

Multiperiod portfolio investment using stochastic programming with conditional value at risk

Hung-Hsin Chen, Chang-Biau Yang*

Department of Computer Science and Engineering, National Sun Yat-sen University, 70 Lienhai Rd., Kaohsiung 80424, Taiwan



ARTICLE INFO

Article history:

Received 18 March 2015

Revised 29 September 2016

Accepted 9 November 2016

Available online 16 November 2016

Keywords:

Multiperiod portfolio investment

Stochastic programming

Conditional value at risk

Moment matching

Superior predictive ability

ABSTRACT

This paper proposes the portfolio stochastic programming (PSP) model and the stagewise portfolio stochastic programming (SPSP) model for investing in stocks in the Taiwan stock market. The SPSP model effectively reduces the computational resources needed to solve the PSP model. Additionally, the conditional value at risk (CVaR) is used as a risk measure in the models. In each period of investment, 200 scenarios are generated to solve the SPSP model. The experimental data set consists of the 50 listed companies with the greatest market capitalization in the Taiwan stock exchange, and the experimental interval began on January 3, 2005 and ended on December 31, 2014, consisting of 2484 trading periods (days) in total. The experimental results show that the SPSP model is insensitive to small variation of the portfolio size and the historical period for estimating statistics. The portfolio size of the SPSP model can be set with two cases: $M = M_c$ and $M \leq M_c$. When $M = M_c$, the M invested target stocks have been predetermined. When $M \leq M_c$, a set of M_c candidate stocks are given, but the M real target stocks have not been decided. The average annualized returns are 13.09% and 12.06% for the two portfolio settings, respectively, which are higher than that of the buy-and-hold (BAH) rule (9.95%). In addition, because the CVaR is considered, both portfolio settings of the SPSP model exhibit higher Sharpe and Sortino ratios than the BAH rule, indicating that the SPSP model provides a higher probability to earn a positive return. The superior predictive ability test is performed to illustrate that the SPSP model can avoid the data-snooping problem.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

This paper proposes the portfolio stochastic programming (PSP) model and the stagewise portfolio stochastic programming (SPSP) model for guiding an investor in allocating assets appropriately to earn higher return with risk management in each period. The portfolio problem is crucial in investment. A rational investor prefers to attain the highest possible return with the lowest possible risk. However, high risk often accompanies an investment yielding a high return. Markowitz [1] proposed the mean-variance model, which serves as the foundation of portfolio theory.

Precisely evaluating the future performance of an asset is difficult, because many uncertain parameters are involved in modeling the problem. Making appropriate decisions and earning a positive return typically require mathematical analysis. To improve decision quality, many researchers have studied mathematical models for overcoming uncertainty. Ben-Tal et al. [2] proposed a robust multi-stage model for solving the portfolio problem. Li and Ng [3] solved

a multistage mean-variance model and provided an analytic solution. Bertsimas and Pachamanova [4] proposed a robust model of multistage portfolio that incorporated transaction costs. Topaloglou et al. [5] used the *conditional value at risk* (CVaR) to control risk in investment, and they used *stochastic programming* (SP) with the CVaR [6] to solve the international asset allocation problem. These researchers mainly developed theoretical models; however, more practical models are required for efficient investment.

The purpose of risk management is to avoid a portfolio that may suffer large losses in investment. The Basel Committee on Banking Supervision recommended using the *value at risk* (VaR) as the basis for modeling market risk. However, the VaR is not a coherent risk measure, and it does not satisfy the subadditivity axiom [7], meaning that the VaR is inconsistent with the financial principle that diversification reduces risk. The CVaR [8] is a coherent risk measure that represents the expected loss below the VaR. Rockafellar and Uryasev [9] introduced the CVaR for a continuous distribution and defined it for a general distribution [10]. Topaloglou et al. [6] applied multistage SP with the CVaR to an international portfolio.

In this paper, we describe the portfolio management problem of multiperiod active stock investment, and then propose the

* Corresponding author.

E-mail address: cbyang@cse.nsysu.edu.tw (C.-B. Yang).

portfolio stochastic programming (PSP) model and the stagewise portfolio stochastic programming (SPSP) model for solving this problem. In multiperiod investment, allocation decisions should be rebalanced in each period. Transaction fees substantially influence investment returns because frequent transactions result in high transaction costs and potentially increase risk exposure. SP is a method for modeling optimization problems involving uncertainty. Deterministic optimization problems are typically formulated by using known parameters [11]. An investor may not always predict the future return of a stock accurately, but may predict the future return with a confidence level. Input parameters denoting uncertainty in SP can be represented by discrete scenarios that describe the joint distribution of random variables, and the expected cost and constraints of these scenarios are added into the model. After the scenarios are generated, the uncertainty problem is transformed into an equivalent deterministic one that can be solved efficiently.

In this paper, the investment is confined to a market consisting of the 50 listed companies with the greatest market capitalization in the Taiwan stock exchange on January 3, 2005, and the goal is to earn a positive return. The SPSP model employs the CVaR to control the investment risk. In addition, the *superior predictive ability* (SPA) test [12] is employed in the daily return process to verify that our models have no data-snooping problems. The experimental interval is from January 3, 2005 to December 31, 2014, which consists of 2484 trading periods (days) in total. For the SPSP model in which the invested target stocks have been predetermined ($M = M_c$), this leads to linear programming. The highest average cumulative return and annualized return are 242.20% and 13.09%, respectively. For the SPSP model in which a set of possible candidate stocks have been determined, but the real target stocks have not been settled ($M \leq M_c$), this leads to mixed integer linear programming. The highest average cumulative return and annualized return are 212.34% and 12.06%, respectively. Both settings of the SPSP model yield higher returns and higher Sharpe and Sortino ratios than the *buy-and-hold* (BAH) rule. This means that our model provides a higher probability of earning positive returns in the future.

The remainder of this paper is organized as follows. Section 2 presents background knowledge, including the SP, CVaR, and SPA test. Section 3 proposes the PSP and SPSP models, which can earn positive returns in the Taiwan stock market. Section 4 presents the experimental results of the SPSP model. Section 5 provides the conclusion of this paper.

2. Preliminaries

This section introduces the SP, the value of information, the CVaR measure, and the data-snooping problem associated with the SPA test.

2.1. Stochastic programming

Modeling uncertainty is widely recognized as an integral component in most real-world decision problems. SP provides an appropriate modeling framework for decision-making after uncertainty has been identified. Decisions have to be made at the moment, but crucial information is not available until the decision has been made [13]. A major advantage of SP is that there is no requirement regarding the assumption of the distributions of random variables.

Determining how to describe uncertainty is a critical task in modeling a problem. When an SP problem is solved, uncertain events are represented by discrete distributions. In a *multiperiod* problem, discrete random variables are typically represented in the form of a tree. A *scenario* is a path from the root to a leaf node;

a *stage* is the moment when a decision is made, and it constitutes a level of the tree; and a *period* is an interval between two time points or the edge length between two neighboring nodes. For example, in the two-stage SP investment model, an investor desires to estimate the joint distribution of M assets returns. Therefore, the scenario tree has only two levels, one root and S leaves, where each leaf node represents a scenario. When the scenario tree is constructed, the S scenarios are sampled simultaneously, and a scenario is an M -dimensional vector in which each element corresponds to a realized return of one asset. A scenario tree consisting of these S scenarios is an approximation of the joint distribution of the risky assets returns. The goal is to create a scenario tree in which the error caused by approximation is minimized and the solvability of the model is retained.

The basic principle of the two-stage SP model is that decisions should be made according to the data available at the time and should not depend on future observations; this is called the nonanticipativity principle [14,15]. The two-stage linear SP model with fixed recourse, called the *recourse problem*, is defined in (1) [14].

$$\begin{aligned} RP = \max & E_{\xi}(g(\mathbf{x}, \xi(\omega))) = \mathbf{c}^T \mathbf{x} + E_{\xi}(\max \{\mathbf{q}(\omega)^T \mathbf{y}(\omega) | \\ & \mathbf{W}\mathbf{y}(\omega) \leq \mathbf{h}(\omega) + \mathbf{T}(\omega)\mathbf{x}, \mathbf{y}(\omega) \geq \mathbf{0}\}), \\ \text{s.t. } & \mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, \end{aligned} \quad (1)$$

where RP is the optimal solution of the SP model; $g(\mathbf{x}, \xi(\omega))$ is a real-valued objective function, which is the sum of the first-stage profit $\mathbf{c}^T \mathbf{x}$ and the optimal second-stage profit $\max \{\mathbf{q}(\omega)^T \mathbf{y}(\omega) | \mathbf{W}\mathbf{y}(\omega) \leq \mathbf{h}(\omega) + \mathbf{T}(\omega)\mathbf{x}, \mathbf{y}(\omega) \geq \mathbf{0}\}$ by given \mathbf{x} and ξ ; $E_{\xi}(\cdot)$ is the expectation with respect to a random vector ξ ; $\xi(\omega) = (\mathbf{q}(\omega), \mathbf{h}(\omega), \mathbf{T}(\omega))$, $\omega \in \Omega$ is a $(m_2 + n_2 + n_2 \times n_1)$ -dimensional random data vector, and Ω is the uncertain data set in the second stage; \mathbf{x} is the m_1 -dimensional decision vector of the first stage; \mathbf{c} is the known m_1 -dimensional coefficient vector of \mathbf{x} ; \mathbf{A} and \mathbf{b} are the known $n_1 \times m_1$ resource matrix and n_1 -dimensional vector of the first stage, respectively; the constraints $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ specify the feasible domain of \mathbf{x} . In the second stage, \mathbf{W} is the known $n_2 \times m_2$ recourse matrix, which is assumed to be fixed here; a number of random events $\omega \in \Omega$ are realized in the second stage. For a given realized event ω , $\mathbf{y}(\omega)$ is the m_2 -dimensional decision vector, and the corresponding second-stage data, $\mathbf{q}(\omega)$, is the m_2 -dimensional coefficient vector of $\mathbf{y}(\omega)$; the n_2 -dimensional resource vector $\mathbf{h}(\omega)$ and the $n_2 \times n_1$ technological matrix $\mathbf{T}(\omega)$ become known. The second-stage constraints $\mathbf{W}\mathbf{y}(\omega) \leq \mathbf{h}(\omega) + \mathbf{T}(\omega)\mathbf{x}$ and $\mathbf{y}(\omega) \geq \mathbf{0}$ specify the feasible domain of $\mathbf{y}(\omega)$.

A decision about vector \mathbf{x} of (1) has to be made *here-and-now* before a realization of the corresponding random data vector ξ becomes known. After a realization of ξ is available, an optimal decision about \mathbf{y} is made by solving the second stage optimization problem, which depends on the first stage decision vector \mathbf{x} and data vector ξ .

A perfect information solution would choose the optimal decisions \mathbf{x} of the first stage for each realization of ξ . For each scenario, once the SP is solved, the corresponding decision vector \mathbf{x} is obtained by foreseeing the scenario which has not happened yet. The expected value of this solution is known as the *wait-and-see* solution as defined in (2) [14].

$$WS = E_{\xi}(\max g(\mathbf{x}, \xi)), \quad (2)$$

where WS is the optimal solution of the perfect information model.

To reduce the large number of possible realizations of scenarios and the number of constraints of the recourse problem described in (1), each random vector $\xi(\omega)$ is replaced by its expected vector $E_{\xi}(\xi)$. The deterministic expected value problem is often solved as an approximating solution of the recourse problem. The model is

defined in (3) [14].

$$EV = \max g(\mathbf{x}, E_{\xi}(\xi)), \quad (3)$$

where EV is the optimal solution of the expected value model.

Let $\bar{\xi} = E_{\xi}(\xi)$. The expected value model yields the first-stage and second-stage solutions. When the first-stage decision vector $\mathbf{x}(\bar{\xi})$ is fixed, $\mathbf{x}(\bar{\xi})$ becomes the parameter of the second stage. A second optimization should then be performed after the random vector ξ is realized. The expected result EEV , obtained from the expected value model of the first-stage decision vector $\mathbf{x}(\bar{\xi})$, is defined in (4) [14].

$$EEV = E_{\xi}(g(\mathbf{x}(\bar{\xi}), \xi)). \quad (4)$$

Madansky [16] established the inequalities for general SP among the optimal solutions RP , WS , and EEV , which are expressed in (5).

$$EEV \leq RP \leq WS. \quad (5)$$

For SP with fixed objective coefficient vector \mathbf{q} and fixed recourse matrix \mathbf{W} , the inequality between the EV and WS as expressed in (6) can be established.

$$WS \leq EV. \quad (6)$$

2.2. Value of information

Analysis of the effect of uncertainty in the SP generally concentrates on the *expected value of perfect information* (EVPI), which is the difference between the solutions of wait-and-see and the recourse problem method. The EVPI is defined in (7) [14].

$$EVPI = WS - RP. \quad (7)$$

From (5), we immediately find that $EVPI \geq 0$. $EVPI$ represents the substantial profit that a decision maker should minimize to gain more accurate information about the future. A small $EVPI$ indicates a low additional profit when we obtain perfect information. To compute the $EVPI$, we should know the real distribution of all random variables in the model in advance. However, this is almost impossible in general applications.

When no more information about the future can be found, the *value of the stochastic solution* (VSS) may be more useful. VSS is defined as the difference between the solutions of the recourse problem and the expected value of the expected value model, given in (8) [14].

$$VSS = RP - EV. \quad (8)$$

From (5), it is clear that $VSS \geq 0$. The VSS represents substantial profit that a decision maker can obtain using a more complicated model for solving a problem as opposed to a reduced model. A small VSS means that the approximation of SP by the program using the expected value instead of a random variable is a good model.

One common approach for solving an SP problem is to assume that the uncertain data ξ can be modeled as a random vector with a known probability distribution, and the distribution can be approximated by a finite number of possible samples or realizations, which are called *scenarios* and denoted as ξ_1, \dots, ξ_S , with respective probabilities p_1, \dots, p_S , where $\sum_{s=1}^S p_s = 1$. Therefore, the recourse problem can be rewritten as (9) [13].

$$\begin{aligned} \max \quad & \mathbf{c}^T \mathbf{x} + \sum_{s=1}^S p_s \mathbf{q}(\omega_s)^T \mathbf{y}(\omega_s) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \preceq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}, \\ & \mathbf{W}\mathbf{y}(\omega_s) \preceq \mathbf{h}(\omega_s) + \mathbf{T}(\omega_s)\mathbf{x}, \quad s = 1, 2, \dots, S, \\ & \mathbf{y}(\omega_s) \geq \mathbf{0}, \quad s = 1, 2, \dots, S, \end{aligned} \quad (9)$$

where $\xi_s = (\mathbf{q}(\omega_s), \mathbf{h}(\omega_s), \mathbf{T}(\omega_s))$ is the s th scenario, and S is the number of scenarios.

After S scenarios have been generated, the original non-deterministic SP can be transformed into a deterministic program, which can typically be solved efficiently. When some of the decision variables must be of integer values, (1) becomes a *stochastic mixed integer programming* (SMIP). An SMIP generally lacks the convexity properties required for applying optimization techniques, and the programming may be NP-hard. Although the SMIP is difficult to solve, the exact solutions of small-scale SMIP can still be obtained by using the branch-and-bound algorithm, dynamic programming, or the cutting plane method, among others.

An SP can be used to manage only discrete samples of limited sizes; hence, the target distribution should be approximated. Before scenarios are generated, the lengths of time periods and number of stages should be determined because not all information is known at the current time. In addition, the number of scenarios to be generated should be determined. In general, the greater the number of scenarios generated is, the more accurate the approximation is, but more computational resources are required. The distribution of a random variable is typically estimated based on historical data, statistical models, expert opinions, or a combination of these information sources. Selecting suitable scenario generation is problem dependent; no single approach to scenario generation is suitable for all types of problems.

Pflug [17] discussed multiperiod scenario properties in the financial market, and proved that the smaller the error between generated scenarios and the real distribution is, the more accurate the optimal values of an SP model are. Høyland and Wallace [18] proposed an optimization model for generating multiperiod scenarios. The model accepts arbitrary statistics of the target random variables. However, the model may lack convexities, and the scenarios could not be generated efficiently. Høyland et al. [19] proposed the *heuristic moment matching* (HMM) method for generating scenarios with given statistics, including the mean, standard deviation, skewness, kurtosis, and correlation matrix. This method is one of the most efficient methods for generating scenarios. Thus, we will apply this method in this paper. Heitsch and Römisch [20] proposed forward and backward algorithms for tree generation, which consists of recursive scenario reduction and bundling steps. The algorithms require discrete approximation of the target process, and the scenarios tree is generated within a given error upper bound. Beraldi et al. [21] proposed a parallel scenario generation method based on [18,19]. Kaut and Wallace [22] criticized the linear correlation in moment matching, and suggested replacing this correlation with a copula function. Kaut [23] proposed the heuristic copula method for approximating the mixed integer program to generate scenarios, and the method requires the marginal distribution of each random variable.

A multistage linear SP can be solved by using dynamic programming, a decomposition algorithm [24], and a progressive hedging algorithm [15]. But there are only a few empirical studies of these models, especially in a long-term problem because of the bloating scenario tree problem. The classical scenario tree branches at each stage. Even with a small number of realizations per stage, the size of the tree grows exponentially with respect to the number of stages. Therefore, solving a long-term problem with exponential growth of the scenarios is intractable.

There are two common methods for reducing scenarios in the multistage linear SP, the expected value and the event spike methods [25]. The expected value method uses the single expected scenario in each stage, and the event spike method uses the scenarios based on the expected decision of the previous stage in the current stage to get approximate solutions of the multistage linear programming. The expected value method largely reduces the scenarios, but the method results in poor approximation of the

target distribution. The event spike method keeps several scenarios in each stage, and the method usually gets better approximation than the expected value method. However, in a long-term programming, the event spike method still requires huge computational resources to get a solution.

In this paper, we generate the scenarios and solve the multi-stage linear SP in stages. The solutions of period $t - 1$ are used as the inputs of period t , so the problem of period t can be viewed as a two-stage SP. This method satisfies the nonanticipativity principle and controls the approximation with reasonable error bounds by using small computational resources.

2.3. Conditional value at risk

In the current financial industry, the VaR is a widely used loss measure for a specific portfolio of risky assets, and it has become the financial standard for measuring risk. The VaR is utilized to estimate the volatility, premium of portfolio returns, and possible loss in investment [26]. The definition of the percentile of a random variable and VaR is provided in (10).

$$VaR(x, \alpha) = \inf\{u | P(x \leq u) \geq 1 - \alpha, \quad 0 \leq \alpha < 1\}, \quad (10)$$

where x is a random variable to be measured, α is a confidence level set by the investor, and $P(\cdot)$ is the probability measure.

By the definition, the VaR is a percentile-based measure that is typically defined as the maximally allowable loss at a certain confidence level $\alpha \times 100\%$. However, the VaR is not a coherent risk measure because it does not satisfy the subadditivity axiom [7], which is inconsistent with the financial principle that diversification facilitates reducing risk.

The CVaR is an improved coherent risk measure that accounts for the expected loss below the VaR [8]. Rockafellar and Uryasev [9] introduced the CVaR for a continuous distribution as shown in (11).

$$CVaR_c(x, \alpha) = E(u | u \leq VaR(x, \alpha)). \quad (11)$$

Rockafellar and Uryasev [10] further defined the CVaR for a general distribution as shown in (12).

$$CVaR(x, \alpha) = \left(1 - \frac{1}{1 - \alpha} \sum_{\{s \in \Omega | x_s \leq z\}} p_s\right)z + \frac{1}{1 - \alpha} \sum_{\{s \in \Omega | x_s \leq z\}} p_s x_s, \quad (12)$$

where $z = VaR(x, \alpha)$, x_s is the realized value of x in the s th scenario, p_s is the probability of the s th scenario, and Ω represents the set of all scenarios.

Topaloglou et al. [5] reformulated the CVaR as shown in (13).

$$\begin{aligned} y_s &= \max(0, z - x_s), \\ z - \frac{1}{1 - \alpha} \sum_{s \in \Omega} p_s y_s &= \left(1 - \frac{1}{1 - \alpha} \sum_{\{s \in \Omega | x_s \leq z\}} p_s\right)z \\ &\quad + \frac{1}{1 - \alpha} \sum_{\{s \in \Omega | x_s \leq z\}} p_s x_s. \end{aligned} \quad (13)$$

Rockafellar and Uryasev [9] showed that the CVaR can be computed by optimizing the linear programming with approximation of the distribution by using a finite number of scenarios. The linear programming for computing the CVaR measure at a confidence level α is shown in (14).

$$\begin{aligned} \max & \quad z - \frac{1}{1 - \alpha} \sum_{s=1}^S p_s y_s \\ \text{s.t.} & \quad y_s \geq z - x_s, \quad y_s \geq 0, \quad s = 1, 2, \dots, S. \end{aligned} \quad (14)$$

2.4. Superior predictive ability test

Data snooping, which is also known as data dredging, or equation fitting, is the use of data mining to uncover relationships in data. Data snooping occurs when a given set of data is used repeatedly for inference or model selection. Many studies have applied statistical and machine learning models to fitting data used for prediction. When such data reuse occurs often, any satisfactory result obtained may simply be due to chance rather than to merit inherent in the method yielding the results [27]. This problem is practically unavoidable in analyzing the data of time series, where only one observation of interest is typically available for analysis. White [27] proposed the *reality check* (RC) test to address the data-snooping problem. Hansen [12] proposed the *superior predictive ability* (SPA) test, thereby improving the power of the RC test.

Given K alternative models with T periods, let $d_{k,t}$ represent the performance measure of model k related to the benchmark model in period t . The null hypothesis, explained in (15), is used to determine whether there is one alternative model superior to the benchmark model. When no model is superior to the benchmark model, the expected value of the relative performance $d_{k,t}$ should be less than or equal to zero.

$$H_0 : \max_{1 \leq k \leq K} \mu_k \leq 0, \quad (15)$$

where μ_k represents the expected value of $d_{k,t}$, $1 \leq k \leq K$.

The SPA statistic [12] is expressed as follows.

$$T^{SPA} = \max \left(\max_{1 \leq k \leq K} \sqrt{T} \frac{\bar{d}_k}{\hat{\omega}_k}, 0 \right), \quad (16)$$

where $\bar{d}_k = \sum_{t=1}^T d_{k,t}/T$, and $\hat{\omega}_k$ is a consistent estimator of the standard deviation of the relative performance of the k th model.

When the performance of most alternative models is inferior to that of the benchmark model, a decrease in the power of the test should be avoided. Hence, in (16), the SPA statistic is set to zero if the value of every $\sqrt{T} \frac{\bar{d}_k}{\hat{\omega}_k}$ is negative.

The mean sample relative performance in the SPA test is expressed in (17) [12].

$$\hat{\mu}_k = \bar{d}_k \times 1 \left(\sqrt{T} \bar{d}_k \leq -\hat{\omega}_k \sqrt{2 \log \log T} \right), \quad (17)$$

where $1(\cdot)$ is the indicator function.

After B times of bootstrap sampling for computing the statistics T^{SPA} , these sample statistics $T_1^{SPA}, T_2^{SPA}, \dots, T_B^{SPA}$ form an empirical distribution, and the P value of the SPA test is equal to $\sum_{b=1}^B 1(T_b^{SPA} > T^{SPA})/B$ [12].

In this paper, an alternative model could be daily return series of a stock or daily portfolio return series, and the benchmark model is a zero return series. The relative performance measure for the SPA test is expressed as follows.

$$\bar{d}_k = \frac{1}{T} \sum_{t=1}^T r_{k,t}, \quad (18)$$

where the average relative performance measure \bar{d}_k is defined as the average return of the k th alternative model, $r_{k,t}$ is the return of the k th alternative model on period (day) t , and T is the number of periods (days). The number B of samples in the SPA test is set as 5000 to obtain precise P values. Because the SPA test requires bootstrap sampling, the P value is random in an interval, and the sampling takes a great deal of time. Therefore, the SPA test is performed only ten runs in each statistical test in this paper.

3. Portfolio stochastic programming for stock investment

Asset allocation is not a one-time task; in fact, it is a continually dynamic process. In this paper, the investment target is the Taiwan

stock market. To predict a stock's future price is difficult because the price is a nonstationary and chaotic time series. Hence, it is difficult for an investor who is not able to evaluate market information precisely to adjust the portfolio to earn a positive return and reduce the unsystematic risk in the market. To earn a positive return, we first propose the *portfolio SP* (PSP) model incorporated with CVaR for the multiperiod portfolio management problem. To efficiently solve the PSP model, we simplify the model to the *stage-wise PSP* (SPSP) model. The PSP model enables a nearly optimal current asset allocation based on a prediction of the future joint distribution of risky asset returns.

Heuristic moment matching (HMM) [19] can be applied to the generation of scenarios, which are the samples of the joint distribution of risky asset returns on the following period (day). Here, the joint distribution of returns on the following day is assumed to be identical to that of the current day, because the distribution does not change quickly. The joint distribution is modeled according to the values of statistics, including the mean, standard deviation, skewness, excess kurtosis, and Pearson correlation matrix of risky asset returns. The definitions of these statistics can be found in the statistics textbooks [28].

In the HMM method, scenarios are generated in two phases iteratively. In the first phase, random samples are drawn from an arbitrary user-specified distribution, and a cubic transform is applied to these samples. After the cubic transform, the mean, standard deviation, skewness, and kurtosis of these transformed

correlation matrix. After the Cholesky decomposition, although the error between the correlation matrix of the transformed samples and the target correlation matrix is minimized, the error between the statistics of the transformed samples and the target might have been enlarged. Therefore, the transformed samples should be subjected to the two phases repeatedly until the errors in both statistics and correlation matrices converge. Although the HMM method does not guarantee convergence of the errors, it typically converges well in practice.

According to the suggestion provided by Klaassen [30], scenarios are examined to ensure that there is no arbitrage opportunity. In addition, we assume that the wealth allocation of each stock may be fractal, and stocks can always be bought and sold in after-hours trading. Fig. 1 shows the flowchart of our investment method.

The PSP is given in (\mathcal{PSP}) . Our PSP model is multistage event integer linear programming, and the formulation of the model is inspired by the research of Bertsimas and Pachamanov [4]. However, they focused on return maximization and the robust formulation of the problem.

Let M denote the number of risky target assets or the portfolio size, and M_c denote the size of the candidate set of risky assets from which M target assets will be selected. Note that $0 \leq M \leq M_c$. If the setting $M = M_c$ is satisfied, constraints (26)–(28) can be ignored because these constraints are implied in constraints (21)–(25). Therefore, (\mathcal{PSP}) with $M = M_c$ degenerates to multistage linear programming.

$$\max \quad \frac{1}{T} \sum_{t=1}^T \left(z_{t+1} - \frac{1}{1-\alpha} \sum_{s_{t+1}=1}^{S_{t+1}} p_{s_{t+1}, t+1} y_{s_{t+1}, t+1} \right) \quad (19)$$

$$\text{s.t.} \quad y_{s_{t+1}, t+1} \geq z_{t+1} - \sum_{m=0}^{M_c} \bar{w}_t^m (1 + r_{s_{t+1}, t+1}^m), \quad y_{s_{t+1}, t+1} \geq 0, \quad s_t = 1, 2, \dots, S_t, \quad t = 1, 2, \dots, T, \quad (20)$$

$$w_{s_t, t}^m = (1 + r_{s_t, t}^m) \bar{w}_{t-1}^m + b_{s_t, t}^m - e_{s_t, t}^m, \quad s_t = 1, 2, \dots, S_t, \quad m = 1, 2, \dots, M_c, \quad t = 1, 2, \dots, T, \quad (21)$$

$$w_{s_t, t}^0 = (1 + r_{s_t, t}^0) \bar{w}_{t-1}^0 - \sum_{m=1}^{M_c} (1 + c_{buy, s_t, t}^m) b_{s_t, t}^m + \sum_{m=1}^{M_c} (1 - c_{sell, s_t, t}^m) e_{s_t, t}^m, \\ s_t = 1, 2, \dots, S_t, \quad t = 1, 2, \dots, T, \quad (22)$$

$$(\mathcal{PSP}) \quad \bar{w}_t^m = \sum_{s_t \in S_t} p_{s_t, t} w_{s_t, t}^m, \quad m = 0, 1, \dots, M_c, \quad t = 1, 2, \dots, T, \quad (23)$$

$$w_{s_t, t}^m \geq 0, \quad b_{s_t, t}^m \geq 0, \quad e_{s_t, t}^m \geq 0, \quad s_t = 1, 2, \dots, S_t, \quad m = 1, 2, \dots, M_c, \quad t = 1, 2, \dots, T, \quad (24)$$

$$w_{s_t, t}^0 \geq 0, \quad s_t = 1, 2, \dots, S_t, \quad t = 1, 2, \dots, T, \quad (25)$$

$$\frac{w_{s_t, t}^m}{\sum_{i=0}^{M_c} \bar{w}_{t-1}^i (1 + r_{s_t, t}^i)} \leq x_{s_t, t}^m, \quad s_t = 1, 2, \dots, S_t, \quad m = 1, 2, \dots, M_c, \quad t = 1, 2, \dots, T, \quad (26)$$

$$\sum_{m=1}^{M_c} x_{s_t, t}^m \leq M, \quad s_t = 1, 2, \dots, S_t, \quad t = 1, 2, \dots, T, \quad (27)$$

$$x_{s_t, t}^m \in \{0, 1\}, \quad s_t = 1, 2, \dots, S_t, \quad m = 1, 2, \dots, M_c, \quad t = 1, 2, \dots, T, \quad (28)$$

samples are approximate to the target statistics, but the correlation matrix of these transformed samples may not approximate the target correlation matrix. In the second phase, the Cholesky decomposition [29] is applied to the transformed samples to approximate the correlation matrix of the transformed samples to the target

where z_{t+1} is the variable z of the CVaR defined in (12) in period $t+1$; M is the number of risky target assets or the portfolio size; M_c is the size of the candidate set of risky assets, from which M target assets will be selected; T denotes the number of investment periods; S_t denotes the number of scenarios in period t ; s_t

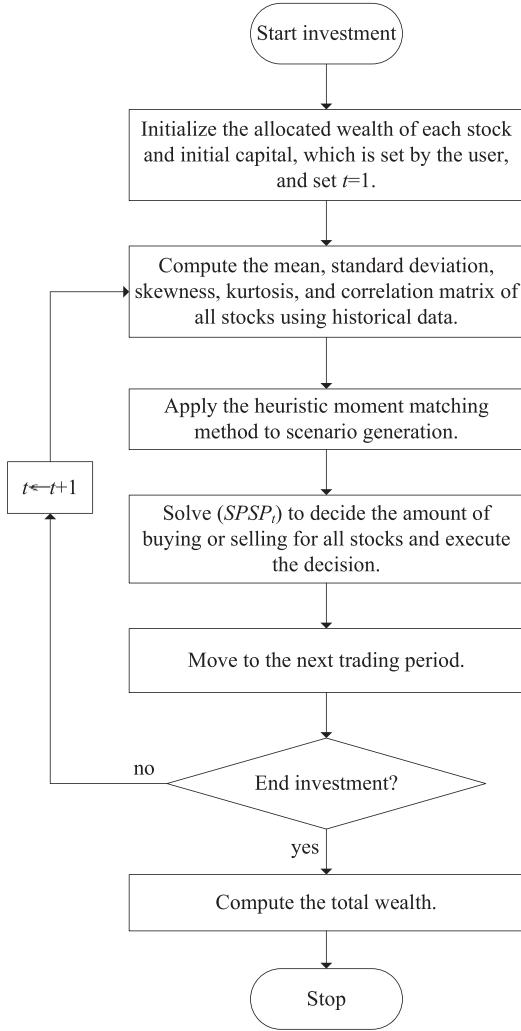


Fig. 1. The flowchart of our investment method.

represents the s_t th scenario in period t ; $r_{s_{t+1},t+1}^m$ is the s_{t+1} th return scenario of the m th risky asset return in period $t+1$; $y_{s_{t+1},t+1}$ is the portfolio wealth of the s_{t+1} th scenario, which involves a loss in excess of the VaR in period $t+1$; $p_{s_{t+1},t+1}$ is the probability of the s_{t+1} th scenario in period $t+1$; $w_{s_t,t}^m$ is the wealth of the s_t th scenario invested in the m th risky asset in period t ; $w_{s_t,t}^0$ is the wealth of the s_t th scenario invested in a risk-free asset in period t ; \bar{w}_t^m is the expected wealth invested in the m th asset in period t ; $b_{s_t,t}^m$ is the amount of money of the s_t th scenario with which an investor buys the m th risky asset in period t ; $e_{s_t,t}^m$ is the amount of money of the s_t th scenario for which an investor sells the m th risky asset in period t ; $c_{buy,s_t,t}^m$ is the tax rate of the s_t th scenario imposed on the purchase of the m th risky asset in period t , and $c_{sell,s_t,t}^m$ is the tax rate of the s_t th scenario imposed on the sale of the m th risky asset in period t . $x_{s_t,t}^m$ is a binary switch variable of the s_t th scenario used to determine whether an investor has invested in the m th risky asset in period t . Note that S_1 contains only one scenario, which is the realized return of all assets before all decisions; \bar{w}_0^m and \bar{w}_0^0 are known constants, where \bar{w}_0^m is typically set as 0, for all $m \geq 1$, and \bar{w}_0^0 is set as the initial capital.

In the (PSP) model, (19) is the objective function, which maximizes the portfolio wealth at a given confidence level α of the CVaR considering all periods. (20) controls the portfolio returns of all scenarios in each period which are lower than the VaR. (21) ex-

presses that the m th risky asset wealth in period t is the return of the previous period plus the value difference of stocks bought and sold in period t . (22) expresses that the wealth of a risk-free asset in period t is the return of the previous period plus the value difference of stocks sold and bought of all risky assets involving a transaction tax. (23) is the expected wealth of the m th asset of the scenarios in S_t in period t . In (24) and (25), the risk-associated and risk-free wealth must be greater than or equal to zero, meaning that short selling is not allowed, and the amounts of stocks bought and sold must both be nonnegative. (26) controls the wealth invested in the m th risky asset in the s_t th scenario. When $x_{s_t,t}^m = 0$, an investor does not invest any capital in the m th risky asset of the s_t th scenario in period t . When $x_t^m = 1$, the amount invested in the m th risky asset of the s_t th scenario in period t cannot exceed the capital that the investor currently holds. (27) expresses that the number of invested assets cannot exceed the portfolio size M .

The unknown decision variables in the (PSP) model are $\{z_{t+1}, y_{s_{t+1},t+1}, w_{s_t,t}^m, w_{s_t,t}^0, \bar{w}_t^m, \bar{w}_t^0, b_{s_t,t}^m, e_{s_t,t}^m, x_{s_t,t}^m, s_t = 1, 2, \dots, S_t, m = 1, 2, \dots, M_c, t = 1, 2, \dots, T\}$. It should be noted that $0 \leq M \leq M_c$. The most important is that (28) forces the model to be multistage mixed integer linear programming. If the special case $M = M_c$ is satisfied, each $x_{s_t,t}^m$ in (28) should be equal to 1. Thus, the constraints (27) and (28) can be ignored. When $M = M_c$, (PSP) degenerates to solve multistage linear programming, which becomes easier. With the consideration of execution efficiency, the implementation of our model is divided into two cases: $M = M_c$ and $M \leq M_c$. The former is much faster than the latter.

Theoretically, the (PSP) model can be solved after all scenarios in all time periods ($t = 1, 2, \dots, T$) have been generated. Directly solving the model (PSP) , however, requires huge computational resources. Besides, the portfolio wealth of model (PSP) in all periods, except the first period, is the expected wealth but not the realized values as shown in (23). An investor usually prefers to know the realized performance of a model instead of its expected values. In practice, estimating the multiperiod joint return distributions of risky assets is difficult, and an investor is typically eager to receive responses or feedback quickly after making investment decisions.

Therefore, to reduce the computational resources needed to solve the (PSP) model, we propose the stagewise PSP (SPSP) model, as expressed in $(SPSP_t)$. The SPSP model entails considering the decision in period t as a two-stage PSP. Thus, the decision variables determined in period $t-1$ become the inputs of period t . Model (PSP) in period t can be solved iteratively by model $(SPSP_t)$.

$$\max z_{t+1} - \frac{1}{1-\alpha} \sum_{s=1}^S p_{s,t+1} y_{s,t+1} \quad (29)$$

$$\text{s.t. } y_{s,t+1} \geq z_{t+1} - \sum_{m=0}^{M_c} w_t^m (1 + r_{s,t+1}^m), \quad y_{s,t+1} \geq 0, \\ s = 1, 2, \dots, S, \quad (30)$$

$$w_t^m = (1 + r_t^m) w_{t-1}^m + b_t^m - e_t^m, \quad m = 1, 2, \dots, M_c, \quad (31)$$

$$w_t^0 = (1 + r_t^0) w_{t-1}^0 - \sum_{m=1}^{M_c} (1 + c_{buy,t}^m) b_t^m + \sum_{m=1}^{M_c} (1 - c_{sell,t}^m) e_t^m, \quad (32)$$

$$(SPSP_t) \quad w_t^m \geq 0, \quad b_t^m \geq 0, \quad e_t^m \geq 0, \quad m = 1, 2, \dots, M_c, \quad (33)$$

$$w_t^0 \geq 0, \quad (34)$$

Algorithm 1: SPSPModel($M_c, M, h, \alpha, capital, P, C_{buy}, C_{sell}$).

input :

M_c : size of the candidate set of risky assets.
 M : portfolio size.
 h : length of the historical period.
 α : confidence level of the CVaR.
 $capital$: initial capital.
 P : matrix of size $S \times T$, where $p_{s,t}$ is the scenarioprobability of scenario s in period t .
 C_{buy} : matrix of size $M \times T$, where $c_{buy,t}^m$ is the purchase tax rate of asset m in period t .
 C_{sell} : matrix of size $M \times T$, where $c_{sell,t}^m$ is the sales tax rate of asset m in period t .

output:

$W^{(p)} = (W_1^{(p)}, \dots, W_T^{(p)})$: wealth vector of the portfolio, where the wealth $W_T^{(p)}$ in period T is nearly maximized.

begin

$w_0^0 \leftarrow capital, w_0^m \leftarrow 0, m = 1, 2, \dots, M_c$.

for ($t \leftarrow 1$ to T) **do**

/* Compute the statistics of the risky assets in the candidate set */

/* Statistics matrix κ : size $M_c \times 4$ */

for $m \leftarrow 1$ to M_c **do**

$\kappa_m \leftarrow$ mean, standard deviation, skewness, and kurtosis matrix of the m th risky assets with returns in the past h periods.

end

$\Sigma \leftarrow$ the correlation matrix of the M_c risky assets in the portfolio using returns in the past h periods.

/* Generate scenarios */

$R_{t+1} \leftarrow S$ scenarios in period $(t+1)$ with (κ, Σ) generated by using HMM, where R_{t+1} is a matrix of size $M_c \times S$, and $r_{s,t+1}^m$ denotes the return of m th asset in s th scenario.

/* Solve the SPSP model */

$\{M, w_{t-1}^m, w_t^0, r_t^m, r_t^0, R_{t+1}, p_{s,t+1}, \alpha, C_{buy,t}^m, C_{sell,t}^m, m = 1, 2, \dots, M_c, s = 1, 2, \dots, S\}$ as the inputs of model $(SPSP_t)$.

$\{w_t^m, w_t^0, b_t^m, e_t^m, y_{s,t+1}, z_{t+1}, x_t^m, m = 1, 2, \dots, M_c, s = 1, 2, \dots, S\} \leftarrow$ solution of model $(SPSP_t)$.

$W_t^{(p)} \leftarrow \sum_{m=0}^{M_c} w_t^m$.

end

return $W^{(p)}$

end

$$\frac{w_t^m}{\sum_{i=0}^{M_c} w_{t-1}^i (1 + r_t^i)} \leq x_t^m, \quad m = 1, 2, \dots, M_c, \quad (35)$$

$$\sum_{m=1}^{M_c} x_t^m \leq M, \quad (36)$$

$$x_t^m \in \{0, 1\}, \quad m = 1, 2, \dots, M_c. \quad (37)$$

where $w_{t-1}^m, 0 \leq m \leq M_c$ is obtained by solving model $(SPSP_{t-1})$. Thus, the calculation of (23) is removed because there is only one realized return r_t^m for each asset in period t , and the symbol s_t in the subscript is omitted. For the first period in model $(SPSP_1)$, the wealth is usually set $w_0^m = 0, 1 \leq m \leq M_c$ and $w_0^0 = 1$, or an arbitrary nonnegative value of wealth. For execution efficiency, the SPSP model is also divided into two cases: $M = M_c$ and $M \leq M_c$.

For more clarity, the period t means a particular time in the investment interval, while the stage t means the moment when a decision is made. More specifically, the decision variables $w_{t-1}^m, m = 0, 1, \dots, M_c$, determined in period $t-1$ become the input of period t , the $z_{t+1}, w_t^m, w_t^0, b_t^m, e_t^m, x_t^m, m = 1, 2, \dots, M_c$, are the decision variables of the stage t (first stage) in period t , and $y_{s,t+1}, s = 1, 2, \dots, S$, are the decision variables of the stage $t+1$ (second stage) in period t .

In each stage, S scenarios are generated by the HMM method. Here, a scenario is an M_c -dimensional vector whose elements each corresponds to the daily return of one risky asset, and the scenario matrix R_{t+1} of size $M_c \times S$ represents all scenario values. To generate scenarios, the HMM requires four M_c -dimensional statistical vectors and an $M_c \times M_c$ correlation matrix Σ . The m th elements of the four M_c -dimensional vectors are the mean, standard deviation, skewness, and kurtosis of the historical daily returns of the

m th risky asset, and each entry of Σ corresponds to the correlation value of the historical daily returns of one risky asset pair in the portfolio in period t . The m th row of R_{t+1} has approximately the same mean, standard deviation, skewness, and kurtosis as the historical daily returns on the m th risky asset. Moreover, the correlation matrix of R_{t+1} is approximately equal to Σ in period t .

The returns of the past h periods (including the current trading period) are used to estimate the statistics of each risky asset. After scenarios in period $t+1$ are generated, the $(SPSP_t)$ model is solved to determine the optimal values of the decision variables $z_{t+1}, y_{s,t+1}, w_t^m, w_t^0, b_t^m$, and e_t^m for each risky asset m and each scenario s in period t . The variables b_t^m and e_t^m are then applied to all risky assets in the portfolio and the wealth w_t^m and w_t^0 is updated accordingly. These steps are repeated until the final period T is reached. The procedure of the SPSP model is formally described in Algorithm 1.

4. Experimental results

The experiments were conducted based on the assumptions that shares of stocks can be bought and sold during after-hours trading in the Taiwan stock market and a fraction of wealth can be invested in stocks. The experimental interval began on January 3, 2005 and ended on December 31, 2014, containing a total of 2484 trading days. The target stocks are the 50 Taiwanese listed companies with the greatest market capitalization on January 3, 2005 in the Taiwan stock exchange, and the data were obtained from the Taiwan Economic Journal [31]. The daily stock return is adjusted according to the dividend of each stock. Thus, the return of a stock is computed according to the price variation of the stock as well as all dividends paid to shareholders.

The experimental environment is a computer equipped with CPU Intel i7-3770, 8G RAM, and the Ubuntu Linux 14.04 operating system. Python Optimization Modeling Objects (Pyomo) [32], an open source software for modeling optimization problems, is used to construct the linear programming and the mixed integer linear programming. Here, the IBM ILOG CPLEX 12.6 optimizer is employed to solve the optimization model.

4.1. Risk measures and statistical tests of target stocks

The Sharpe ratio characterizes the excess return per excess return deviation, sometimes called the *risk*. In general, the higher the Sharpe ratio of a target asset is, the higher the probability is that the target asset will yield a positive return in the future. The formula for calculating the Sharpe ratio S_p is expressed in (38) [33].

$$S_p = \frac{E(r_t) - r_f}{\sigma_p}, \quad (38)$$

where S_p is the Sharpe ratio, r_t is the target return in period t , r_f is the risk-free return, and σ_p is the standard deviation of the difference between the target daily returns and the risk-free return. When $\sigma_p = 0$, S_p is set as 0. Here, r_f was set as 0 in our experiments.

The definition of the Sortino ratio is similar to that of the Sharpe ratio, but the Sortino ratio considers only the downside deviation [34]. Chaudhry and Johnson [35] suggested that the Sortino ratio may be a superior performance measurement compared with the Sharpe ratio when the distribution of excess returns is skewed. The formula for calculating the Sortino ratio S_n is expressed in (39) [34].

$$S_n = \frac{E(r_t) - r_{MAR}}{\sigma_n},$$

$$\sigma_n = \sqrt{\frac{1}{T} \sum_{t=1}^T (r_t \times 1(r_t < r_{MAR}))^2}, \quad (39)$$

where r_t is the target return in period t , r_{MAR} is the minimal acceptable return set by the investor, σ_n is the downside deviation, T is the number of returns, and $1(\cdot)$ is the indicator function. When $\sigma_n = 0$, S_n is set as 0. r_{MAR} was set as 0 in all experiments. Thus, only negative returns were considered as unfavorable aspects of price fluctuation.

The Jarque-Bera (JB) test [36] is a goodness-of-fit test to check if the sample data comes from a normal distribution according to the skewness and kurtosis, and the null hypothesis is that the data comes from the normal distribution. Past studies usually assumed that the return distribution of a financial asset is a normal distribution. However, more and more empirical evidence rejects this assumption, especially for financial time series.

A time series or stochastic process which has unit roots is non-stationary. Testing for unit roots is a crucial aspect of time series analysis because the presence of unit roots determines how to proceed for a correct statistical inference. The *augmented Dickey-Fuller* (ADF) test [37] is a test for a unit root in a univariate time series, where the null hypothesis is that the series has a unit root.

The return series are usually stationary because the return is the percentage of relative change between the current price and the previous price, and thus first-order difference is implied in the equation. Note that a nonstationary series can usually be transformed to a stationary one by differencing [38].

The hypothesis tests, including SPA, JB and ADF tests, may involve many random results for the same target. Here the maximum P value of each test is reported. The symbols *, **, and *** before the P value of each statistical test in the following tables indicate that the results are significant at the 10%, 5%, and 1% levels, respectively.

Table 1 lists the 50 stocks and their statistics of daily returns. In **Table 1**, 15 stocks suffer some loss in the experimental interval if the buy-and-hold rule is applied. The largest loss is stock 2475 with the return of -95.36%, meaning that we lost most of the money invested in stock 2475. Conversely, the highest profit is stock 1216 with the return of 527.97%, which earns five times the invested money. From the mean μ of the daily return, we can see that most stocks have positive means, indicating that these stocks can usually earn a positive profit. However, the standard deviation σ of all stocks is greater than 1%, representing that the fluctuation of the daily return is very large. Thus, it is difficult for an investor to determine the appropriate timing for buying and selling stocks. Most values of the skewness γ_1 are positive, indicating that the return distribution is asymmetric and the mode is smaller than the mean. All values of the excess kurtosis γ_2 are positive, indicating that the distribution is more peaked than the normal distribution, and it is a fat-tailed distribution. The Sharpe ratio S_p and Sortino ratio S_n are both indicators of excess return per risk, and the higher the values are, the better the stocks are. The P values P_{JB} of the JB test show that all stock distributions are not normal, and the P values P_{ADF} of the ADF test show that there is no unit root in all daily return series. The most important are the P values P_{SPA} of the SPA test, showing that only six stocks (marked by ***), with 99% confidence level, can earn a profit without any data-snooping problem. The results of the SPA test show that much effort can still be made in portfolio investment.

4.2. Results of the SPSP model with predetermined target stocks

As mentioned before, the SPSP model is divided into two cases: $M = M_c$ and $M \leq M_c$. The former one is linear programming, and is faster than the latter, which is mixed integer linear programming. In this subsection, we discuss the results of SPSP with $M = M_c$ (referred to as SPSP compact setting), meaning that the invested target stocks have been predetermined. Since $M = M_c$, we use M in terms of M_c when there is no ambiguity.

In all experiments, the initial capital is set to one dollar. On each period (day) t , the HMM method is applied to generate $S = 200$ scenarios of the joint returns of the portfolio on period (day) $t + 1$ according to the estimated mean, standard deviation, skewness, kurtosis vectors, and correlation matrix. The transaction tax rates $c_{buy,t}^m$ and $c_{sell,t}^m$ are set as 0.1425% and 0.4425% (real tax rates in Taiwan) for each period t and each stock m , respectively, and each risk-free return r_t^0 is set as 0 in model ($SPSP_t$).

In the experiments of the SPSP compact setting, the portfolio size $M \in \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$, the number of historical days $h \in \{50, 60, \dots, 230, 240\}$, and the confidence level $\alpha \in \{50\%, 55\%, \dots, 90\%, 95\%\}$ were considered. Hence, $10 \times 20 \times 10 = 2000$ trials in total were conducted. The portfolio size M (for example $M = 10$) means that the stocks in **Table 1** with ranks from 1 to $M = 10$ are selected for the portfolio. A length of historical period h shorter than 50 was not examined here, because a previous study [39] showed that the statistics of return distributions cannot be estimated accurately using short historical periods.

Because the scenarios generated by the HMM method are random, each trial was performed five times to test the stability of our models against the scenario generation method.

One trial requires about 600 to 2400 s, depending on the portfolio size M . In all runs, both the SPSP compact and general settings have the largest $\pm 9\%$ difference in cumulative returns with the same parameter combination (M, h, α) , indicating that the SPSP model is insensitive to the random scenarios generated by the HMM method. For this reason, we show the average values of the five-run results for better representation unless stated otherwise.

In **Fig. 2**, we see that the SPSP compact setting is stable against small variations in historical periods h when $h \geq 100$. The model

Table 1

The daily return statistics of the 50 target stocks of Taiwan for Jan. 3, 2005 through Dec. 31, 2014. These stocks are ranked in descending order by their market capitalization on January 3, 2005. R_C : cumulative return, R_A : annualized return, μ : mean, σ : standard deviation, γ_1 : skewness, γ_2 : excess kurtosis, S_p : Sharpe ratio, S_n : Sortino ratio, P_{SPA} : P value of SPA, P_{JB} : P value of JB, and P_{ADF} : P value of ADF.

rank	symbol	$R_C(\%)$	$R_A(\%)$	$\mu(\%)$	$\sigma(\%)$	γ_1	γ_2	$S_p(\%)$	$S_n(\%)$	$P_{SPA}(\%)$	$P_{JB}(\%)$	$P_{ADF}(\%)$
1	2330	356.11	16.39	0.0763	1.7463	0.14	2.16	4.37	6.50	*** 0.16	*** 0.00	*** 0.00
2	2412	149.44	9.57	0.0419	1.0116	0.60	8.69	4.14	6.27	*** 0.26	*** 0.00	*** 0.00
3	2882	26.36	2.37	0.0297	2.0180	0.15	2.45	1.47	2.15	23.66	*** 0.00	*** 0.00
4	6505	100.14	7.18	0.0403	1.5737	0.17	3.17	2.56	3.76	* 5.94	*** 0.00	*** 0.00
5	2317	165.04	10.24	0.0616	2.1150	0.07	1.61	2.91	4.25	* 9.68	*** 0.00	*** 0.00
6	2303	-3.73	-0.38	0.0196	2.0610	0.33	1.68	0.95	1.42	31.82	*** 0.00	*** 0.00
7	2002	82.58	6.21	0.0345	1.4329	0.31	3.87	2.41	3.57	13.44	*** 0.00	*** 0.00
8	1303	158.77	9.97	0.0526	1.6923	0.12	2.38	3.11	4.54	* 7.52	*** 0.00	*** 0.00
9	1326	142.95	9.28	0.0483	1.5883	0.30	3.35	3.04	4.53	* 8.46	*** 0.00	*** 0.00
10	1301	181.81	10.92	0.0542	1.5836	0.22	2.69	3.42	5.09	* 5.20	*** 0.00	*** 0.00
11	2881	152.71	9.71	0.0568	1.9751	0.09	2.57	2.88	4.19	* 5.14	*** 0.00	*** 0.00
12	2886	93.10	6.80	0.0442	1.8837	0.10	2.72	2.35	3.42	11.62	*** 0.00	*** 0.00
13	2409	-47.71	-6.28	0.0041	2.4596	0.14	0.75	0.17	0.24	48.26	*** 0.00	*** 0.00
14	2891	40.68	3.47	0.0368	2.1464	0.10	2.19	1.71	2.48	19.64	*** 0.00	*** 0.00
15	2357	125.18	8.46	0.0544	2.0869	0.05	1.60	2.61	3.80	* 9.26	*** 0.00	*** 0.00
16	2382	176.41	10.70	0.0615	2.0303	0.26	1.46	3.03	4.56	** 3.82	*** 0.00	*** 0.00
17	3045	316.73	15.34	0.0671	1.3878	0.25	3.42	4.83	7.29	*** 0.10	*** 0.00	*** 0.00
18	2883	-6.10	-0.63	0.0161	1.9297	0.12	2.67	0.83	1.20	33.64	*** 0.00	*** 0.00
19	2454	313.96	15.26	0.0869	2.4401	0.07	0.86	3.56	5.22	** 3.06	*** 0.00	*** 0.00
20	2880	31.19	2.75	0.0258	1.7255	0.16	3.91	1.50	2.17	17.00	*** 0.00	*** 0.00
21	2892	42.91	3.63	0.0311	1.8301	0.28	3.50	1.70	2.50	18.88	*** 0.00	*** 0.00
22	4904	201.68	11.67	0.0539	1.3734	0.28	3.37	3.92	5.89	*** 0.70	*** 0.00	*** 0.00
23	2887	-16.24	-1.76	0.0158	2.1430	0.15	2.43	0.74	1.07	38.76	*** 0.00	*** 0.00
24	2353	-33.87	-4.05	0.0095	2.2902	0.16	1.19	0.42	0.61	43.76	*** 0.00	*** 0.00
25	2324	33.29	2.92	0.0350	2.1666	0.11	1.44	1.62	2.35	19.94	*** 0.00	*** 0.00
26	2801	32.82	2.88	0.0290	1.8787	0.21	2.55	1.55	2.26	22.46	*** 0.00	*** 0.00
27	1402	154.20	9.78	0.0641	2.3088	0.18	1.37	2.78	4.10	* 7.24	*** 0.00	*** 0.00
28	2311	281.04	14.31	0.0814	2.3487	0.17	0.97	3.47	5.17	** 3.46	*** 0.00	*** 0.00
29	2475	-95.36	-26.43	-0.0807	2.9277	0.19	0.49	-2.76	-3.94	88.66	*** 0.00	*** 0.00
30	2888	-55.26	-7.73	-0.0069	2.2579	0.12	1.94	-0.31	-0.44	57.26	*** 0.00	*** 0.00
31	2408	-85.78	-17.72	-0.0228	3.3383	0.12	0.03	-0.68	-0.98	62.08	* 5.94	*** 0.00
32	2308	526.67	20.14	0.0972	2.1587	0.04	1.30	4.50	6.62	*** 0.66	*** 0.00	*** 0.00
33	2301	64.02	5.07	0.0395	1.9805	-0.00	2.10	2.00	2.88	15.78	*** 0.00	*** 0.00
34	2352	-68.92	-11.03	-0.0102	2.7169	0.18	0.88	-0.38	-0.55	57.20	*** 0.00	*** 0.00
35	2603	20.34	1.87	0.0315	2.1985	0.27	1.76	1.43	2.14	21.06	*** 0.00	*** 0.00
36	2884	47.33	3.95	0.0335	1.8906	0.09	2.51	1.77	2.57	19.10	*** 0.00	*** 0.00
37	2890	7.98	0.77	0.0245	2.0711	0.11	2.15	1.18	1.71	27.84	*** 0.00	*** 0.00
38	2609	2.16	0.21	0.0264	2.2653	0.28	1.53	1.17	1.73	29.54	*** 0.00	*** 0.00
39	9904	187.08	11.12	0.0610	1.9280	0.24	2.02	3.16	4.73	** 4.20	*** 0.00	*** 0.00
40	2610	-31.13	-3.66	0.0067	2.0864	0.20	2.44	0.32	0.47	45.10	*** 0.00	*** 0.00
41	1216	527.97	20.17	0.0949	2.0500	0.23	2.15	4.63	7.01	*** 0.56	*** 0.00	*** 0.00
42	1101	262.45	13.74	0.0758	2.1922	0.12	1.74	3.46	5.13	** 4.46	*** 0.00	*** 0.00
43	2325	273.47	14.08	0.0807	2.3564	0.19	1.00	3.43	5.11	** 2.40	*** 0.00	*** 0.00
44	2344	-12.14	-1.29	0.0350	2.8394	0.21	0.62	1.23	1.82	29.30	*** 0.00	*** 0.00
45	2323	-73.33	-12.38	-0.0258	2.3412	0.17	1.72	-1.10	-1.58	73.54	*** 0.00	*** 0.00
46	2371	-70.99	-11.64	-0.0180	2.5261	0.17	1.25	-0.71	-1.03	64.96	*** 0.00	*** 0.00
47	2204	-0.37	-0.04	0.0237	2.1865	0.21	2.18	1.08	1.59	31.66	*** 0.00	*** 0.00
48	1605	-22.57	-2.53	0.0177	2.3662	0.09	1.35	0.75	1.08	35.76	*** 0.00	*** 0.00
49	2615	79.82	6.04	0.0471	2.1706	0.15	1.47	2.17	3.19	11.96	*** 0.00	*** 0.00
50	2201	60.79	4.86	0.0439	2.2294	0.32	2.10	1.97	2.96	21.48	*** 0.00	*** 0.00
	avg	95.91	3.89	0.0361	2.0861	0.18	2.11	1.93	2.86			
	dev	145.14	9.46	0.0336	0.4098	0.10	1.30	1.63	2.42			

can earn a positive return when the confidence level of CVaR is in the range $50\% \leq \alpha \leq 85\%$. When α is low, representing that investing action is more aggressive, it usually brings a higher profit. The compact setting cannot yield high returns when $\alpha \geq 90\%$, because conservative investors usually do not buy a promising stock in a price-rising trend.

To make sure the compact setting can earn a positive return without the data-snooping problem, we apply the SPA test to all trials of the compact setting. The maximum P values of the five runs are shown in Fig. 3. As the results show, when $50\% \leq \alpha \leq 85\%$ and $h \geq 100$, most of the P values are less than 1%. This means that the compact setting can earn a positive return in most parameter combinations with a high confidence level. Thus, it is a robust model for portfolio investment.

Table 2 lists the parameter combinations, which yield the highest average cumulative returns at various portfolio sizes M . All the annualized returns in the table are higher than 10%, and the standard deviations of daily returns are less than 1%, indicating that the SPSP compact setting has a higher return and smaller price fluctuation than every single stock listed in **Table 1**. The skewness γ_1 is positive in half of the parameter combinations, while the excess kurtosis γ_2 is higher than a single stock, indicating the return distribution of the portfolio is more peaked than the distribution of a single stock. The Sharpe and Sortino ratios also show that the SPSP compact setting gets a better return per risk measure than a single stock. The daily returns of the best parameter combinations in **Table 2** are not normal distributions and stationary series according to the P values of the JB test and ADF test, respectively.

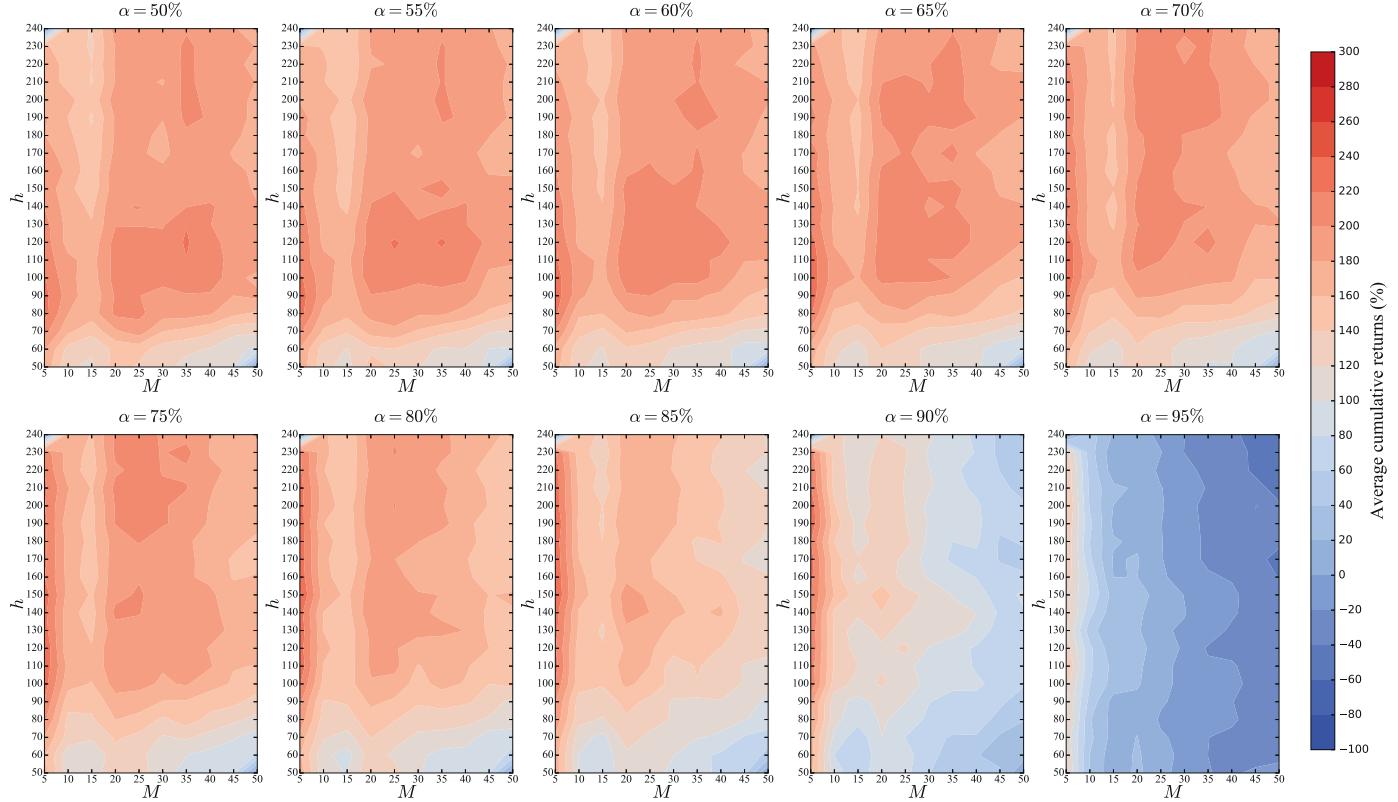


Fig. 2. The contour of the average cumulative returns of the SPSP model with $M = M_c$ (the SPSP compact setting) at various values of confidence level α of the CVaR.

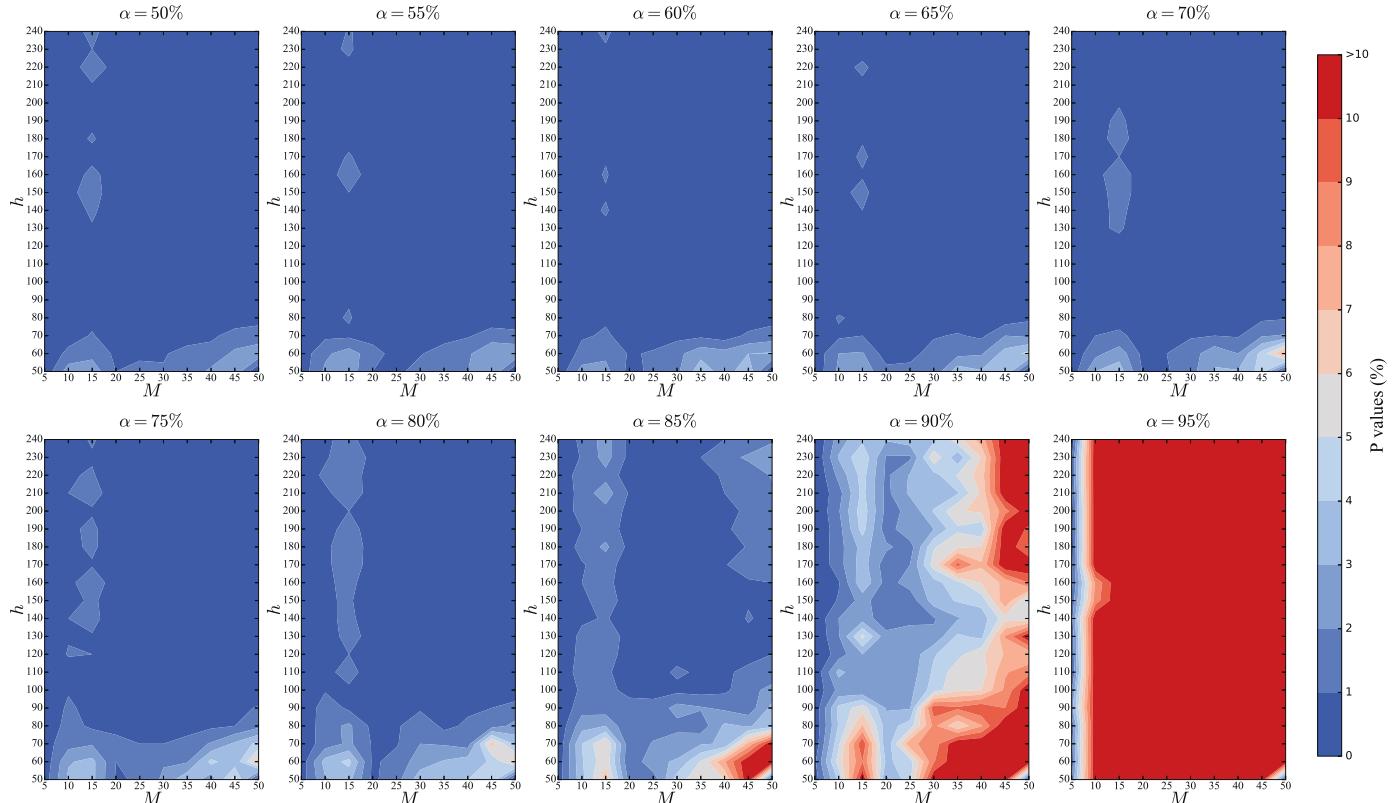


Fig. 3. The contour of the maximum P values of a five-run SPA test in the SPSP model with $M = M_c$ (the SPSP compact setting) at various values of confidence level α of the CVaR.

Table 2

The parameter combinations and statistics of the SPSP model with $M = M_c$ (the SPSP compact setting) having the highest average cumulative return at various values of the portfolio size M .

(M, h, α)	$R_C(\%)$	$R_A(\%)$	$\mu(\%)$	$\sigma(\%)$	γ_1	γ_2	$S_p(\%)$	$S_n(\%)$	$P_{SPA} (\%)$	$P_{JB} (\%)$	$P_{ADF} (\%)$
(5, 150, 80%)	242.20	13.09	0.0539	0.9263	0.41	7.42	5.82	8.78	*** 0.10	*** 0.00	*** 0.00
(10, 90, 50%)	190.82	11.27	0.0472	0.9113	0.40	7.59	5.18	7.75	*** 0.30	*** 0.00	*** 0.00
(15, 100, 65%)	178.82	10.80	0.0453	0.8955	0.31	7.03	5.06	7.52	*** 0.60	*** 0.00	*** 0.00
(20, 110, 60%)	215.07	12.16	0.0500	0.8658	0.14	5.28	5.78	8.56	*** 0.10	*** 0.00	*** 0.00
(25, 120, 55%)	222.59	12.43	0.0508	0.8503	0.02	4.99	5.98	8.78	*** 0.10	*** 0.00	*** 0.00
(30, 190, 70%)	211.24	12.02	0.0493	0.8469	-0.01	5.34	5.83	8.53	*** 0.30	*** 0.00	*** 0.00
(35, 120, 55%)	222.73	12.43	0.0508	0.8470	-0.08	4.60	6.00	8.73	*** 0.20	*** 0.00	*** 0.00
(40, 100, 50%)	214.69	12.15	0.0499	0.8560	-0.12	4.28	5.83	8.45	*** 0.20	*** 0.00	*** 0.00
(45, 120, 55%)	196.72	11.49	0.0474	0.8501	-0.12	4.42	5.58	8.08	*** 0.00	*** 0.00	*** 0.00
(50, 120, 55%)	197.98	11.54	0.0476	0.8506	-0.11	4.57	5.60	8.11	*** 0.40	*** 0.00	*** 0.00

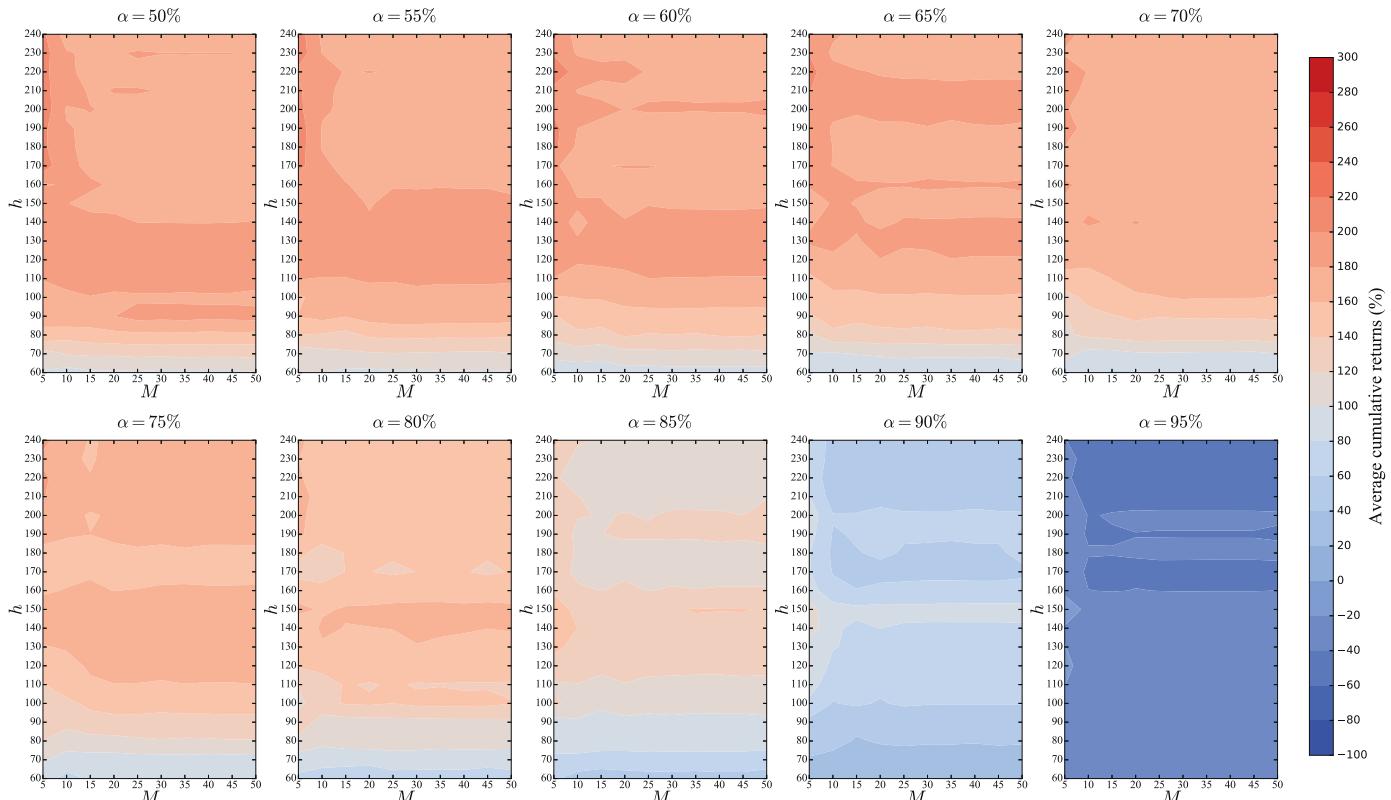


Fig. 4. The contour of the average cumulative returns of the SPSP general setting at various values of confidence level α of the CVaR.

By the P values of the SPA test, the best parameter combinations of the SPSP compact setting can earn a positive return without any data-snooping problem.

4.3. Results of the SPSP model with a set of candidate stocks

For the SPSP model with $0 \leq M \leq M_c$ (referred to as SPSP general setting), a set of M_c candidate stocks are given, but the M invested stocks have not been decided yet. In other words, at most M stocks of the candidate set are invested in each period. In the experiment of the SPSP general setting, all 50 stocks ($M_c = 50$) listed in Table 1 form the set of candidate stocks. The required time for one trial is about 2400 to 6000 seconds.

Figs. 4 and 5 show the average cumulative returns and the maximum P values of the P_{SPA} of the five runs in the SPSP general setting, respectively. The results of the SPSP general setting have similar conclusions as the SPSP compact setting. For most parameter combinations, both settings can earn a positive return without any data-snooping problem. Comparing Figs. 2 and 4, the SPSP general setting usually earns more stable returns than the SPSP compact

setting, because the SPSP general setting utilizes more stocks to generate scenarios than the SPSP compact setting does.

Table 3 lists the parameter combinations of the highest average cumulative returns at various portfolio sizes M . All the annualized returns are higher than 11%, and all the standard deviations of daily return are less than 0.9%. Every skewness γ_1 is negative, indicating that it is the left-skewness distribution, and the mode is larger than the mean. The excess kurtosis γ_2 is higher than a single stock, but usually lower than the corresponding value of the same portfolio size M in Table 2. Comparing to the annualized return R_A and standard deviation σ in Table 2, the annualized returns have similar values, but most of the standard deviations in the SPSP general setting are lower, indicating that the SPSP general setting has lower price fluctuation than the SPSP compact setting. The daily returns of the best parameter combinations in Table 3 are not normal distributions and stationary series according to the P values of the JB test and ADF test, respectively. According to the P values of the SPA test, the best parameter combinations of the SPSP general setting can also earn a positive return without any data-snooping problem.

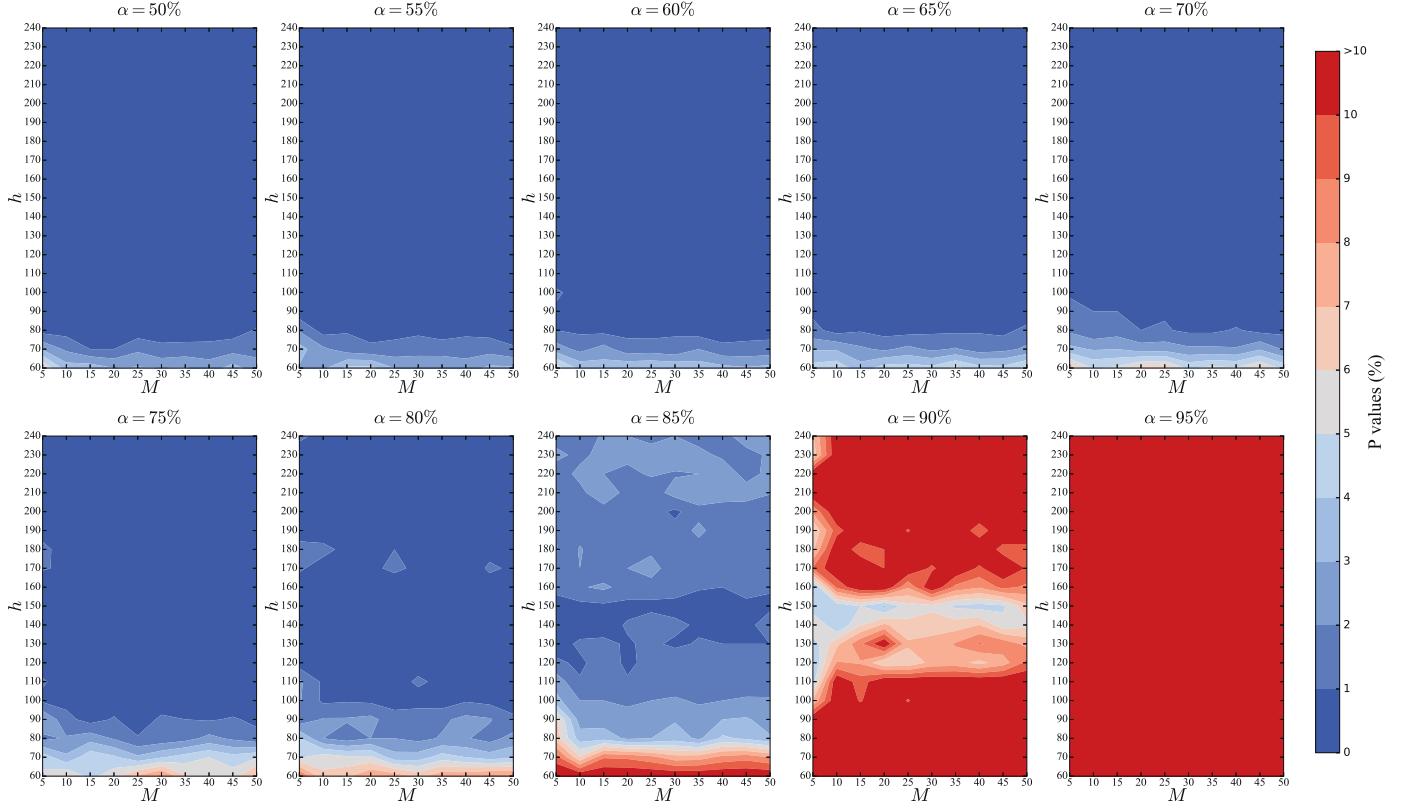


Fig. 5. The contour of the maximum P values of a five-run SPA test in the SPSP general setting at various values of confidence level α of the CVaR.

Table 3

The parameter combinations and statistics of the SPSP general setting having the highest cumulative average returns at various portfolio sizes M .

(M, h, α)	$R_C(\%)$	$R_A(\%)$	$\mu(\%)$	$\sigma(\%)$	γ_1	γ_2	$S_p(\%)$	$S_n(\%)$	$P_{SPA} (\%)$	$P_{JB} (\%)$	$P_{ADF} (\%)$
(5, 200, 50%)	212.34	12.06	0.0495	0.8583	-0.03	5.35	5.77	8.47	*** 0.00	*** 0.00	*** 0.00
(10, 130, 50%)	195.91	11.46	0.0474	0.8544	-0.09	4.66	5.55	8.05	*** 0.30	*** 0.00	*** 0.00
(15, 120, 50%)	194.30	11.40	0.0472	0.8565	-0.12	4.52	5.51	7.96	*** 0.10	*** 0.00	*** 0.00
(20, 120, 50%)	196.84	11.49	0.0475	0.8556	-0.13	4.53	5.55	8.02	*** 0.10	*** 0.00	*** 0.00
(25, 120, 55%)	198.43	11.55	0.0477	0.8504	-0.11	4.56	5.61	8.12	*** 0.20	*** 0.00	*** 0.00
(30, 120, 55%)	199.37	11.59	0.0478	0.8507	-0.11	4.55	5.62	8.14	*** 0.10	*** 0.00	*** 0.00
(35, 120, 55%)	199.37	11.59	0.0478	0.8508	-0.11	4.55	5.62	8.14	*** 0.20	*** 0.00	*** 0.00
(40, 120, 55%)	198.43	11.55	0.0477	0.8507	-0.11	4.56	5.61	8.12	*** 0.10	*** 0.00	*** 0.00
(45, 120, 55%)	198.73	11.57	0.0477	0.8506	-0.11	4.56	5.61	8.13	*** 0.10	*** 0.00	*** 0.00
(50, 120, 55%)	198.17	11.54	0.0477	0.8504	-0.11	4.56	5.61	8.12	*** 0.00	*** 0.00	*** 0.00

Table 4

The cumulative returns and statistics of the BAH rule in the experimental interval.

M	$R_C(\%)$	$R_A(\%)$	$\mu(\%)$	$\sigma(\%)$	γ_1	γ_2	$S_p(\%)$	$S_n(\%)$	$P_{SPA} (\%)$	$P_{JB} (\%)$	$P_{ADF} (\%)$
5	158.28	9.95	0.0457	1.2250	-0.08	2.95	3.73	5.36	** 2.24	*** 0.00	*** 0.00
10	134.91	8.92	0.0415	1.1967	0.04	3.04	3.47	5.01	** 4.74	*** 0.00	*** 0.00
15	113.95	7.90	0.0381	1.2216	-0.03	3.05	3.12	4.46	* 6.38	*** 0.00	*** 0.00
20	126.78	8.53	0.0400	1.1868	-0.07	3.12	3.37	4.81	** 4.88	*** 0.00	*** 0.00
25	110.41	7.72	0.0369	1.1759	-0.08	3.29	3.14	4.47	* 6.54	*** 0.00	*** 0.00
30	102.47	7.31	0.0358	1.2173	-0.10	3.13	2.94	4.17	* 8.00	*** 0.00	*** 0.00
35	100.74	7.22	0.0355	1.2213	-0.15	3.12	2.91	4.10	* 8.38	*** 0.00	*** 0.00
40	93.41	6.82	0.0341	1.2245	-0.16	3.19	2.78	3.91	* 9.66	*** 0.00	*** 0.00
45	104.63	7.42	0.0366	1.2442	-0.18	3.15	2.94	4.13	* 7.96	*** 0.00	*** 0.00
50	95.05	6.91	0.0348	1.2570	-0.19	3.13	2.77	3.88	* 9.58	*** 0.00	*** 0.00

4.4. Performance comparison

Under the buy-and-hold (BAH) rule, an investor buys and allocates equal capital to the selected stocks at the beginning of investment, and sells these stocks at the end of investment. Table 4 shows the cumulative returns and statistics of the BAH rule.

Fig. 6(a)–(c) show the wealth processes of the parameter combinations of the SPSP compact setting, the SPSP general setting, and the BAH rule in Tables 2, 3, and 4, respectively. As shown in the figures, both the SPSP compact and general settings yield higher returns and lower volatility than the BAH rule. The SPSP compact setting yields higher returns than the SPSP general

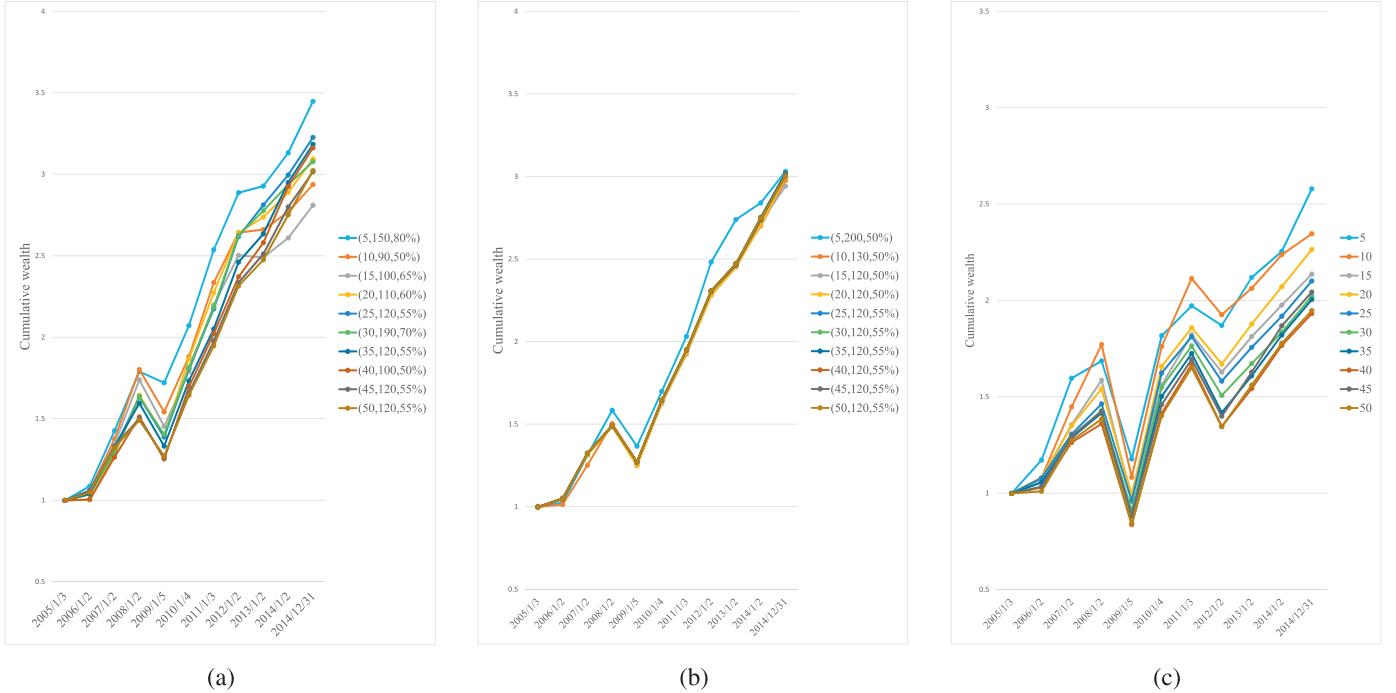


Fig. 6. The cumulative wealth with the best parameter combinations at various portfolio sizes M , where the initial capital is one dollar. (a) The SPSP compact setting ($M = M_c$). (b) The SPSP general setting ($M \leq M_c$). (c) The BAH rule.

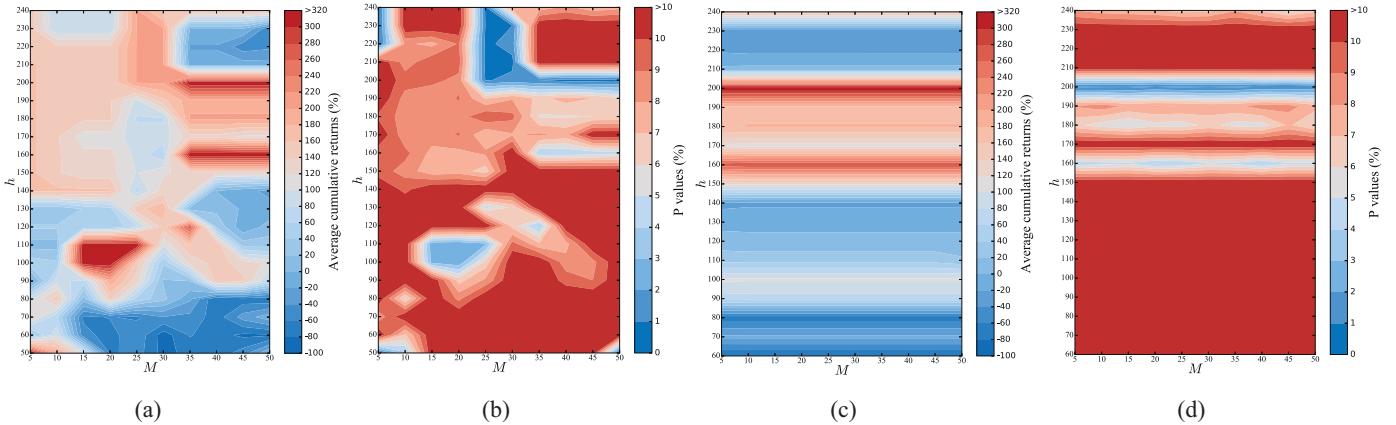


Fig. 7. The contour of average cumulative returns and maximum P values of five-run SPA test in the SPSP model by using expected scenarios. (a) The returns of the SPSP compact setting ($M = M_c$). (b) The P values of the SPSP compact setting. (c) The returns of the SPSP general setting ($M \leq M_c$). (d) The P values of the SPSP general setting.

setting, while the wealth process of the SPSP general setting is more stable, because more stocks are considered during scenario generation in the SPSP general setting than in the SPSP compact setting.

Comparing Tables 2–4, all the cumulative returns, standard deviation of daily return, Sharpe ratio, and Sortino ratio of the SPSP compact setting and general setting outperform the corresponding values of the BAH rule, because the CVaR is considered during decision making in both models. In the three tables, the highest average annualized returns of the SPSP compact setting and general setting are 13.09% with $(M, h, \alpha) = (5, 150, 80\%)$ and 12.06% with $(M, h, \alpha) = (5, 200, 50\%)$, respectively, which are higher than that of the buy-and-hold (BAH) rule (9.95%).

According to (3), we can get an approximate solution by replacing the generated scenarios (variable $r_{s,t+1}^m$) by the expected scenario ($\bar{r}_{t+1}^m = \sum_{s=1}^S r_{s,t+1}^m / S$) in each period to get the first-stage solutions of the $(SPSP_t)$ model. The second-stage variable $y_{s,t+1}$ in model $(SPSP_t)$ must have a value of zero because there is only

one scenario. Thus, the first-stage variables and the objective value are independent of the parameter α . In addition, because only one scenario is used, it results in the investment of the one stock with the highest expected return. In summary, the SPSP model with the expected scenario is independent of the first-stage variables, objective value, parameter α , and portfolio size M .

Comparing Figs. 7 (a), (b), 2, and 3 in the SPSP compact setting, and Figs. 7 (c), (d), 4, and 5 in the SPSP general setting, some parameter combinations with the expected scenarios yield a high positive return. However, the expected models are more sensitive to small variation of the historical periods h , indicating that an investor cannot easily search appropriate parameters for the expected models. The P values also show that the expected model cannot earn a positive return in most parameter combinations.

Because the SPSP model in each period (day) is a two-stage model, the VSS of each period is calculated, and the average daily VSS of the experimental interval is shown in Figs. 8 and 9. As expressed in (8), the average daily VSS represents the average

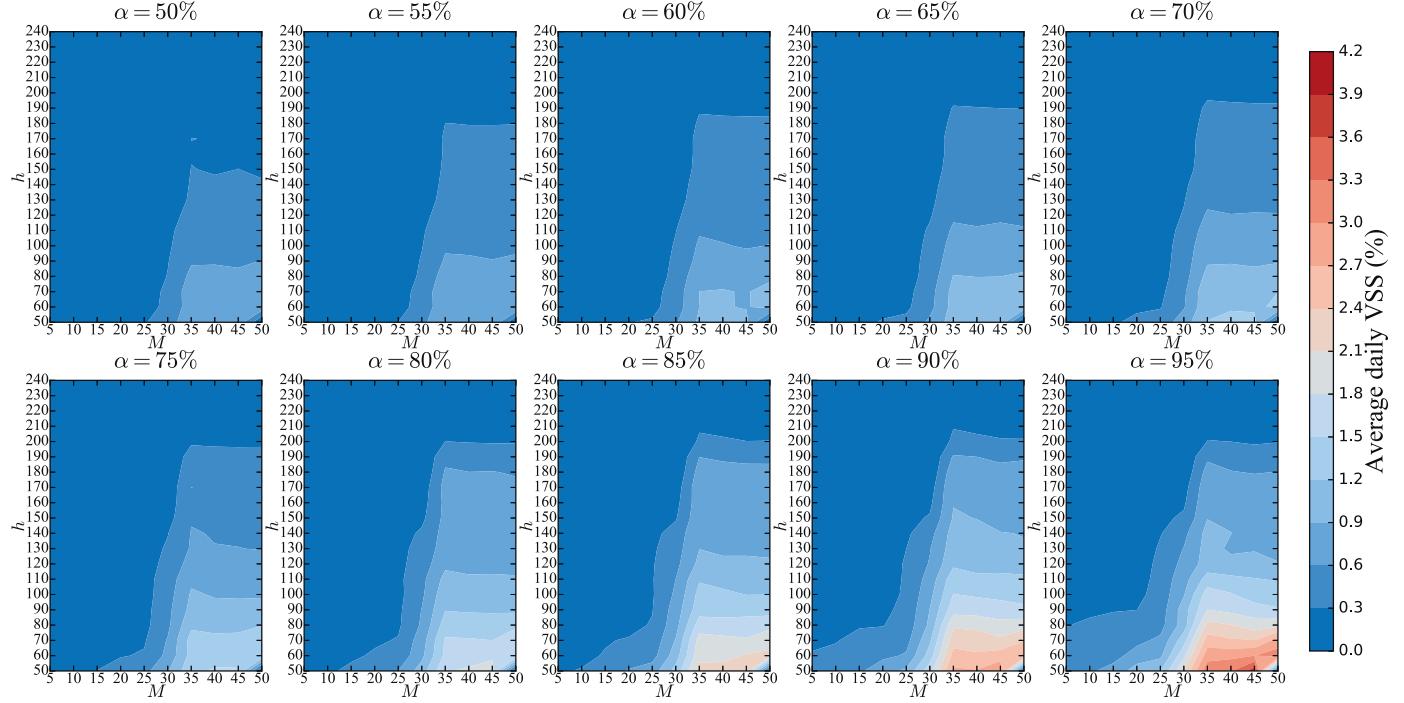


Fig. 8. Average daily VSS of the SPSP compact setting ($M = M_c$) at various portfolio sizes M .

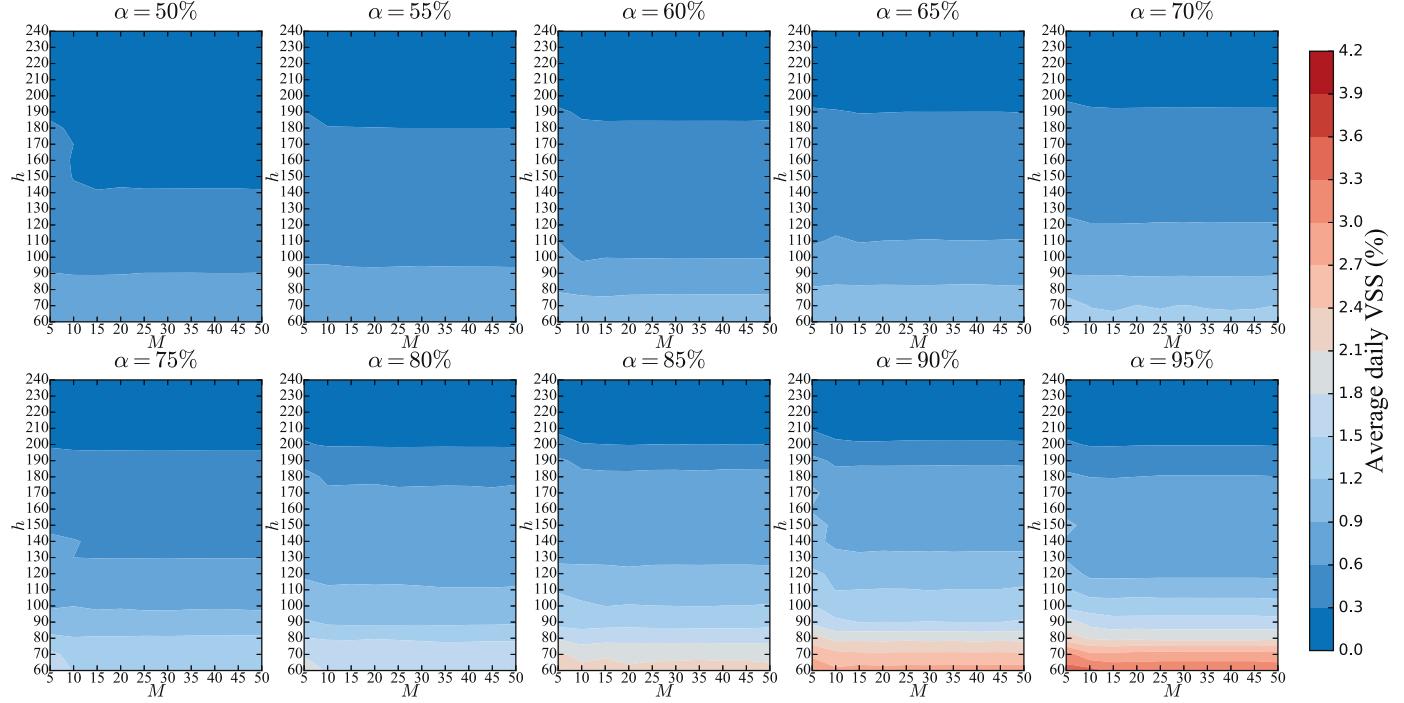


Fig. 9. Average daily VSS of the SPSP general setting ($M \leq M_c$) at various portfolio sizes M .

substantial profit of the SPSP compact and general settings minus their corresponding expected models in the decision of each period (day). Figs. 8 and 9 indicate that the SPSP compact and general settings (complicated models) can earn higher returns than the expected models (simple models), especially when the joint distribution of the expected model is difficult to estimate. Although the average daily VSS in the two figures is not very high, the effect of daily compounding will earn significant profit in a long-term investment. As the VSS of the SPSP compact setting shown in Fig. 8, the average daily VSS is larger than 0.3% in the lower right corner.

The reason for this is that the more the specified stocks and the shorter the historical periods are, the more difficult it is to estimate the joint return distribution. The SPSP general setting always uses a fixed number of stocks ($M_c = 50$) to generate scenarios, so the average daily VSS in Fig. 9 is not sensitive to the portfolio size M .

Directly solving the PSP model is very intractable, since a huge amount of computation is required. For example, in our preliminary tests, directly solving the PSP model with the compact setting of portfolio size $M = M_c = 5$ in a trial of one year

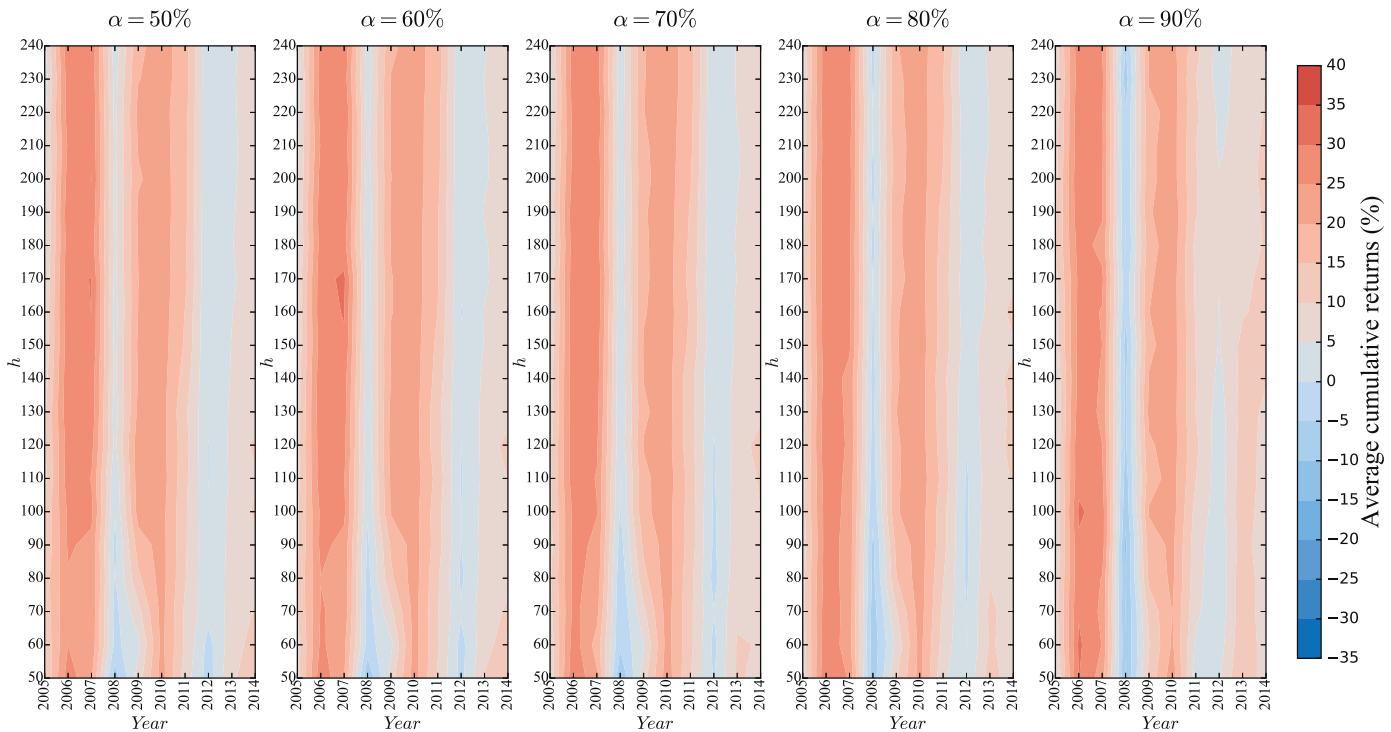


Fig. 10. The contour of the average yearly cumulative returns of the SPSP model with portfolio size $M = M_c = 5$ (the PSP compact setting) at various values of confidence level α of the CVaR.

investment costs about four hours and three gigabyte of memory space. If all stocks ($M_c = 50$) are considered, the PSP model requires at least 10 times the memory space and hours of the model with $M = M_c = 5$. Thus, we simplify the experiment procedure. We performed the experiments of yearly investment on the PSP compact setting with a small portfolio size of $M = 5$. In the yearly investment, the capital is invested in the selected stocks on the first trading day of the year, and the stocks are sold out on the last trading day of the same year. We perform each trial ($M = 5, h, \alpha$) five runs in yearly investment, where the number of historical days $h \in \{50, 60, \dots, 230, 240\}$, and the confidence level $\alpha \in \{50\%, 60\%, 70\%, 80\%, 90\%\}$. So, there are $1 \times 20 \times 5 = 100$ trials in a year.

Figs. 11 and 10 shows the average yearly cumulative returns of PSP and SPSP compact setting with $M = M_c = 5$, respectively. The cumulative returns of the PSP compact setting usually get higher values than the SPSP model, because the PSP compact setting considers the expected wealth as shown in (23) instead of the realized wealth in all stages as described in (23). A rational investor prefers the realized performance of a model to the expected one. In addition, Fig. 11 shows that the model tends to earn higher returns in short historical periods h than in longer periods, which means the PSP compact setting is more sensitive to h than the SPSP compact setting. Because of the requirement of computational resources, the realized wealth against the expected wealth, and the sensitivity to historical periods, the SPSP compact and general settings are more robust and more suitable than the multistage model in real investment.

5. Conclusion

We propose the PSP and SPSP models for multiperiod investment in stocks; the SPSP model effectively reduces the computational resources for solving the PSP model. The portfolio size M of the SPSP model can be set with two cases. When $M = M_c$ (the SPSP compact setting), the M invested target stocks have been predetermined. When $M \leq M_c$ (the SPSP general setting), a set of M_c

candidate stocks are given, but the M real target stocks have not been decided yet. The SPSP model consists of two main phases: (1) to generate scenarios of joint returns on period $t + 1$ according to the mean, standard deviation, skewness, kurtosis, and correlation matrix of the historical data before period t ; and (2) to solve the SPSP model for determining the optimal buying amount and selling amount for each stock in the portfolio.

The SPSP compact and general settings achieve average cumulative returns of 242.20% (annualized return: 13.09%) and 212.34% (12.06%) with the best parameter combinations $(M, h, \alpha) = (5, 150, 80\%)$ and $(5, 200, 50\%)$, respectively. Experimental results show that both the SPSP compact and general settings yield higher returns and less volatility than the single stock and the BAH rule, and both models are not sensitive to small variations of portfolio size M and historical periods h . In addition, the Sharpe and Sortino ratios indicate that both the SPSP compact and general settings outperform the BAH rule because the CVaR is considered as the risk measurement. According to the P values of the SPA test, both models effectively avoid the data-snooping problem. The average daily VSS of the SPSP compact setting shows, when the joint distribution is difficult to estimate, the SPSP compact setting with a variety of scenarios can earn higher returns than one expected scenario.

Directly solving the multistage model (the PSP model) requires much more computing time and memory space than required by the stagewise model (the SPSP model). Additionally, the PSP model is more sensitive to the historical length h than the SPSP model. Thus, the SPSP model is more robust and more suitable than the PSP model in practice. Although the SPSP general setting more accurately approximates real investment than the SPSP compact setting does, solving the SPSP general setting requires more computation time.

In the future, developing a method for generating more accurate scenarios might be worthwhile, because the HMM method considers only static statistical properties. However, time series have been proved to contain autocorrelation properties. In

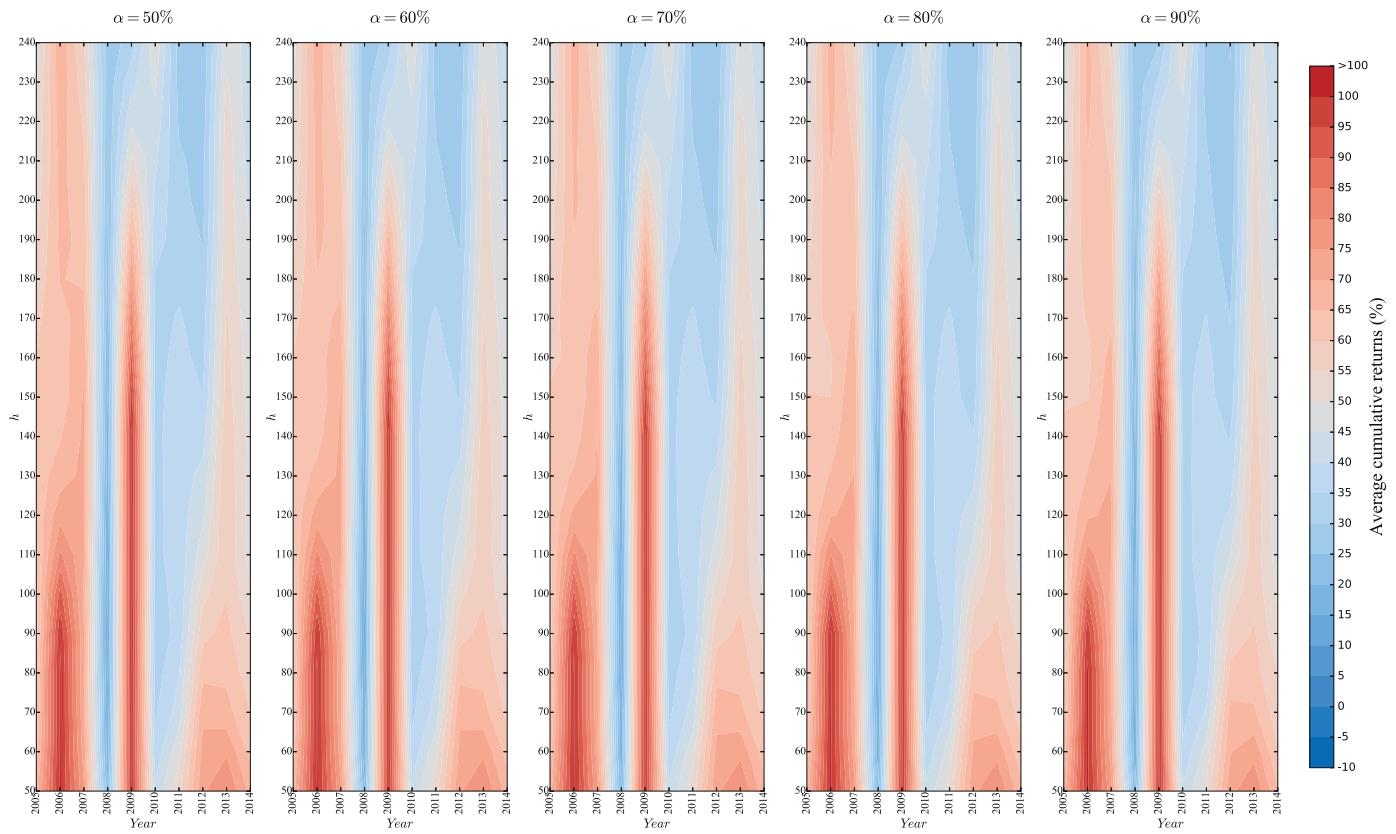


Fig. 11. The contour of the average yearly cumulative returns contour of the PSP model with portfolio size $M = M_c = 5$ (the PSP compact setting) at various values of confidence level α of the CVaR.

addition, a more accurate method may have to involve more constraints for approximating the real investment.

Acknowledgments

This research was partially supported by the National Science Council of Taiwan under contract NSC 101-2221-E-110-020-MY3.

References

- [1] Markowitz HM. Portfolio selection. *J Finance* 1952;7:77–91. doi:[10.1111/j.1540-6261.1952.tb01525.x](https://doi.org/10.1111/j.1540-6261.1952.tb01525.x).
- [2] Ben-Tal A, Margalit T, Nemirovski A. Robust modeling of multi-stage portfolio problems. In: Frenk H, Roos K, Terlaky T, Zhang S, editors. High performance optimization. Applied Optimization, 33. Springer US; 2000. p. 303–28. doi:[10.1007/978-1-4757-3216-0](https://doi.org/10.1007/978-1-4757-3216-0).
- [3] Li D, Ng W-L. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Math Finance* 2000;10(3):387–406. doi:[10.1111/1467-9965.00100](https://doi.org/10.1111/1467-9965.00100).
- [4] Bertsimas D, Pachamanova D. Robust multiperiod portfolio management in the presence of transaction costs. *Comput Oper Res* 2008;35(1):3–17. doi:[10.1016/j.cor.2006.02.011](https://doi.org/10.1016/j.cor.2006.02.011).
- [5] Topaloglu N, Vladimirou H, Zenios SA. CVaR models with selective hedging for international asset allocation. *J Bank Finance* 2002;26(7):1535–61. doi:[10.1016/S0378-4266\(02\)00289-3](https://doi.org/10.1016/S0378-4266(02)00289-3).
- [6] Topaloglu N, Vladimirou H, Zenios SA. A dynamic stochastic programming model for international portfolio management. *Eur J Oper Res* 2008;185(3):1501–24. doi:[10.1016/j.ejor.2005.07.035](https://doi.org/10.1016/j.ejor.2005.07.035).
- [7] Artzner P, Delbaen F, Eber J-M, Heath D. Coherent measures of risk. *Math Finance* 1999;9(3):203–28. doi:[10.1111/1467-9965.00068](https://doi.org/10.1111/1467-9965.00068).
- [8] Pflug G. Some remarks on the value-at-risk and the conditional value-at-risk. In: Uryasev S, editor. Probabilistic constrained optimization: methodology and applications. Nonconvex Optimization and Its Applications, 49. Springer US; 2000. p. 272–81. doi:[10.1007/978-1-4757-3150-7_15](https://doi.org/10.1007/978-1-4757-3150-7_15).
- [9] Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. *J Risk* 2000;2(3):21–42.
- [10] Rockafellar RT, Uryasev S. Conditional value-at-risk for general loss distributions. *J Banking Finance* 2002;26(7):1443–71. doi:[10.1016/S0378-4266\(02\)00271-6](https://doi.org/10.1016/S0378-4266(02)00271-6).
- [11] Shapiro A, Dentcheva D, Ruszczynski A. Lectures on stochastic programming: modeling and theory. SIAM-Society for Industrial and Applied Mathematics; 2009. doi:[10.1137/1.9780898718751](https://doi.org/10.1137/1.9780898718751).
- [12] Hansen PR. A test for superior predictive ability. *J Bus Econ Stat* 2005;23(4):365–80. doi:[10.1198/073500105000000063](https://doi.org/10.1198/073500105000000063).
- [13] Birge JR, Louveaux F. Introduction to stochastic programming. Springer Series in Operations Research and Financial Engineering. Springer New York; 2011. doi:[10.1007/978-1-4614-0237-4](https://doi.org/10.1007/978-1-4614-0237-4).
- [14] Wets RJ-B. Stochastic programs with fixed recourse: The equivalent deterministic program. *SIAM Rev* 1974;16(3):309–39. doi:[10.1137/1016053](https://doi.org/10.1137/1016053).
- [15] Rockafellar RT, Wets RJ-B. Scenarios and policy aggregation in optimization under uncertainty. *Math Oper Res* 1991;16(1):119–47. doi:[10.1287/moor.16.1.119](https://doi.org/10.1287/moor.16.1.119).
- [16] Madansky A. Inequalities for stochastic linear programming problems. *Manag Sci* 1960;6(2):197–204. doi:[10.1287/mnsc.6.2.197](https://doi.org/10.1287/mnsc.6.2.197).
- [17] Pflug GC. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Math Program* 2001;89(2):251–71. doi:[10.1007/PL00011398](https://doi.org/10.1007/PL00011398).
- [18] Høyland K, Wallace SW. Generating scenario trees for multistage decision problems. *Manag Sci* 2001;47(2):295–307. doi:[10.1287/mnsc.47.2.295.9834](https://doi.org/10.1287/mnsc.47.2.295.9834).
- [19] Høyland K, Kaut M, Wallace SW. A heuristic for moment-matching scenario generation. *Comput Optim Appl* 2003;24(2):169–85. doi:[10.1023/A:1021853807313](https://doi.org/10.1023/A:1021853807313).
- [20] Heitsch H, Römisch W. Scenario tree modeling for multistage stochastic programs. *Math Program* 2009;118(2):371–406. doi:[10.1007/s10107-007-0197-2](https://doi.org/10.1007/s10107-007-0197-2).
- [21] Beraldi P, De Simone F, Violi A. Generating scenario trees: A parallel integrated simulation-optimization approach. *J Comput Appl Math* 2010;233(9):2322–31. doi:[10.1016/j.cam.2009.10.017](https://doi.org/10.1016/j.cam.2009.10.017).
- [22] Kaut M, Wallace SW. Shape-based scenario generation using copulas. *Comput Manag Sci* 2011;8(1-2):181–99. doi:[10.1007/s10287-009-0110-y](https://doi.org/10.1007/s10287-009-0110-y).
- [23] Kaut M. A copula-based heuristic for scenario generation. *Comput Manag Sci* 2014;11(4):503–16. doi:[10.1007/s10287-013-0184-4](https://doi.org/10.1007/s10287-013-0184-4).
- [24] Birge JR. Decomposition and partitioning methods for multistage stochastic linear programs. *Oper Res* 1985;33(5):989–1007. doi:[10.1287/opre.33.5.989](https://doi.org/10.1287/opre.33.5.989).
- [25] Beltran-Royo C, Escudero LF, Monge JF, Rodriguez-Ravines RE. An effective heuristic for multistage linear programming with a stochastic right-hand side. *Comput Oper Res* 2014;51:237–250. doi:[10.1016/j.cor.2014.06.010](https://doi.org/10.1016/j.cor.2014.06.010).
- [26] Jorion P. Value at risk: The new benchmark for managing financial risk. 3rd ed. McGraw-Hill New York; 2007.
- [27] White H. A reality check for data snooping. *Econometrica* 2000;68(5):1097–126. doi:[10.1111/1468-0262.00152](https://doi.org/10.1111/1468-0262.00152).
- [28] Elliot Tanis RVH. Probability and statistical inference. 8th ed. Pearson; 2009.

- [29] Watkins DS. Fundamentals of matrix computations. 2nd ed. John Wiley & Sons; 2004. doi:[10.1002/0471249718](https://doi.org/10.1002/0471249718).
- [30] Klaassen P. Comment on generating scenario trees for multistage decision problems. *Manag Sci* 2002;48(11):1512–16. doi:[10.1287/mnsc.48.11.1512.261](https://doi.org/10.1287/mnsc.48.11.1512.261).
- [31] Taiwan Economic Journal Co., Ltd, TEJ. <http://www.tej.com.tw/twsite/>.
- [32] Hart WE, Watson J-P, Woodruff DL. Pyomo: modeling and solving mathematical programs in python. *Math Program Comput* 2011;3(3):219–60. doi:[10.1007/s12532-011-0026-8](https://doi.org/10.1007/s12532-011-0026-8).
- [33] Sharpe WF. Mutual fund performance. *J Bus* 1966;39(1):119–38. doi:[10.1086/294846](https://doi.org/10.1086/294846).
- [34] Sortino FA, Price LN. Performance measurement in a downside risk framework. *J Invest* 1994;3(3):59–64. doi:[10.3905/joi.3.3.59](https://doi.org/10.3905/joi.3.3.59).
- [35] Chaudhry A, Johnson HL. The efficacy of the sortino ratio and other benchmarked performance measures under skewed return distributions. *Australian J Manag* 2008;32(3):485–502. doi:[10.1177/031289620803200306](https://doi.org/10.1177/031289620803200306).
- [36] Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ Lett* 1980;6(3):255–9. doi:[10.1016/0165-1765\(80\)90024-5](https://doi.org/10.1016/0165-1765(80)90024-5).
- [37] Said SE, Dickey DA. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 1984;71(3):599–607. doi:[10.1093/biomet/71.3.599](https://doi.org/10.1093/biomet/71.3.599).
- [38] Wei WWS. Time series analysis: Univariate and multivariate methods. 2nd ed. Pearson Addison Wesley; 2006. ISBN 9780321322166.
- [39] Chen H-H, Yang C-B. Stock investment with stochastic programming. In: Proceedings of the 13th conference on information technology and applications in outlying islands. Pingtung, Taiwan; 2014.