_t^{(j)} x_t^{(j)} - r_c |\beta_t^{(j)} - \beta_{t-1}^{(j)}|(1 + x_t^{(j)}) \quad \text{subject to} \sum_{j=1}^{m} \beta_t^{(j)} = 1 \tag{14}$$

and the new objective function

$$\max_{\psi, \phi} S_p = \frac{M(R_T)}{\sqrt{V(R_T)}} \quad \text{subject to} \begin{cases} \alpha_t = \zeta_t = g(\mathbf{y_t}, \psi), \\ \beta_t^{(j)} = \xi_t^{(j)} / \sum_{r=1}^{m} \xi_t^{(r)}, \\ \xi_t = f(\mathbf{y_t}, \phi). \end{cases} \tag{15}$$

In implementation, the algorithm in Table 1 could be adopted except the two terms $\nabla_{\zeta_T} S_p$ and $\nabla_{\xi_T^{(j)}} S_p$ become respectively

$$\nabla_{\zeta_T} S_p = \frac{\left[V(R_T) - M(R_T)\left(R_T - M(R_T) - \frac{1}{T}\sum_{t=1}^{T}(R_t - M(R_t)) \right) \right]}{T\sqrt{[V(R_T)]^3}}$$
$$\cdot \left(\sum_{j=1}^{m} \left[\frac{\xi_T^{(j)} x_T^{(j)}}{\sum_{r=1}^{m} \xi_T^{(r)}} - r_c \left| \frac{\xi_T^{(j)}}{\sum_{r=1}^{m} \xi_T^{(r)}} - \frac{\xi_{T-1}^{(j)}}{\sum_{r=1}^{m} \xi_{T-1}^{(r)}} \right| (1 + x_T^{(j)}) \right] - r^f \right)$$

$$\nabla_{\xi_T^{(j)}} S_p = \frac{\left[V(R_T) - M(R_T)\left(R_T - M(R_T) - \frac{1}{T}\sum_{t=1}^{T}(R_t - M(R_t)) \right) \right] \zeta_T x_T^{(j)} \left(\sum_{r=1}^{m} \xi_T^{(r)} - \xi_T^{(j)} \right)}{T\sqrt{[V(R_T)]^3} \left(\sum_{r=1}^{m} \xi_T^{(r)} \right)^2}$$

Simulation For the purpose of simulation we fix the rate of transaction cost at $r_c = 0.1\%$ and short selling is not applicable to the return-based approach. Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 1(c) while daily risk-return statistics of the portfolios are given in Table 3(c).
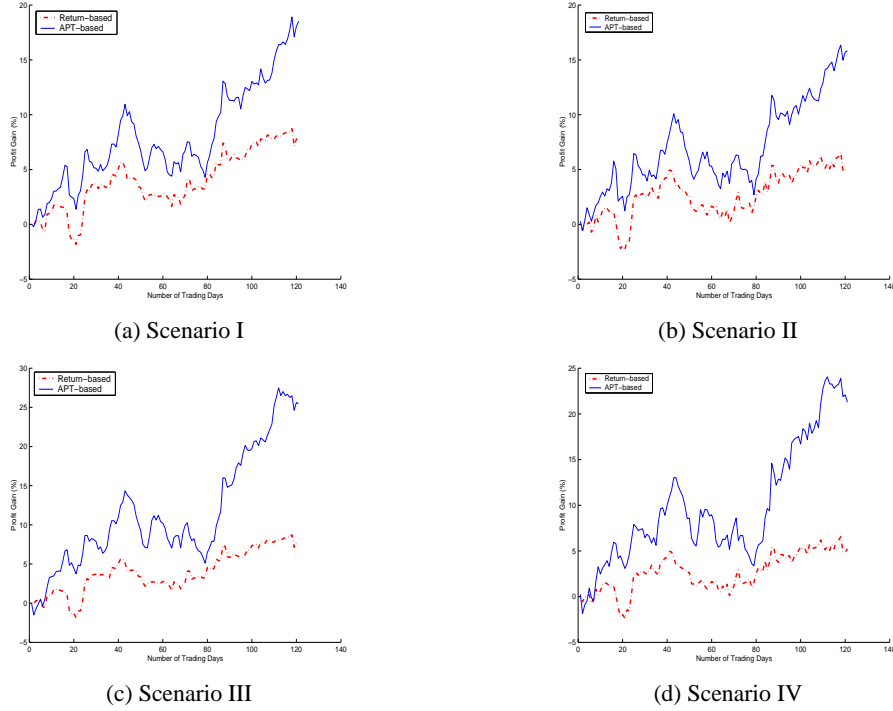

(a) Scenario I

(b) Scenario II

(c) Scenario III

(d) Scenario IV

Fig. 1. Comparative profit gain of APT-based and return-based portfolios for different scenarios

4.5 Scenario IV: HAS Transaction Cost and Short Sale IS Permitted

Scenario IV differs from Scenario I in that the effects of both transaction cost and short sale on portfolio selection have to be treated appropriately. As a result, we have

$$R_t = (1 - \alpha_t)r^f + \alpha_t \sum_{j=1}^{m} \left[\beta_t^{(j)} x_t^{(j)} - r_c |\beta_t^{(j)} - \beta_{t-1}^{(j)}| (1 + x_t^{(j)}) \right], \tag{16}$$

$$\text{subject to} \quad \sum_{j=1}^{m} \beta_t^{(j)} = 1$$

Here we have the objective function same as (15). The APT-based algorithm in Table 1 could still be adopted in this case except that the two terms $\nabla_{\zeta_T} S_p$ and $\nabla_{\xi_T^{(j)}} S_p$ become respectively

$$\nabla_{\zeta_T} S_p = \frac{\left[V(R_T) - M(R_T)\left(R_T - M(R_T) - \frac{1}{T}\sum_{t=1}^{T}(R_t - M(R_t))\right)\right]}{T\sqrt{[V(R_T)]^3}} \left(\frac{\sum_{j=1}^{m} \xi_T^{(j)} x_T^{(j)}}{\sum_{j=1}^{m} \xi_T^{(j)}} - r^f \right),$$

$$\nabla_{\xi_T^{(j)}} S_p = \left[V(R_T) - M(R_T)\left(R_T - M(R_T) - \frac{1}{T}\sum_{t=1}^{T}(R_t - M(R_t))\right)\right]$$
$$\cdot \zeta_T \left[x_T^{(j)} - r_c \text{sign}\left(\xi_T^{(j)} - \xi_T^{(j-1)}\right)\right]$$
$$\cdot \left[\sum_{r=1}^{m} \xi_T^{(r)} - \xi_T^{(j)}\right] \Big/ \left[T\sqrt{[V(R_T)]^3}\left(\sum_{r=1}^{m} \xi_T^{(r)}\right)^2 \right]$$

Table 3. Daily risk-return statistics of the portfolio for different scenarios

(a) Scenario I

	Return-based Portfolio	APT-based Portfolio	Change in S_p
Mean Return	0.06%	0.14%	–
Risk	0.48%	0.81%	–
S_p	0.1250	0.1728	↑ 38.24%

(b) Scenario II

	Return-based Portfolio	APT-based Portfolio	Change in S_p
Mean Return	0.04%	0.12%	–
Risk	0.42%	0.73%	–
S_p	0.0952	0.1644	↑ 72.69%

(c) Scenario III

	Return-based Portfolio	APT-based Portfolio	Change in S_p
Mean Return	0.06%	0.19%	–
Risk	0.48%	0.92%	–
S_p	0.1250	0.2065	↑ 65.20%

(d) Scenario IV

	Return-based Portfolio	APT-based Portfolio	Change in S_p
Mean Return	0.04%	0.16%	–
Risk	0.42%	0.88%	–
S_p	0.0952	0.1818	↑ 90.97%

Simulation In simulation short selling is not applicable to the return-based approach. Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 1(d) while daily risk-return statistics of the portfolios are given in Table 3(d).

4.6 Performance Evaluation

To summarize the experimental results of the above four scenarios, we have noted the following two phenomena. First, the APT-based portfolio in general performs better than the return-based portfolio if the scope of comparison is limited to within each scenario, as evidenced by higher S_p attained in Table 3. It should be noted that higher S_p may arise as a consequence of one of the following situations: i) higher expected return, lower overall volatility; ii) higher expected return, same overall volatility; iii) same expected return, lower overall volatility; iv) both expected return increase or decrease, with expected return increases (decreases) at a faster (lower) rate than overall volatility. Second, if we compare the performance of APT-based portfolios across all the four scenarios, especially the portfolio Sharpe ratio of scenario III against I (↑ 19.50%) and scenario IV against II (↑ 10.58%), we may conclude that performance may be further improved whenever short sale is permitted.

The first phenomenon reveals the fact that independent hidden factors may be more effective in controlling portfolio weights. Possible rationales include dimensionality reduction as

there are usually only a few hidden factors for a large number of securities. What seems to be a more important revelation is that the classical APT [11] model is still helpful here.

Although short selling is expensive for individual investors and not generally permissible for most institutional investors [1] in many markets, relevant experimental results reveal the hypothetical potential benefit such facility might add to the portfolio returns. The benefit mainly arises from the exploitation of downside trend in market price in addition to upward movement. This in turn reduces the chance that the fund is left idle due to declining stock prices for most stocks, which is more or less a phenomenon when the general market atmosphere is gloomy.

5 APT-Based Portfolio Management by Modified Portfolio Sharpe Ratio

In this section, we consider three variants of the portfolio Sharpe Ratio. After the Sharpe ratio was proposed [8] in 1966, many researchers realized that variance was not an appropriate measure of risk since it counted positive returns as risk. Specifically, Fishburn in [18] used the Lower Partial Moment (LPM) of returns called downside risk to replace the traditional variance measure. Similar approach had been extended in [15] to include portfolios. Specifically, we consider portfolio expected downside risk V_T^- which is represented by

$$V_T^- = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^m \sum_{j=1}^m \beta_t^{(i)} \beta_t^{(j)} \int_{-\infty}^0 \int_{-\infty}^0 x_t^{(i)} x_t^{(j)} p(x_t^{(i)}, x_t^{(j)}) dx_t^{(i)} dx_t^{(j)}$$
$$= \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^m \sum_{j=1}^m \beta_t^{(i)} \beta_t^{(j)} D_t^{(i,j)}$$

where β_i and x_i denote the portfolio weight and return of the i-th risky security respectively, and $D_t^{(i,j)}$ is a constant.

In addition to considering the downside risk, the so-called portfolio expected upside volatility V_T^+ can be defined similarly as

$$V_T^+ = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^m \sum_{j=1}^m \beta_t^{(i)} \beta_t^{(j)} \int_0^\infty \int_0^\infty x_t^{(i)} x_t^{(j)} p(x_t^{(i)}, x_t^{(j)}) dx_t^{(i)} dx_t^{(j)}$$
$$= \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^m \sum_{j=1}^m \beta_t^{(i)} \beta_t^{(j)} U_t^{(i,j)}$$

where β_i and x_i denote the portfolio weight and return of the i-th risky security respectively, and $U_t^{(i,j)}$ is a constant.

5.1 Modified Sharpe Ratio with Minimum Downside Risk and Maximum Upside Volatility

Given that portfolio variance can be broken down into portfolio downside risk and upside volatility, it is desirable to consider the maximization of the upside volatility and minimization of the downside risk at the same time in calculating the optimal portfolio. In other words, we can consider maximization of the following improved Sharpe ratio S_p'

$$\max_{\psi,\phi} S_p' = \frac{M(R_T) + V_T^+}{V_T^-} \text{ subject to } \begin{cases} \alpha_t = \exp(\zeta_t), \\ \zeta_t = g(\mathbf{y_t}, \psi), \\ \beta_t^{(j)} = \exp(\xi_t^{(j)}) / \sum_{r=1}^m \exp(\xi_t^{(r)}), \\ \xi_t = f(\mathbf{y_t}, \phi). \end{cases} \quad (17)$$

In implementation, the algorithm in Table 1 could be adopted except the two terms $\nabla_{\zeta_T} S_p$ and $\nabla_{\xi_T^{(j)}} S_p$ become respectively

$$\nabla_{\zeta_T} S'_p = \frac{\sum_{j=1}^{m} \exp{(\xi_T^{(j)})} \left(\sum_{j=1}^{m} \exp{(\xi_T^{(j)})} x_T^{(j)} - r^f \sum_{j=1}^{m} \exp{(\xi_T^{(j)})} \right) \exp{(\zeta_T)}}{\sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \exp{(\xi_T^{(i)})} \exp{(\xi_T^{(j)})} D_t^{(i,j)}},$$

$$\nabla_{\xi_T^{(j)}} S'_p = \left[\sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \exp{(\xi_T^{(i)})} \exp{(\xi_T^{(j)})} D_t^{(i,j)} \left(\exp{(\zeta_T)} x_T^{(j)} + \frac{\sum_{i=1}^{m} \exp{(\xi_T^{(i)})} U_T^{(i,j)}}{\sum_{r=1}^{m} \exp{(\xi_T^{(r)})}} \right) \right.$$

$$\left. - \left(\sum_{t=1}^{T} R_t + \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \exp{(\xi_T^{(i)})} \exp{(\xi_T^{(j)})} U_t^{(i,j)} \right) \frac{\sum_{i=1}^{m} \exp{(\xi_T^{(i)})} D_T^{(i,j)}}{\sum_{r=1}^{m} \exp{(\xi_T^{(r)})}} \right]$$

$$\cdot \left[\left(\sum_{r=1}^{m} \exp{(\xi_T^{(r)})} - \exp{(\xi_T^{(j)})} \right) \exp{(\xi_T^{(j)})} \right]$$

$$\bigg/ \left[\left(\sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \exp{(\xi_T^{(i)})} \exp{(\xi_T^{(j)})} D_t^{(i,j)} \right) \left(\sum_{r=1}^{m} \exp{(\xi_T^{(r)})} \right) \right]^2$$

Simulation We implement the modified Sharpe ratio simulation using the same set of data described before and the APT-based approach in Scenario I as benchmark for comparison. Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 2(a) while daily risk-return statistics of the portfolios are given in Table 4.

Table 4. Daily Risk-return statistics of portfolio under original and modified Sharpe ratio

	Original Sharpe ratio	Modified Sharpe ratio
Mean Return	0.14%	0.24%
Risk	0.81%	1.13%
Upside Volatility	–	0.43%
Downside Risk	–	0.35%
S_p	0.1728	1.9143

5.2 Risk Minimization with Control of Expected Return

Some conservative investors are more concerned about risk than return. Therefore, a more appropriate investment strategy may be to minimize risk while controlling the expected return. Particularly, this can be achieved by setting the expected return in (25) to be a constant specified by the investor and the optimization essentially becomes a minimization of downside risk and a maximization of upside volatility.

$$\max_{\psi, \phi} S'_p = \frac{r + V_T^+}{V_T^-} \quad \text{subject to} \begin{cases} \alpha_t = \exp{(\zeta_t)}, \\ \zeta_t = g(\mathbf{y_t}, \psi), \\ \beta_t^{(j)} = \exp{(\xi_t^{(j)})} / \sum_{r=1}^{m} \exp{(\xi_t^{(r)})}, \\ \xi_t = f(\mathbf{y_t}, \phi), \\ M(R_T) = r. \end{cases} \quad (18)$$

To solve the above optimization problem with equality constraints, we adopt the augmented Lagrangian method. Specifically, for the equality constrained problem

$$\begin{aligned} \text{maximize} \quad & f(\mathbf{x}) \quad \text{with respect to } \mathbf{x} \\ \text{subject to} \quad & h(\mathbf{x}) = 0, \end{aligned} \quad (19)$$

the augmented Lagrangian function can be written as

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda h(\mathbf{x}) - \frac{1}{2}c[h(\mathbf{x})]^2 \tag{20}$$

where λ is the Lagrange multiplier, c is the penalty parameter. Then a sequence of minimizations of the form

$$\begin{aligned} \text{maximize} \quad & L_{c_k}(\mathbf{x}, \lambda_k) \quad \text{with respect to } \mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in \mathbb{R}^n \end{aligned} \tag{21}$$

is performed where $\{c_k\}$ is a sequence of positive penalty parameters sequence with

$$\begin{aligned} 0 < c_k < c_{k+1} \quad & \forall k \\ c_k \to \infty \quad & \text{as } k \to \infty. \end{aligned} \tag{22}$$

The multiplier sequence $\{\lambda_k\}$ is generated by the iteration

$$\{\lambda_{k+1}\} = \lambda_k + c_k h(\hat{\mathbf{x}}) \tag{23}$$

where $\hat{\mathbf{x}}$ is the solution of (21).

Here the augmented Lagrangian is given by

$$L = \frac{r + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{m}\beta_t^{(i)}\beta_t^{(j)}U_t^{(i,j)}}{\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{m}\beta_t^{(i)}\beta_t^{(j)}D_t^{(i,j)}} - \lambda\left(\frac{1}{T}\sum_{t=1}^{T}R_t - r\right) - \frac{c}{2}\left(\frac{1}{T}\sum_{t=1}^{T}R_t - r\right)^2 \tag{24}$$

In implementation, the algorithm in Table 1 could be adopted except the two terms $\nabla_{\zeta_T}S_p$ and $\nabla_{\xi_T^{(j)}}S_p$ are replaced by $\nabla_{\zeta_T}L$ and $\nabla_{\xi_T^{(j)}}L$ respectively, where

$$\nabla_{\zeta_T}L = \frac{\exp(\zeta_T)}{T}\left(r^f - \frac{\sum_{r=1}^{m}\exp(\xi_T^{(r)})x_T^{(r)}}{\sum_{r=1}^{m}\exp(\xi_T^{(r)})}\right)\left(\lambda + c\big(M(R_T) - r\big)\right)$$

$$\begin{aligned} \nabla_{\xi_T^{(j)}}L = &\left[\left(\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{m}\exp(\xi_T^{(i)})\exp(\xi_T^{(j)})D_t^{(i,j)}\left(\frac{\sum_{i=1}^{m}\exp(\xi_T^{(i)})U_T^{(i,j)}}{\sum_{r=1}^{m}\exp(\xi_T^{(r)})}\right)\right.\right. \\ &\left.- \left(\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{m}\exp(\xi_T^{(i)})\exp(\xi_T^{(j)})U_t^{(i,j)}\right)\frac{\sum_{i=1}^{m}\exp(\xi_T^{(i)})D_T^{(i,j)}}{\sum_{r=1}^{m}\exp(\xi_T^{(r)})}\right) \\ &\left./\left(\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{m}\exp(\xi_T^{(i)})\exp(\xi_T^{(j)})D_t^{(i,j)}\right)^2\right. \\ &\left.- \frac{\exp(\zeta_T)x_T^{(j)}}{T}\left(\lambda + c\big(M(R_T) - r\big)\right)\right] \cdot \frac{\left(\sum_{r=1}^{m}\exp(\xi_T^{(r)}) - \exp(\xi_T^{(j)})\right)\exp(\xi_T^{(j)})}{\left[\sum_{r=1}^{m}\exp(\xi_T^{(r)})\right]^2} \end{aligned}$$

Simulation We simulate the modified Sharpe ratio with control of expected return approach and use the modified Sharpe ratio approach in the previous subsection as benchmark. The predetermined expected return used for the simulation is $r = 0.15\%$. Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 2(b) while daily risk-return statistics of the portfolios are given in Table 5.

5.3 Return Maximization with Control of Expected Downside Risk

Some aggressive investors are more concerned about return than risk. Therefore, a strategy that may better serve them is to maximize the expected return while controlling the expected
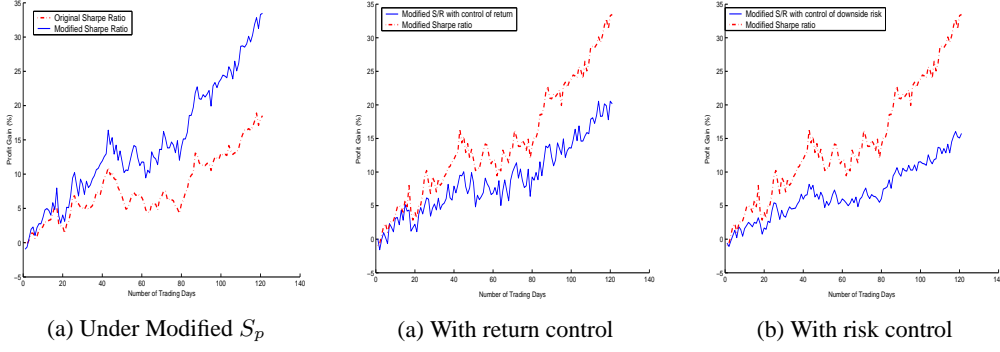
(a) Under Modified S_p (a) With return control   (b) With risk control

Fig. 2. Comparative profit gain of portfolio under modified Sharpe ratio and with control of expected return and risk

Table 5. Risk-return statistics of portfolio with control of expected return

	Modified Sharpe ratio	Modified Sharpe ratio with control of expected return
Mean Return	0.24%	0.17%
Risk	1.13%	0.79%
Upside Volatility	0.43%	0.30%
Downside Risk	0.35%	0.23%
S_p	1.9143	2.0435

downside risk. In particular, this can be achieved by setting the expected downside risk in (25) to be a constant specified by the investor and the optimization essentially becomes a maximization of expected return and upside volatility.

$$
\max_{\psi,\phi} S_p' = \frac{M(R_T) + V_T^+}{v} \ \text{ subject to } \begin{cases} \alpha_t = \exp\left(\zeta_t\right), \\ \zeta_t = g(\mathbf{y_t}, \psi), \\ \beta_t^{(j)} = \exp{(\xi_t^{(j)})} / \sum_{r=1}^m \exp{(\xi_t^{(r)})}, \\ \xi_t = f(\mathbf{y_t}, \phi), \\ V_T^- = v. \end{cases} \quad (25)
$$

Here the augmented Lagrangian is given by

$$
L = \frac{1}{Tv} \Big(\sum_{t=1}^{T} R_t + \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \beta_t^{(i)} \beta_t^{(j)} U_t^{(i,j)} \Big) - \lambda \Big(\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \beta_t^{(i)} \beta_t^{(j)} D_t^{(i,j)} - v \Big)
$$
$$
- \frac{c}{2} \Big(\frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{m} \beta_t^{(i)} \beta_t^{(j)} D_t^{(i,j)} - v \Big)^2 \quad (26)
$$

In implementation, the algorithm in Table 1 could be adopted except the two terms $\nabla_{\zeta_T} S_p$ and $\nabla_{\xi_T^{(j)}} S_p$ are replaced by $\nabla_{\zeta_T} L$ and $\nabla_{\xi_T^{(j)}} L$ respectively, where

$$\nabla_{\zeta_T} L = \frac{\exp\left(\zeta_T\right)}{Tv} \left(\frac{\sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right) x_T^{(r)}}{\sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right)} - r^f \right)$$

$$\nabla_{\xi_T^{(j)}} L = \left[\frac{\exp\left(\zeta_T\right) Tv \sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right) x_T^{(j)} + \sum_{i=1}^{m} \exp\left(\xi_T^{(i)}\right) U_T^{(i,j)}}{Tv \sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right)} \right.$$
$$\left. - \frac{\sum_{i=1}^{m} \exp\left(\xi_T^{(i)}\right) D_T^{(i,j)}}{T \sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right)} \left(\lambda + c\left(V_T^{-} - v\right) \right) \right] \cdot \frac{\left(\sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right) - \exp\left(\xi_T^{(j)}\right) \right) \exp\left(\xi_T^{(j)}\right)}{\left[\sum_{r=1}^{m} \exp\left(\xi_T^{(r)}\right) \right]^2}$$

Simulation We simulate the modified Sharpe ratio with control of expected downside risk approach and use the modified Sharpe ratio approach in the previous subsection as benchmark. The predetermined expected downside risk used for the simulation is $v = 0.20\%$. Graphical comparison of profit gain between the two approaches using test data is shown in Fig. 2(c) while daily risk-return statistics of the portfolios are given in Table 6.

Table 6. Risk-return statistics of portfolio with control of expected downside risk

	Modified Sharpe ratio	Modified Sharpe ratio with control of downside risk
Mean Return	0.24%	0.15%
Risk	1.13%	0.71%
Upside Volatility	0.43%	0.23%
Downside Risk	0.35%	0.19%
S_p	1.9143	0.2000

5.4 Performance Evaluation

The investment strategy with control of expected return is well-suited for risk-averse investors. By comparing the statistics shown in Table 5 with that of Table 4, we can see that not only is the expected return under control, but also is risk lowered. As a result, S_p remains more or less constant. The phenomenon is consistent with the intuition that risk and return go hand-in-hand in the same direction due to the nonexistence of free lunch. Similar reasoning could also be extended to include the case of aggressive profit-seeking investors, by comparing the statistics shown in Table 6 with that of Table 4.

6 Conclusion

In this paper, we introduce how the gaussian TFA model originally devoted for further study of the APT could be utilized for adaptive portfolio management. Furthermore, we study three variants of the APT-based Sharpe ratio maximization technique that utilize the concept of portfolio downside risk and upside volatility tailored-made for the special need of risk-averse as well as aggressive profit-seeking investors. Simulation results reveal that APT-based portfolio management in general excels return-based portfolio management and portfolio returns may

be somehow enhanced by short selling, especially when the general market climate is not that favorable.

References

1. Chan, L., K.J., Lakonishok, J.: On portfolio optimization: Forecasting covariances and choosing the risk model. The Review of Financial Studies **12** (1999) 937–974
2. Haugh, M., Lo, A.: Computational challenges in portfolio management. Computing in Science & Engineering **3** (2001) 54–59
3. Choey, M., Weigend, A.S.: Nonlinear trading models through sharpe ratio optimization. International Journal of Neural Systems **8** (1997) 417–431
4. Moody, J., Wu, L.: Optimization of trading systems and portfolios. Proc. of Computational Intelligence for Financial Engineering (CIFEr'97) (1997) 300–307
5. Xu, L., Cheung, Y.M.: Adaptive supervised learning decision networks for traders and portfolios. Proc. of Computational Intelligence for Financial Engineering (CIFEr'97) (1997) 206–212
6. Neuneier, R.: Optimal asset allocation using adaptive dynamic programming. Advances in Neural Information Processing Systems **8** (1996) 952–958
7. Moody, J., Wu, L., Liao, Y., Saffell, M.: Performance functions and reinforcement learning for trading systems and portfolios. Journal of Forecasting **17** (1998) 441–470
8. Sharpe, W.F.: Mutual fund performance. Journal of Business **39** (1966) 119–138
9. Sharpe, W.F.: The sharpe ratio - properly used, it can improve investment. Journal of Portfolio Management (1994) 49–58
10. Xu, L.: Temporal byy learning for state space approach, hidden markov model and blind source separation. IEEE Trans. on Signal Processing **48** (2000) 2132–2144
11. Ross, S.: The arbitrage theory of capital asset pricing. Journal of Economic Theory **13** (1976) 341–360
12. Roll, R., Ross, S.: An empirical investigation of the arbitrage pricing theory. Journal of Finance **35** (1980) 1073–1103
13. Roll, R., Ross, S.: The arbitrage pricing theory approach to strategic portfolio planning. Financial Analysts Journal **40** (1984) 14–26
14. Xu, L.: Byy harmony learning, independent state space and generalized apt financial analyses. IEEE Transactions on Neural Networks **12** (2001) 822–849
15. Hung, K.K., Cheung, C.C., Xu, L.: New sharpe-ratio-related methods for portfolio selection. Proc. of Computational Intelligence for Financial Engineering (CIFEr 2000) (2000) 34–37
16. Xu, L.: Rbf nets, mixture experts, and bayesian ying-yang learning. Neurocomputing **19** (1998) 223–257
17. Chiu, K.C., Xu, L.: Financial apt-based gaussian tfa learning for adaptive portfolio management. in J.R. Dorronsoro (Ed.): Artificial Neural Networks- ICANN'2002, LNCS 2415 (2002) 1019–1024
18. Fishburn, P.C.: Mean-risk analysis with risk associated with below target returns. The American Economic Review **67** (1977) 116–126