CIRJE-F-1022

A Return Prediction-based Investment with Particle Filtering and Anomaly Detection

Masafumi Nakano
Graduate School of Economics, The University of Tokyo

Akihiko Takahashi
The University of Tokyo

Soichiro Takahashi
Graduate School of Economics, The University of Tokyo

September 2016

A Return Prediction-based Investment with Particle Filtering and Anomaly Detection *

Masafumi Nakano, Akihiko Takahashi, Soichiro Takahashi †

September 1, 2016

Abstract

This paper proposes a new stochastic volatility model with time-varying expected return, which enables us to predict returns based on exponential moving averages of the past returns frequently used in practice. Particularly, exploiting a particle filter in a self-organizing state space framework, we demonstrate that a simple return prediction-based strategy is superior to well-known strategies such as equally-weighted, minimum-variance and risk parity portfolios, which do not depend on return prediction.

In addition, we develop three types of anomaly detectors that are easily implemented in the algorithm of the particle filter and apply them to investment decision. As a result, our model robustly outperforms the exponential moving average.

Our dataset is monthly total returns of global assets such as stocks, bonds and REITs, and investment performances are evaluated with various statistics such as compound returns, Sharpe ratios, Sortino ratios or drawdowns.

Keywords: return prediction, particle filtering, anomaly detection, exponential moving averages, stochastic volatility, state space models

*The authors are grateful to Takahiko Suenaga and Takeshi Hakamada at GCI Asset Management Inc. for valuable comments.

†Graduate School of Economics, University of Tokyo

1 Introduction

To predict asset returns is one of the most important issues in finance. For example, Lettau & Ludvigson (2001) state "it is now widely accepted that excess returns are predictable by variables such as dividend-price ratios, earnings-price ratios, dividend-earnings ratios, and an assortment of other financial indicators." In contrast, a comprehensive empirical study by Welch & Goyal (2008) reports that predictor variables of the equity premium suggested by the academic literature predict poorly both in-sample and out-of-sample.

On the other hand, in Bayesian time series analysis the prediction of volatility and correlation is a major topic and in particular there are numerous studies on stochastic volatility models (e.g., Taylor (1986)) which are estimated by Markov chain Monte Carlo methods or sequential Monte Carlo methods known as particle filters because of non-linearity of the models. In an application to practical investment problems, factor stochastic volatility models (e.g., Aguilar, Prado, Huerta, & West (1998) and Pitt & Shephard (1999)) are often used, which make us possible to estimate correlation between high-dimensional asset returns by a small number of factors (e.g., Aguilar & West (2000) and Zhou, Nakajima, & West (2014)).

In terms of investment in practice, it seems to be much more important to predict returns than volatility in spite of its difficulty. According to Chopra & Ziemba (1993), in a mean-variance portfolio, estimation errors in mean are at least 10 times as important as those in variance. Johannes, Korteweg, & Polson (2014) discuss returns and volatility prediction simultaneously. They provide a stochastic volatility model with a return predictor of payout yield, and consider one risky asset portfolio based on expected returns and volatility estimated by a particle filter. As a result, they point out that it improves portfolio performance to incorporate time-varying properties of both of expected returns and volatility into a model.

While there are a lot of academic works about return predictions, moving averages (MAs) of past returns are often used in practice for its simplicity of implementation. In particular, exponential moving averages (EMAs) have a feature that they put higher weight on more recent returns, which are similar to the estimates of expected returns based on a filtering method because it sequentially modifies the distribution by the most recent data. Then, in order to justify a filtering approach to return prediction from a practical point of view, it seems necessary to outperform those practically used simple methods such as EMAs.

For these reasons, we develop a new stochastic volatility model with time-varying expected return which represents the EMA dynamics by extending a famous AR(1) process. Then, using a particle filtering method in a self-organizing state-space framework, we sequentially estimate states and parameters of our models and predict asset return, which is applied to an investment problem.

In order to clarify the improvement effects by our approach, we test a simple return prediction-based strategy. Roughly speaking, it is the strategy that if there is an asset which has the highest positive expected return, an investor bets all on it, and if all predicted returns are negative, he/she does not bet on risky assets at all. We confirm this return prediction-based strategy achieves higher performance than the well-known strategies such as equally-weighted, minimum-variance, and risk parity portfolio.

In addition, we introduce three anomaly detection methods which can be easily implemented in the algorithm of the particle filter. Anomaly detection seems useful though there are few application examples in financial literature. In our investment case, realized asset returns sometimes largely deviate from the models, whence it is inappropriate to implement predictions. Therefore, we can use it effectively to exclude the assets for which anomalies are detected. As a result, we discover the return prediction-based strategy with the anomaly detectors robustly outperforms the same strategy based on the EMA.

When using the EMAs of past returns, one of the most practically important issues is the way of deciding a smoothing factor which represents the degree of weighting decrease. We provide two solutions to this problem. One is resorting to estimation by a particle filter in a self-organizing state space framework. The other is aggregating models for various fixed smoothing factors. With the second approach, we avoid using an statistical optimal value because it does not necessarily indicate the optimality in terms of investment performance. Rather, we prospect the so-called ensemble learning effect known in the field of machine learning by combining various cases. The effectiveness of these new approaches is also checked.

Our investment universe is designed by U.S. and Japanese REIT as well as international bonds and equities with a riskless asset. Taking trading costs into account, investment performance is evaluated by various statistics such as compound returns, standard deviation, downside deviation, Sharpe ratios, Sortino ratios, maximum drawdowns, and average drawdowns.

The remainder of the paper is organized as follows. Section 2 introduces the models that represent the dynamics of asset returns. Section 3 provides the state space representation of the models and the algorithm of our particle filter with anomaly detection. Section 4 describes the basic setting in estimation and the results. After explaining our investment strategy and setup, Section 5 reports investment performances. Finally, Section 6 concludes.

2 Model

In this paper, we suppose that the dynamics of asset returns $y = \{y_t \ ; \ t = 0, 1, \cdots, T\}$ are specified as follows.

$$
\begin{aligned}
&y_t = \mu_t + \exp(x_t/2)\epsilon_t, \quad \epsilon_t \sim i.i.d.\ N(0,1), \quad t \geq 0, \\
&x_t = \bar{x} + \phi_x(x_{t-1} - \bar{x}) + \sigma_x \xi_t, \quad \xi_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\
&x_0 \sim N(\bar{x}, \sigma_x^2/(1-\phi_x^2)),
\end{aligned}
$$

where $\bar{x}$, ϕ_x, and σ_x are constant unknown parameters, $|\phi_x| < 1$, $Cov(\epsilon_t, \eta_t) = 0$, and $\mu = \{\mu_t \ ; \ t = 0, \cdots, T\}$ and $x = \{x_t \ ; \ t = 0, \cdots, T\}$ are stochastic processes which represent the dynamics of expected returns and volatility. Clearly, this belongs to a class of notable stochastic volatility models (e.g., Taylor (1986)) with expected returns $\{\mu_t\}_t$. As for μ_t, one of the most popular modelings in time series analysis is AR(1) :

$$\begin{aligned} &\mu_t = \bar{\mu} + \phi_\mu(\mu_{t-1} - \bar{\mu}) + \sigma_\mu \eta_t, \quad \eta_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\ &\mu_0 \sim N(\bar{\mu}, \sigma_\mu^2/(1-\phi_\mu^2)), \end{aligned} \tag{2.1}$$

where $\bar{\mu}$, ϕ_μ, and σ_μ are constant unknown parameters, $|\phi_\mu| < 1$, and $Cov(\epsilon_t, \eta_t) = Cov(\xi_t, \eta_t) = 0$.

On the other hand, when fund managers estimate expected returns in practice, moving averages (MAs) of asset returns are often used, and in particular exponential moving averages (EMAs) are considered to be an effective alternative. Here we define EMAs $\bar{\mu} = \{\bar{\mu}_t \ ; \ t = 0, \cdots, T\}$ of asset returns $y = \{y_t \ ; \ t = 0, \cdots, T\}$ as follows.

$$\begin{aligned} &\bar{\mu}_t = \beta y_{t-1} + (1-\beta)\bar{\mu}_{t-1}, \quad t \geq 1, \\ &\bar{\mu}_0 = \alpha \in \mathbb{R}, \end{aligned} \tag{2.2}$$

where $\beta \in (0,1)$ is called a constant smoothing factor which represents the degree of weighting decrease and α is a constant, that will be specified later.

Now let us introduce the model of μ based on the EMA.

$$\begin{aligned} &\mu_t = y_{t-1} + \phi_\mu(\mu_{t-1} - y_{t-1}) + \sigma_\mu \eta_t, \quad \eta_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\ &\mu_0 \sim N(\alpha, \sigma_\mu^2/(1-\phi_\mu^2)), \end{aligned}$$

where we set $E[\mu_0] = \alpha$ so as to be consistent with the definition of $\bar{\mu}_0$. Obviously this process is obtained from setting $\bar{\mu} = y_{t-1}$ in AR(1) (2.1), and it coincides with the above-defined EMA $\bar{\mu}$ if we exclude the noise term $\sigma_\mu \eta_t$ and put $\beta = 1 - \phi_\mu$. In other words, we introduce the expected return μ model with the EMA dynamics by extending a well-known AR(1) process. Making use of this model, we attempt to overcome the investment strategies based on estimates by the simple EMA in the following.

Here we emphasize the significance of our new modeling. By expressing the practically used EMA as a time-series model, we can get access to a variety of statistical tools, which brings us great benefits. Anomaly detection is one of the most impressive examples. In addition, we demonstrate later two novel statistical solutions to the problem that we are not able to know in advance an optimal smoothing factor β.

Our investment universe includes not only stocks or REITs whose returns change sharply over time but also bonds whose cumulative returns grow stably. Therefore we also exploit a constant μ model in order to express their dynamics well, that is, a case of $\mu_t \equiv \bar{\mu}$.

As the closing part of this section, we summarize and name the models used below.

- CMSV (Constant Mean and Stochastic Volatilily):

$$
\begin{aligned}
&y_t = \bar{\mu} + \exp(x_t/2)\epsilon_t, \quad \epsilon_t \sim i.i.d.\ N(0,1), \quad t \geq 0, \\
&x_t = \bar{x} + \phi_x(x_{t-1} - \bar{x}) + \sigma_x \xi_t, \quad \xi_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\
&x_0 \sim N(\bar{x}, \sigma_x^2/(1-\phi_x^2)),
\end{aligned}
\tag{2.3}
$$

- SMSV (Stochastic Mean and Stochastic Volatility):

$$
\begin{aligned}
&y_t = \mu_t + \exp(x_t/2)\epsilon_t, \quad \epsilon_t \sim i.i.d.\ N(0,1), \quad t \geq 0, \\
&x_t = \bar{x} + \phi_x(x_{t-1} - \bar{x}) + \sigma_x \xi_t, \quad \xi_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\
&\mu_t = \bar{\mu} + \phi_\mu(\mu_{t-1} - \bar{\mu}) + \sigma_\mu \eta_t, \quad \eta_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\
&x_0 \sim N(\bar{x}, \sigma_x^2/(1-\phi_x^2)), \quad \mu_0 \sim N(\bar{\mu}, \sigma_\mu^2/(1-\phi_\mu^2)),
\end{aligned}
\tag{2.4}
$$

- SMSV+EMA:

$$
\begin{aligned}
&y_t = \mu_t + \exp(x_t/2)\epsilon_t, \quad \epsilon_t \sim i.i.d.\ N(0,1), \quad t \geq 0, \\
&x_t = \bar{x} + \phi_x(x_{t-1} - \bar{x}) + \sigma_x \xi_t, \quad \xi_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\
&\mu_t = y_{t-1} + \phi_\mu(\mu_{t-1} - y_{t-1}) + \sigma_\mu \eta_t, \quad \eta_t \sim i.i.d.\ N(0,1), \quad t \geq 1, \\
&x_0 \sim N(\bar{x}, \sigma_x^2/(1-\phi_x^2)), \quad \mu_0 \sim N(\alpha, \sigma_\mu^2/(1-\phi_\mu^2)),
\end{aligned}
\tag{2.5}
$$

where $\bar{x}$, $\bar{\mu}$, ϕ_x, ϕ_μ, σ_x and σ_μ are constant parameters, $Cov(\epsilon_t, \eta_t) = Cov(\xi_t, \eta_t) = Cov(\epsilon_t, \xi_t) = 0$, $|\phi_\mu| < 1$, and $|\phi_\mu| < 1$.

3 Particle filtering with Anomaly Detection

3.1 Particle filtering

In general, we cannot directly observe expected returns and volatility. In order to estimate these unobservable variables, we introduce a general state space model that consists of the following system and observation model.

$$
\begin{aligned}
&Y_t = H(Z_t, u_t), \qquad \text{[observation model]} \\
&Z_t = F(Z_{t-1}, v_t), \qquad \text{[system model]}
\end{aligned}
\tag{3.6}
$$

where Z_t denotes a n-dimensional unobservable state vector and Y_t denotes a m-dimensional observation vector at time t. $H : R^n \times R^m \to R^m$ and $F : R^n \times R^n \to R^n$ are non-linear functions in general. u_t and v_t denote the observational noise and the system noise respectively.

In our models, (2.3)-(2.5), we can easily apply this general state space model by regarding the first equation as the observation model and the other equations as the system model with $Y_t = y_t$ and $Z_t = (\mu_t, x_t)$ ($Z_t = x_t$ in CMSV (2.3)).

We utilize Monte Carlo filter (MCF) developed by Kitagawa (1996) for state estimation. Given the above state space representation, the algorithm of MCF is described as follows.

the algorithm of MCF

1. Generate the initial state vector $\{f_0^{(1)}, \cdots, f_0^{(L)}\}$ where L is the number of particles.
2. Apply the following steps (a)∼(d) to each time $t = 1, \cdots, T$.
   (a) Generate system noise $v_t^{(j)}$, $j = 1, \cdots, L$.
   (b) Compute for each $j = 1, \cdots, L$ $p_t^{(j)} = F(f_{t-1}^{(j)}, v_t^{(j)})$.
   (c) Evaluate the weights of particles $\{p_t^{(1)}, \cdots, p_t^{(L)}\}$ using likelihood function as $\alpha_t^{(j)} \equiv p(Y_t|p_t^{(j)})$, $j = 1, \cdots, L$.
   (d) Resample $\{f_t^{(1)}, \cdots, f_t^{(L)}\}$ from $\{p_t^{(1)}, \cdots, p_t^{(L)}\}$. More precisely, resample each $f_t^{(i)}$, $i = 1, \cdots, L$ from $\{p_t^{(1)}, \cdots, p_t^{(L)}\}$ with the probability given by

$$\text{Prob.}(f_t^{(i)} = p_t^{(j)}|Y_t) = \frac{\alpha_t^{(j)}}{\sum_{k=1}^{L} \alpha_t^{(k)}}, \quad j = 1, \cdots, L.$$

Here the likelihood at time t, $p(Y_t|Z_t)$, is approximately calculated by

$$p(Y_t|Z_t) \approx \frac{1}{L} \sum_{k=1}^{L} \alpha_t^{(k)} \tag{3.7}$$

and this approximation is used in anomaly detection discussed in the next subsection.

The unknown parameters vector θ_t is sequentially estimated by augmenting the state vector as $Z_t = (\mu_t, x_t, \theta_t)$ ($Z_t = (x_t, \theta_t)$ in CMSV (2.3)). If the transition of θ_t follows

$$\theta_t = \theta_{t-1}, \tag{3.8}$$

this algorithm will degenerate in the sense that almost all of the particles quickly reach zero weight. Besides, parameter estimation does not work when the true values are not included in particles generated by initial distribution. Then, we add an artificial noise ζ_t to the equation (3.8):

$$\theta_t = \theta_{t-1} + \zeta_t.$$

Kitagawa (1998) named this framework as a "self-organizing state space model". Here, it is necessary to specify the distribution of the artificial noise ζ_t conditioned on θ_{t-1}.

In this paper, we use the Kernel Smoothing (KS) method developed by West (1993a,1993b) and Liu & West (2001). In the KS method, the distribution of ζ_t conditioned on θ_{t-1} is given by

$$p(\zeta_t|\theta_{t-1}) \sim N((a-1)(\theta_{t-1} - \bar{\theta}_{t-1}), (1-a^2)V_{t-1}),$$

i.e. the conditional distribution of θ_t is

$$p(\theta_t|\theta_{t-1}) \sim N(a\theta_{t-1} + (1-a)\bar{\theta}_{t-1}, (1-a^2)V_{t-1}),$$

where $a = (3\delta - 1)/2\delta$. $\bar{\theta}_t$ and V_t represent the mean and variance of the particles $\{\theta_t^i\}_{i=1\cdots L}$ respectively. δ is a shrinkage factor which usually takes 0.95-0.99. Here, we set $\delta = 0.98$.

We adopt this algorithm based on the following reasons. It is well known that there are mainly two types of algorithms in particle filters (PF): one which first implements one-step-ahead prediction and then resampling at each time step, and the other which first implements resampling. For example, Gordon, Salmond, & Smith (1993) and Kitagawa (1996) developed the former algorithm and they are called Bootstrap filter (BF) and Monte Carlo filter (MCF) respectively. On the other hand, Pitt & Shephard (1999) proposed the latter algorithm to estimate more efficiently, that is called Auxiliary Particle filter (APF). In the above-mentioned Liu & West (2001), APF was combined with the Kernel Smoothing method, which is known as Liu and West filter (LWF). Taking into consideration these facts, our algorithm can be regarded as the particle filter which combines MCF and KS. This algorithm is easier to implement than APF or LWF because it is based on MCF. Also, numerical experiments about Markov Switching SV model in Rios & Lopes (2013) shows that the estimation error and computational complexity of our scheme is no less than those of LWF. For these reasons, our algorithm seems to be valid.

3.2 Anomaly Detection

We test three types of anomaly detectors based on previous researches such as Chang (2014), Cai, Hong, Wu, & Liu (2013), Patil, Das, & Pecht (2012) and Knorn & Leith (2008). If an anomaly is detected for an asset return at time t, y_t, we exclude the asset from investment universe at time t, which makes us possible to enhance our investment performances.

First, we utilize the log-likelihoods. Here we define the log-likelihood at time t by $l(y_t) \equiv \log p(y_t|Z_t)$ which is approximately calculated with taking log of the right hand side of the equation (3.7). In our PF algorithm, we are able to obtain $\{l(y_t)\}_t$ for each asset. Then, if $l(y_t)$ takes a lower value than a predetermined threshold at time t, we regard y_t as an anomaly. Now this threshold is set to be about the 5 percentile of $\{l(y_t)\}_{t=6,\cdots,47}$ which corresponds to the second lowest value in $\{l(y_t)\}_{t=6,\cdots,47}$. We omit the first six month of the log-likelihoods because this period seems to be strongly affected by initial distributions.

Second, we employ the traditional Hotelling approach. In this approach, we define an anomaly indicator $a(y_t)$ by the negative log-likelihood $-l(y_t)$. Since the distribution of y_t conditional on states μ_t and x_t at time t is normal with our models, we rewrite $a(y_t)$ by excluding the parts indifferent with y_t as follows.

$$a(y_t) = \exp(-x_t)(y_t - \mu_t)^2. \tag{3.9}$$

Again, this can be approximately calculated in our estimation algorithm. Then the remaining task is to determine a threshold as in the first approach. Here, we make use of the fact that this $a(y_t)$ asymptotically follows a chi-squared distribution with one degrees of freedom $\chi^2(1)$. Although the number of our data is not large enough to apply this asymptotic property, it seems to be an alternative method of deciding an effective threshold. Therefore we put it as the 95 percentile of $\chi^2(1)$.

Last, we test a method by the one-step-ahead predictive distribution of asset returns, $p(y_{t+1}|y_{1:t})$ where $y_{1:t} \equiv (y_1, \cdots, y_t)$. As described in Section 4.2, in our PF algorithm we can obtain the approximation of predictive distribution $p(y_{t+1}|y_{1:t})$ based on Monte Carlo simulation. Then we are able to detect an anomaly by calculating the 2.5 and 97.5 percentiles of $p(y_{t+1}|y_{1:t})$ to exclude the realized return y_{t+1} outsides these percentiles.

In the following, we call those three anomaly detection methods AD1,AD2, and AD3 respectively. Note that all the methods are implemented in our PF algorithm quite easily.

4 Estimation

4.1 Data

We use monthly total returns of 8 indexes corresponding to stocks, bonds, and REITs as listed in Table 1. Hereafter we employ the abbreviations of the index names in this table. The time period of the return data is 156 months (i.e. $T = 155$), from April 2003 to March 2016. An asset return y_t is given by $y_t = 100 \times (P_t/P_{t-1} - 1)$ where P_t denotes the asset price at time t. Our data are downloaded from Bloomberg in JPY-denominated form so that we consider the global investment without currency hedging. Since there exist ETFs which correspond to these indexes except US Bond, it can be said that they are tradable assets. Table 2 shows the descriptive statistics of the asset returns.

Table 1: Data

Index name	Ticker (Bloomberg)	Abbreviated name
Tokyo Stock Price Index	TPXDDVD.Index	JP Equity
Tokyo Stock Exchange REIT Index	TPXDREIT.Index	JP REIT
S&P500	SPTR.Index	US Equity
Morgan Stanley REIT Index	RMS.G.Index	US REIT
FTSE Developed ex North America Net Tax (US RIC) Index	TGPVAN33.Index	Developed Equity
FTSE Emerging Total Return Index	FTS5ALEM.Index	Emerging Equity
Barclays US Treasury 10 Year TERM Index	BCEY4T.INDEX	US Bond
JPMorgan Emerging Market Bond Index	JPEIGLBL.INDEX	Emerging Bond

Table 2: Descriptive statistics

	Mean	Variance	Skew	Kurtosis
JP Equity	0.629	27.807	-0.407	0.855
JP REIT	0.953	32.729	-0.206	3.531
US Equity	0.853	27.869	-0.724	1.764
Developed Equity	0.825	34.851	-0.933	2.384
US REIT	1.193	53.847	-0.941	6.002
Emerging Equity	1.216	51.668	-0.833	2.598
US Bond	0.430	6.613	-0.181	0.720
Emerging Bond	0.726	12.153	-1.528	8.993

4.2 Estimation Setup

In Section 5, we implement an investment strategy based on our estimates of expected returns, and compare their performances with the benchmark EMA cases. Notice that when we calculate an EMA of past returns, it is necessary to decide a smoothing factor β. Here we test the cases of $\beta = 0.1, \cdots, 0.9$, and estimate states and parameters of the SMSV+EMA models under fixed ϕ_μ values which correspond to $1 - \beta$ values. Hereafter, SMSV+EMA for $\beta = 0.1, \cdots, 0.9$ are called SMSV+EMA1,$\cdots$,SMSV+EMA9 respectively. Of course, we also consider the case that β (or ϕ_μ) itself is estimated by a self-organizing state space framework and name it SMSV+EMASO.[1]

Our particle filter are executed individually for each asset. In other words, we do not pay any attention to correlations in predicting returns. It makes us possible to reduce estimation error and computational complexity though incorporating them is one of the possible future researches.

The initial state variables are drawn from their initial distributions described in Section 2. As for a constant α in the EMA (2.2) and the SMSV+EMA model (2.5), we put it the sample mean of the first two years of asset returns, $\{y_t\}_{t=0,\cdots,23}$. The initial distributions of parameters are as follows.

- CMSV : $\bar{\mu} \sim U(-5, 5),\ \bar{x} \sim U(-1, 5),\ \frac{\phi_x+1}{2} \sim B(20, 1.5),\ \sigma_x \sim U(0, 2)$,
- SMSV : $\bar{\mu} \sim U(-5, 5),\ \bar{x} \sim U(-1, 5),\ \phi_\mu \sim U(0, 1),\ \frac{\phi_x+1}{2} \sim B(20, 1.5),\ \sigma_x \sim U(0, 2)$,
- SMSVSO+EMA : $\bar{x} \sim U(-1, 5),\ \phi_\mu \sim U(0, 1),\ \frac{\phi_x+1}{2} \sim B(20, 1.5),\ \sigma_x \sim U(0, 2)$,
- SMSV+EMA1,$\cdots$,SMSV+EMA9 : $\bar{x} \sim U(-1, 5),\ \frac{\phi_x+1}{2} \sim B(20, 1.5),\ \sigma_x \sim U(0, 2)$,

where $U(a, b)$ denotes the uniform distribution on the open interval (a, b), and $B(a, b)$ denotes the beta distribution with shape paremeters a and b. With regard to a parameter σ_μ, in CMSV, SMSV and SMSV+EMASO, we use the uniform distribution between zero and one standard deviation of asset returns over the first four years, $\{y_t\}_{t=0,\cdots,47}$. In SMSV+EMA1,$\cdots$,SMSV+EMA9, we use the same distribution between zero and one standard deviation of EMAs at $\beta = 0.1, \cdots, 0.9$ over the first four years respectively.

Since we use the monthly data, the number of observations is relatively small. Therefore, we set the number of particles 1,000,000 to obtain robust estimation. Even if the number of particles is such large, it takes a few minutes to execute our algorithm for each asset. [2]

We utilize the mean of the one-step-ahead predictive distribution $p(y_t|y_{1:t-1})$ as the estimate of an expected return at time $t-1$ which is necessary for implementing our investment strategy introduced in Section 5. Remark that the mean of $p(y_t|y_{1:t-1})$ does not necessarily equal to that of $p(\mu_t|y_{1:t-1})$ at time $t-1$.

[1] "SO" in EMASO is an abbreviation of "self-organizing".

[2] We use Intel(R) Xeon(R) CPU X5675 @ 3.07GHz.

In this paper, we calculate $p(y_t|y_{1:t-1})$ as follows. First, notice that

$$p(y_t|y_{1:t-1}) \propto p(y_t|Z_{t-1})p(Z_{t-1}|y_{1:t-1}) = p(y_t|Z_t)p(Z_t|Z_{t-1})p(Z_{t-1}|y_{1:t-1}). \quad (4.10)$$

Since $p(Z_{t-1}|y_{1:t-1})$ is the filter distribution at time $t-1$, we can approximate $p(y_t|y_{1:t-1})$ based on Monte Carlo simulation by adding the new operations (b') and (b") between (b) and (c) in the previous section. Given the samples $\{f_{t-1}^{(j)}\}_{j=1,2,\cdots,L}$ from $p(Z_{t-1}|y_{1:t-1})$ at time $t-1$, the operations (b') and (b") are the following:

(b') Generate observational noise $\{u_t^{(j)}\}_{j=1,\cdots,L}$,

(b") Using the observation model, calculate $y_t^{(j)} = H(p_t^{(j)}, u_t^{(j)})$ for each $j = 1, \cdots, L$.

Here $\{u_t^{(j)}\}_{j=1,\cdots,L}$ may be correlated to $\{v_t^{(j)}\}_{j=1,\cdots,L}$ in general, and $y_t^{(j)}$ can be regarded as samples from distribution $p(y_t|Z_t)$. The estimate of the expected return at time $t-1$ is given by calculating the mean of the particles $\{y_t^{(j)}\}_{j=1,\cdots,L}$ which are obtained from these steps.

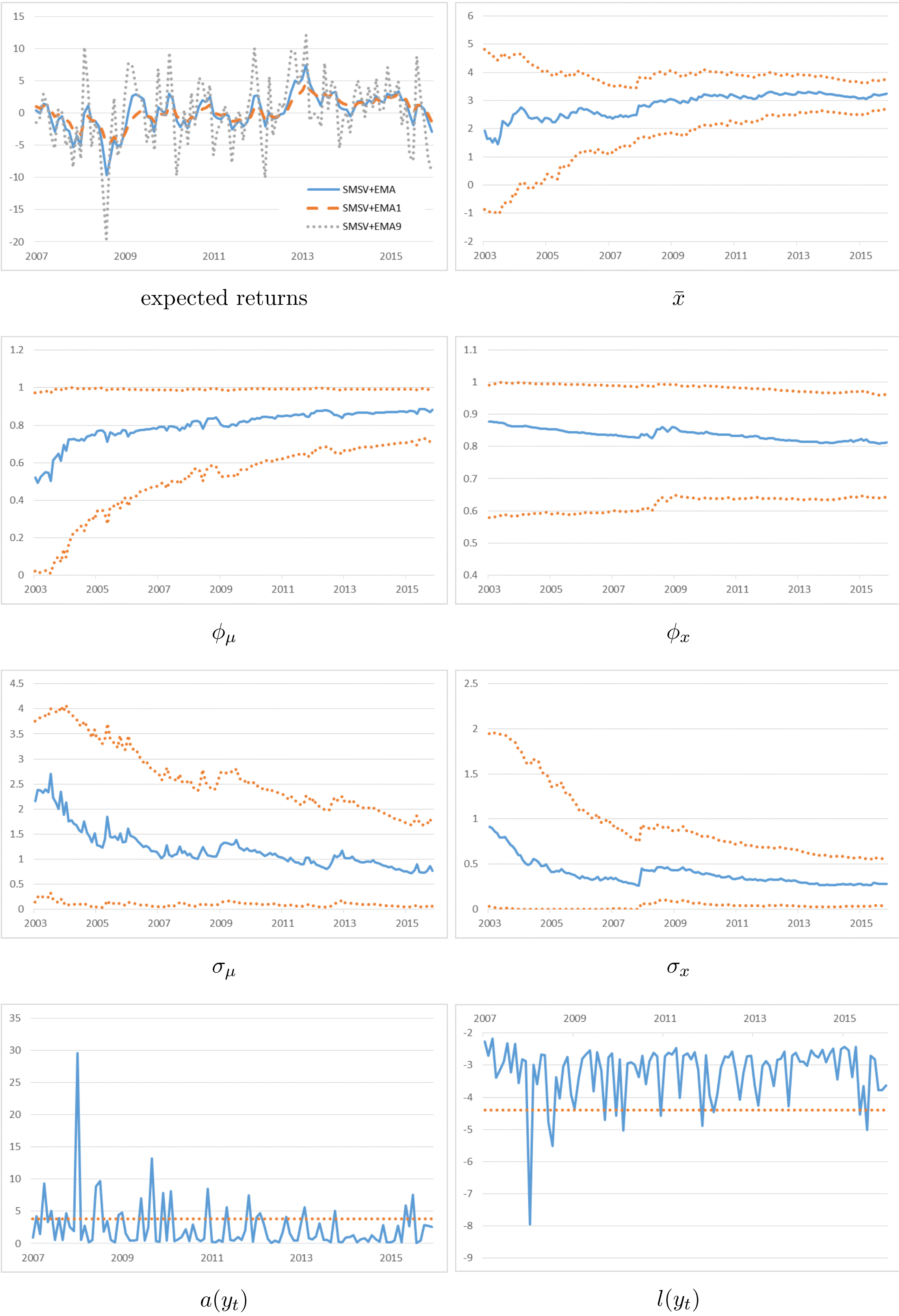
4.3 Estimation Result

Figure 1 displays the estimation results of JP Equity in the SMSV+EMASO model. For parameter learning, although there exists a parameter whose learning speed is slow such as ϕ_x, our sequential estimation is likely to perform properly as a whole. Here the solid line and the dot lines denote posterior medians and 95% credible intervals in each panel of parameters. From these results, we regard the first four years $t = 0, \cdots, 47$ as the learning period, and set the starting point of investment $t = 47$. Then as for other estimation results—expected returns, $a(y_t)$ and $l(y_t)$—we plot just over $t = 48, \cdots, 155$ because it is misleading to describe the values before this period in terms of reliability for estimation.

In the figure of expected returns, the cases of SMSV+EMA1 and SMSV+EMA9 are also illustrated for comparison. The path of SMSV+EMASO is similar to that of SMSV+EMA1 rather than SMSV+EMA9, which is consistent with the result of β. That is, a smoothing factor β which corresponds $1 - \phi_\mu$ is estimated in low range, 0.1-0.3.

Last, we check anomaly indicator $a(y_t)$ and $l(y_t)$. In the both plots, dot lines represent AD thresholds, and so it means anomaly for $a(y_t)$ or $l(y_t)$ to be outside these lines. Our anomaly indicators seem to work as intended.

Figure 1: Estimation results of JP Equity in the SMSV+EMASO model



5 Investment

5.1 Multi-Bang-Bang strategy

As stated in the preceding section, we start investment at $t = 47$. Then in the following, we reset a time index t which starts from the 48th data with $T = 108$ $(= 155 - 47)$ for the simplicity of notation.

Suppose that there exists a risk-free asset in the financial market and the risk-free rate is zero. In order to clarify the effects of using the particle filtering method, let us consider the following simple multi-Bang-Bang strategy only based on the estimates of expected returns $\boldsymbol{m}_t = (m_{t,1}, \cdots, m_{t,N})'$, where $m_{t,i}$ denotes the one for the i-th risky asset. Roughly speaking, it is the strategy that an investor holds the asset which has the highest expected return unless it is negative. We also put no-short-sale constraint.

Multi-Bang-Bang strategy

Apply the following steps to each time $t = 0, 1, \cdots, T-1$.

1. Find an index $i_{max} \in \{1, \cdots, N\}$ such that $m_{t,i_{max}} = \max\{m_{t,1}, \cdots, m_{t,N}\}$.
2. Risky asset weights $\{\omega_{t,i}\}_{i=1,\cdots,N}$ are decided based on the following rule :

   i) if $m_{t,i_{max}} > 0$,

$$\omega_{t,i} = \begin{cases} 1 \quad , \quad i = i_{max}, \\ 0 \quad , \quad i \neq i_{max}. \end{cases}$$

   ii) otherwise,

$$\omega_{t,i} = 0 \quad , \quad i \in \{1, \cdots, N\}.$$

3. Put a risk-free asset weight $\omega_{t,N+1} = 1 - \sum_{i=1}^{N} \omega_{t,i}$.

As stated before, if we apply anomaly detection, the assets for which anomalies are detected are excluded from the above index set $\{1, \cdots, N\}$.

5.2 Investment Setup

Since we cannot neglect transaction costs which arise from portfolio weight update in practice, we define the portfolio value $\{V_t\}_{t=0,\cdots,T}$ and the portfolio return $\{R_t\}_{t=1,\cdots,T}$ with transaction costs as follows.

$$\begin{aligned} V_{t+1} &= V_t(1 + \boldsymbol{\omega}_t{}' \boldsymbol{y}_{t+1}) - \sum_{i=1}^{N} c_i |\omega_{t,i} V_t - \omega_{t-1,i} V_{t-1}(1 + y_{t,i})| \ , \quad V_0 = 1, \\ R_{t+1} &= \frac{V_{t+1}}{V_t} - 1, \end{aligned} \tag{5.11}$$

where $\boldsymbol{y}_t \equiv (y_{t,1}, \cdots, y_{t,N}, 0)$, $\boldsymbol{\omega}_t \equiv (\omega_{t,1}, \cdots, \omega_{t,N}, \omega_{t,N+1})$, and c_i denotes the transaction spread. In the return vector $\boldsymbol{y}_t$, $y_{t,i}$ indicates the i-th risky asset return at time t, and the last element 0 represents the risk-free rate. As there exist ETFs corresponding to the indexes, c_i can be interpreted as a bid-ask spread. Then transaction cost for an asset i equals to c_i times the additional amount of money which is necessary for taking new portfolio position. Here, we assume that c_i is 10 *bp* for all risky assets and there is no transaction for the riskless asset.

The second term of equation (5.11) means the transaction cost at time t. We introduce this penalty term based on the following ideas. $\omega_{t-1,i}V_{t-1}(1 + y_{t,i})$ and $\omega_{t,i}V_t$ indicate the value of the i-th risky asset before and after the position change at time t respectively as $\omega_{t,i}$ means a portfolio weight of the i-th risky asset during $[t, t + 1)$. That is, $|\omega_{t,i}V_t - \omega_{t-1,i}V_{t-1}(1 + y_{t,i})|$ represents the necessary amount of money for the position change of asset i at time t. Hence, the total transaction cost at time t equals to the sum of $c_i|\omega_{t,i}V_t - \omega_{t-1,i}V_{t-1}(1 + y_{t,i})|$ for all i.

We evaluate the portfolio performances from various statistics: the cumulative return, the compound return, the standard deviation, the downside deviation, the Sharpe ratio, Sortino ratio, the maximum drawdown, and the average drawdown. We calculate these performance statistics on annual basis except the cumulative return, the maximum drawdown and the average drawdown.

- Compound Return: This is the annualized geometric mean of the portfolio returns $\{R_t\}$.

$$\mathit{Compound\ Return} \equiv \left\{ \prod_{t=1}^{T} (1 + R_t) \right\}^{12/T} - 1.$$

- Standard Deviation:

$$\mathit{Standard\ Deviation} \equiv \left\{ \frac{12}{T} \sum_{t=1}^{T} (R_t - \bar{R})^2 \right\}^{1/2},$$

$$\bar{R} \equiv \frac{1}{T} \sum_{t=1}^{T} R_t.$$

- Downside Deviation:

$$\mathit{Downside\ Deviation} \equiv \left\{ \frac{12}{T} \sum_{t=1}^{T} \min(0, R_t)^2 \right\}^{1/2}.$$

- Sharpe Ratio: This is usually defined as portfolio excess return divided by portfolio standard deviation. Since a risk-free rate is zero in this paper, we use the following value as the Sharpe ratio.

$$\mathit{Sharpe\ Ratio} \equiv \mathit{Arithmetic\ Return}/\mathit{Standard\ Deviation},$$
$$\mathit{Arithmetic\ Return} \equiv 12\bar{R}.$$

- Sortino Ratio: This ratio does not regard upside volatility as a risk and penalizes only downside volatility while the Sharpe ratio penalizes both upside and downside volatility equally, which is often pointed out as a weakness of the Sharpe Ratio.

$$\mathit{Sortino\ Ratio} \equiv \mathit{Arithmetic\ Return}/\mathit{Downside\ Deviation}.$$

- Maximum Drawdown, Average Drawdown: The drawdown is the decline from the past peak value to the present value. The maximum drawdown and the average drawdown are defined as follows.

$$Maximum\ Drawdown \equiv \max_{1\le t\le T} \frac{M_t - V_t}{M_t},$$
$$Average\ Drawdown \equiv \frac{1}{T}\sum_{t=1}^{T} \frac{M_t - V_t}{M_t},$$

where $M_t \equiv \max_{0\le s\le t} V_s$. In general, portfolio performance depends on the investment timing. The maximum drawdown contributes to the performance analysis because it is independent of the investment timing given the horizon $[0, T]$.

5.3 Investment Performance

When making use of EMAs, one of the most practically important problems is how to decide a smoothing factor β which largely affects the behavior of EMA dynamics. Here we provide two approaches to this problem. One is of course to estimate β in a self-organizing state space framework. The other is not to decide a particular β value but to combine the various cases of $\beta = 0.1, \cdots, 0.9$, which we call SMSV+EMA1-9 hereafter. Precisely, in SMSV+EMA1-9, the expected return at time t of each asset is set to be the average of expected returns at time t under SMSV+EMA1,$\cdots$,SMSV+EMA9.

In applying anomaly detection to this SMSV+EMA1-9 model for each asset, we take an average by only the models for which anomalies are not detected. For example, when anomalies are detected for the models except for SMSV+EMA1 and SMSV+EMA3, we calculate the mean of SMSV+EMA1 and SMSV+EMA3 as the expected return at time t of the asset. If all of the models, SMSV+EMA1,$\cdots$,SMSV+EMA9, tell us that the asset is in anomaly, we stop to invest it at that time. That is, we intend to apply the predictor for each asset only when it seems to work well.

Table 3: Investment Performance of traditional strategies

	Compound return(%)	Standard deviation(%)	Downside deviation(%)	Sharpe ratios(%)	Sortino ratios(%)	Maximum drawdown(%)	Average drawdown(%)
Buy and Hold							
JP Equity	-0.66	19.77	14.25	6.67	9.26	56.23	32.13
JP REIT	1.81	21.98	15.34	19.62	28.12	67.56	35.30
US Equity	5.96	20.58	14.56	38.77	54.80	59.79	23.70
US REIT	4.37	28.58	20.75	30.22	41.63	73.94	27.67
Developed Equity	-0.05	23.07	17.19	11.80	15.83	63.04	30.13
Emerging Equity	1.37	27.17	19.67	19.29	26.65	67.67	28.53
US Bond	5.72	9.16	5.71	65.44	104.95	15.25	4.50
Emerging Bond	6.18	13.30	9.68	52.12	71.63	33.00	7.26
Other strategies							
Equal Weight	3.91	17.26	12.61	31.25	42.76	53.63	20.02
(Stock & REIT : Bond) = 3:7 [3]	5.39	12.25	8.75	49.25	68.93	34.69	10.71
(Stock & REIT : Bond) = 7:3 [4]	4.10	16.63	12.13	32.85	45.02	51.59	18.99
Minimum Variance	6.21	9.00	5.41	71.56	119.05	14.09	4.31
Risk Parity [5]	4.77	13.47	9.55	41.55	58.60	41.37	15.27

While Table 3 reports the investment performances of several famous strategies, Table 4 and Table 5 do the ones of our multi-Bang-Bang strategy. Note that in Table 4 we use CMSV (2.3) for the bonds, that is US Bond and Emerging Bond, as stated in Section 2. First of all, by comparing these tables, we confirm that our simple return prediction-based strategy shows much higher performances in most cases than other basic portfolios including equally-weighted, minimum-variance and risk parity, which do not depend on any return predictions.

Next let us take a close look at Table 4 and Table 5. It is observed that the performances in Table 4 are generally higher than those in Table 5, which implies the importance of model choice. When we compare EMA and SMSV+EMA for each β value (for example, EMA1 and SMSV+EMA1), we notice that although our SMSV+EMA improves the simple EMA in general, some performances get worse such as in $\beta = 0.1$. By exploiting anomaly detection, overall improvement is attained, and our models outperform EMAs in almost all cases.[6] From these results, it is safe to say that our anomaly detection based on particle filter robustly works well. This is because our anomaly detectors make it possible to refine the investment universe. Excluding the assets which deviate from the model, fund managers are able to concentrate on the more desirable investment universe.

As stated above, since we cannot know the optimal β value in advance, we try two approaches: SMSV+EMASO and SMSV+EMA1-9. In Table 5, the performances of our methods do not work much better than those of the fixed β cases. On the other hand, in Table 4 our methods interestingly show much higher performances than the most of the fixed β cases. Remember that in Table 4 we use the CMSV model for the bonds. Again, we have confirmed the importance of suitable model choice, which clearly enhances our attempts. Remark that it is reasonable to assume constant mean models for bonds as mentioned in Section 2 though we report both cases for comparison.

Comparing these two methods, SMSV+EMA1-9 works better than SMSV+EMASO. This result seems a little bit strange. It is natural that SMSV+EMA attains higher performance than SMSV+EMA1-9 because SMSV+EMASO optimizes β in terms of likelihood and SMSV+EMA1-9 just takes an average of estimates under various β models, SMSV+EMA1,$\cdots$, SMSV+EMA9. This can be interpreted as the effect of ensemble learning in the field of machine learning. The ensemble learning theory emphasizes that combining predictors enables us to reinforce predictive ability rather than using single complex model. Now we combine nine β cases so that better performance is achieved. In that sense, this result supports the effectiveness of ensemble learning even if it is implemented in such a simple way.

Furthermore, we do not know in advance which anomaly detector performs the best. Then, we also test mixtures of our three anomaly detectors in the case of SMSV+EMA1-

[3] That is, $\boldsymbol{\omega}_t \equiv (\frac{0.3}{6}, \frac{0.3}{6}, \frac{0.3}{6}, \frac{0.3}{6}, \frac{0.3}{6}, \frac{0.3}{6}, \frac{0.7}{2}, \frac{0.7}{2}, 0)$ for all t.
[4] That is, $\boldsymbol{\omega}_t \equiv (\frac{0.7}{6}, \frac{0.7}{6}, \frac{0.7}{6}, \frac{0.7}{6}, \frac{0.7}{6}, \frac{0.7}{6}, \frac{0.3}{2}, \frac{0.3}{2}, 0)$ for all t.
[5] See, for example, Bruder and Roncalli (2012)
[6] The only exception is the case of $\beta = 0.1$ in Table 5.

9, namely, ADmix1, ADmix2, and ADmix3: ADmix1 tells us anomaly if at least one detector warns it. ADmix2 does when at least two out of three detectors warn anomalies, and ADmix3 says anomaly only when all three detectors do. As a result, in Table 5 the most conservative method (i.e. ADmix1) provides the best performance, while in Table 4 we observe little difference in three methods.

Overall, our model, SMSV+EMA1-9 with ADmix1 that is, SMSV with no bias on a smoothing factor and a conservative mixture of anomaly detections seems to create robust performance in our numerical experiments.

6 Conclusion

In this paper, we have proposed a new stochastic volatility model, which enables to predict returns based on the exponential moving averages of asset returns. Specifically, in the framework of a self-organizing state space model, we have estimated states and parameters of our models by the particle filtering method. In addition, we developed three anomaly detectors which judge whether the models really express the dynamics of asset returns.

We apply these return prediction and anomaly detection methods to the investment problem. In order to clarify their effects, we employed the simple investment strategy and assessed performances by various statistics. As a result, our approach has turned out to attain higher performances than the same strategy based on the exponential moving averages.

Moreover, we have shown our investment scheme outperforms other strategies such as equally-weighted, minimum-variance and risk parity portfolios which do not depend on the estimates of expected returns. Further, in order to overcome a practically well-known issue how to determine a smoothing factor in the exponential moving average, this paper has presented two solutions which properly work.

Although we have focused on the simple investment strategy only based on expected returns to clarify the effectiveness of our scheme for predicting returns, it seems valuable to apply it to other strategies requiring estimates of higher order moments (e.g. mean-variance portfolios), which is one of our future research topics.

Table 4: Investment Performances of multi-Bang-Bang strategy: the bond model is CMSV

	Compound return(%)	Standard deviation(%)	Downside deviation(%)	Sharpe ratios(%)	Sortino ratios(%)	Maximum drawdown(%)	Average drawdown(%)
SMSV	5.41	25.34	17.16	33.99	50.20	62.21	21.03
SMSV with AD1	4.00	25.84	18.18	28.90	41.07	64.50	26.22
SMSV with AD2	8.27	17.31	8.74	54.16	107.24	34.98	5.37
SMSV with AD3	5.65	25.82	17.72	34.98	50.96	64.50	23.51
EMA1	5.02	21.24	13.63	33.63	52.41	34.39	13.13
SMSV+EMA1	4.91	22.22	15.05	32.80	48.44	50.28	17.43
SMSV+EMA1 with AD1	5.48	21.28	13.75	35.65	55.20	44.44	15.38
SMSV+EMA1 with AD2	6.21	20.81	13.22	39.33	61.90	42.79	14.40
SMSV+EMA1 with AD3	8.11	21.53	13.56	47.00	74.64	42.25	12.49
EMA2	4.19	20.29	12.90	30.26	47.61	33.01	12.97
SMSV+EMA2	10.63	21.37	13.63	58.18	91.24	47.55	10.66
SMSV+EMA2 with AD1	10.24	19.70	11.87	59.39	98.53	43.19	12.27
SMSV+EMA2 with AD2	11.19	19.77	11.64	63.59	108.03	42.85	10.46
SMSV+EMA2 with AD3	14.36	20.52	11.88	75.81	130.98	39.08	8.79
EMA3	9.14	18.98	10.97	55.49	96.01	34.56	10.81
SMSV+EMA3	12.57	21.38	12.95	66.27	109.43	45.33	10.05
SMSV+EMA3 with AD1	14.96	19.75	11.02	80.65	144.48	36.33	9.09
SMSV+EMA3 with AD2	14.65	19.22	10.70	80.95	145.33	35.68	8.29
SMSV+EMA3 with AD3	17.47	20.47	11.00	89.15	165.87	34.83	7.92
EMA4	12.01	20.01	11.16	66.67	119.56	38.44	9.61
SMSV+EMA4	13.23	20.84	11.95	70.18	122.41	37.53	8.17
SMSV+EMA4 with AD1	15.66	19.14	9.79	85.72	167.50	26.94	6.97
SMSV+EMA4 with AD2	14.15	18.31	9.50	81.53	157.10	27.78	6.57
SMSV+EMA4 with AD3	17.89	19.89	9.79	92.87	188.57	26.94	6.70
EMA5	12.89	18.59	9.08	74.43	152.30	21.30	6.40
SMSV+EMA5	11.55	20.25	11.59	64.19	112.14	37.53	8.59
SMSV+EMA5 with AD1	15.75	18.92	9.34	86.86	175.90	26.94	6.52
SMSV+EMA5 with AD2	14.88	18.65	9.31	83.73	167.76	28.74	6.49
SMSV+EMA5 with AD3	17.97	19.68	9.34	94.00	198.00	26.94	6.11
EMA6	10.52	18.92	9.96	62.17	118.14	28.85	7.57
SMSV+EMA6	12.55	20.08	11.32	69.05	122.51	34.22	7.72
SMSV+EMA6 with AD1	17.08	18.73	9.01	93.72	194.84	23.07	5.72
SMSV+EMA6 with AD2	15.99	18.42	8.93	89.85	185.40	24.97	5.95
SMSV+EMA6 with AD3	17.04	19.10	9.02	92.02	194.88	23.07	5.66
EMA7	10.49	18.81	9.75	62.25	120.09	26.19	7.03
SMSV+EMA7	12.62	20.26	11.42	68.90	122.25	34.22	7.82
SMSV+EMA7 with AD1	15.57	19.00	9.18	85.71	177.40	25.41	6.51
SMSV+EMA7 with AD2	16.61	18.20	8.59	93.67	198.54	25.78	6.10
SMSV+EMA7 with AD3	16.97	19.32	9.17	90.86	191.40	23.07	5.87
EMA8	12.46	19.39	10.06	70.18	135.32	25.15	6.74
SMSV+EMA8	14.93	21.22	11.68	76.37	138.70	34.40	7.02
SMSV+EMA8 with AD1	18.64	20.02	9.51	95.57	201.24	23.96	5.75
SMSV+EMA8 with AD2	22.63	19.16	8.36	116.48	266.90	22.78	5.01
SMSV+EMA8 with AD3	18.32	20.06	9.51	94.04	198.27	23.96	5.86
EMA9	12.08	20.15	11.18	66.67	120.16	30.09	7.58
SMSV+EMA9	12.08	22.21	13.25	62.62	104.92	44.39	9.05
SMSV+EMA9 with AD1	18.00	20.00	9.52	92.91	195.22	23.03	6.18
SMSV+EMA9 with AD2	17.92	17.75	8.55	102.07	211.74	27.38	5.80
SMSV+EMA9 with AD3	18.12	20.02	9.52	93.31	196.31	23.03	6.18
SMSV+EMASO	13.73	20.86	12.89	72.45	117.26	47.55	10.43
SMSV+EMASO with AD1	15.42	19.61	11.11	83.18	146.80	40.36	10.60
SMSV+EMASO with AD2	13.92	18.75	10.66	79.07	139.02	40.41	9.75
SMSV+EMASO with AD3	17.68	20.05	11.12	91.56	165.16	39.08	8.13
EMA1-9	10.54	19.38	10.52	61.32	112.93	31.40	8.71
SMSV+EMA1-9	13.88	20.33	11.53	74.26	131.00	36.59	7.93
SMSV+EMA1-9 with AD1	17.70	19.40	9.39	93.86	194.03	24.09	5.82
SMSV+EMA1-9 with AD2	19.38	19.84	9.65	99.44	204.56	24.92	5.75
SMSV+EMA1-9 with AD3	18.55	19.44	9.39	97.47	201.76	25.84	5.73
SMSV+EMA1-9 with ADmix1	18.27	19.75	9.69	95.01	193.73	26.50	5.96
SMSV+EMA1-9 with ADmix2	17.48	19.44	9.41	92.74	191.64	25.84	6.17
SMSV+EMA1-9 with ADmix3	18.46	19.41	9.39	97.22	200.90	25.84	5.73

Table 5: Investment Performances of multi-Bang-Bang strategy

	Compound return(%)	Standard deviation(%)	Downside deviation(%)	Sharpe ratios(%)	Sortino ratios(%)	Maximum drawdown(%)	Average drawdown(%)
SMSV	4.74	25.30	17.16	31.44	46.35	62.21	21.64
SMSV with AD1	3.26	25.79	18.18	26.14	37.09	64.71	27.19
SMSV with AD2	9.86	14.97	5.82	69.85	179.81	22.05	2.48
SMSV with AD3	4.90	25.77	17.72	32.22	46.86	64.71	24.30
EMA1	5.02	21.24	13.63	33.63	52.41	34.39	13.13
SMSV+EMA1	4.80	20.47	13.22	33.09	51.25	38.57	15.59
SMSV+EMA1 with AD1	2.82	20.53	13.42	23.73	36.31	47.76	21.92
SMSV+EMA1 with AD2	2.69	20.51	13.46	23.10	35.21	46.11	22.06
SMSV+EMA1 with AD3	5.75	20.76	13.20	37.27	58.64	42.57	18.08
EMA2	4.19	20.29	12.90	30.26	47.61	33.01	12.97
SMSV+EMA2	10.49	19.27	11.33	61.41	104.43	33.01	9.18
SMSV+EMA2 with AD1	7.63	18.74	11.33	48.56	80.35	41.57	14.89
SMSV+EMA2 with AD2	10.95	18.52	10.39	65.36	116.48	36.60	10.60
SMSV+EMA2 with AD3	11.18	19.59	11.33	63.93	110.53	35.96	9.68
EMA3	9.14	18.98	10.97	55.49	96.01	34.56	10.81
SMSV+EMA3	10.66	19.08	10.97	62.64	108.98	34.56	10.58
SMSV+EMA3 with AD1	13.74	19.04	10.50	77.24	140.11	31.89	9.55
SMSV+EMA3 with AD2	14.40	18.34	9.79	82.67	154.94	32.57	8.88
SMSV+EMA3 with AD3	13.66	19.37	10.82	75.92	135.83	33.68	9.79
EMA4	12.01	20.01	11.16	66.67	119.56	38.44	9.61
SMSV+EMA4	14.90	19.11	9.38	82.27	167.57	21.30	6.82
SMSV+EMA4 with AD1	15.25	18.55	9.20	85.88	173.24	21.30	6.68
SMSV+EMA4 with AD2	15.49	17.52	8.41	91.09	189.69	22.97	6.02
SMSV+EMA4 with AD3	17.79	19.32	9.19	94.56	198.68	21.30	6.24
EMA5	12.89	18.59	9.08	74.43	152.30	21.30	6.40
SMSV+EMA5	12.95	18.51	8.97	74.94	154.70	21.30	6.20
SMSV+EMA5 with AD1	15.41	18.36	8.76	87.32	183.07	21.30	6.12
SMSV+EMA5 with AD2	15.79	17.55	8.20	92.41	197.92	25.27	6.49
SMSV+EMA5 with AD3	17.61	19.14	8.76	94.43	206.38	21.30	5.59
EMA6	10.52	18.92	9.96	62.17	118.14	28.85	7.57
SMSV+EMA6	10.97	18.74	9.65	64.77	125.80	26.19	6.99
SMSV+EMA6 with AD1	13.70	18.62	9.46	78.28	154.07	26.12	6.46
SMSV+EMA6 with AD2	14.56	17.75	8.77	85.55	173.15	29.78	6.82
SMSV+EMA6 with AD3	13.64	18.99	9.47	76.77	153.94	26.05	6.34
EMA7	10.49	18.81	9.75	62.25	120.09	26.19	7.03
SMSV+EMA7	10.01	19.13	10.09	59.27	112.42	32.85	8.67
SMSV+EMA7 with AD1	11.98	19.09	9.93	68.73	132.09	35.47	8.72
SMSV+EMA7 with AD2	12.77	17.99	9.20	75.80	148.22	36.33	8.91
SMSV+EMA7 with AD3	12.52	19.41	9.94	70.38	137.51	32.71	8.16
EMA8	12.46	19.39	10.06	70.18	135.32	25.15	6.74
SMSV+EMA8	12.39	19.37	10.05	69.90	134.67	25.15	6.70
SMSV+EMA8 with AD1	14.42	19.32	9.90	79.42	154.99	25.66	6.03
SMSV+EMA8 with AD2	21.06	18.98	8.39	110.55	250.10	26.55	4.81
SMSV+EMA8 with AD3	13.92	19.36	9.92	76.97	150.28	25.66	6.33
EMA9	12.08	20.15	11.18	66.67	120.16	30.09	7.58
SMSV+EMA9	12.08	20.15	11.18	66.67	120.16	30.09	7.58
SMSV+EMA9 with AD1	16.93	19.00	9.08	91.92	192.37	21.85	6.10
SMSV+EMA9 with AD2	17.32	17.80	8.97	99.00	196.32	31.15	6.69
SMSV+EMA9 with AD3	15.58	19.11	9.26	85.36	176.23	21.85	6.29
SMSV+EMASO	10.59	19.30	11.36	61.82	105.00	42.09	16.94
SMSV+EMASO with AD1	12.48	19.53	11.26	70.05	121.49	46.40	18.65
SMSV+EMASO with AD2	11.52	17.89	10.27	69.99	121.86	42.74	12.97
SMSV+EMASO with AD3	13.11	19.60	11.27	72.74	126.53	42.67	16.36
EMA1-9	10.54	19.38	10.52	61.32	112.93	31.40	8.71
SMSV+EMA1-9	11.14	19.51	10.54	63.79	118.08	32.98	9.05
SMSV+EMA1-9 with AD1	13.13	19.99	10.70	71.70	133.89	34.19	8.99
SMSV+EMA1-9 with AD2	17.05	20.40	10.73	87.55	166.37	34.19	8.24
SMSV+EMA1-9 with AD3	12.65	20.09	10.83	69.29	128.54	38.44	10.02
SMSV+EMA1-9 with ADmix1	15.81	20.30	10.76	82.55	155.73	36.35	8.73
SMSV+EMA1-9 with ADmix2	12.92	20.02	10.72	70.68	131.96	35.70	9.41
SMSV+EMA1-9 with ADmix3	12.56	20.06	10.83	69.00	127.78	38.44	10.02

References

[1] Aguilar, O., Huerta, G., Prado, R., & West, M. (1998). Bayesian inference on latent structure in time series. Bayesian Statistics, 6(1), 1-16.

[2] Aguilar, O. & West, M.(2000). Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. Journal of Business and Economic Statistics, 18, 338-57.

[3] Bruder, B., & Roncalli, T. (2012). Managing risk exposures using the risk budgeting approach. Available at SSRN 2009778.

[4] Cai, L., Hong, J., Wu, X., & Liu, E. (2013). Particle Filtering with Observation Anomaly Detection in Wireless Sensor Networks. Journal of of Computational Information Systems, 18, 7273-7280.

[5] Chang, G. (2014). Robust Kalman filtering based on Mahalanobis distance as outlier judging criterion. Journal of Geodesy, 88(4), 391-401.

[6] Chopra, V. K., & Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. Journal of Portfolio Management, Winter 1993, 19, 2

[7] Gordon, N. J., Salmond, D. J., & Smith, A. F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In Radar and Signal Processing, IEE Proceedings F (Vol. 140, No. 2, pp. 107-113). IET.

[8] Johannes, M., Korteweg, A., & Polson, N. (2014). Sequential learning, predictability, and optimal portfolio returns. The Journal of Finance, 69(2), 611-644.

[9] Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. Journal of computational and graphical statistics, 5(1), 1-25.

[10] Kitagawa, G. (1998). A self-organizing state-space model. Journal of the American Statistical Association, 1203-1215.

[11] Knorn, F., & Leith, D. J. (2008). Adaptive kalman filtering for anomaly detection in software appliances. In INFOCOM Workshops 2008, IEEE (pp. 1-6). IEEE.

[12] Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. the Journal of Finance, 56(3), 815-849.

[13] Liu, J., & West, M. (2001). Combined parameter and state estimation in simulation-based filtering. In Sequential Monte Carlo methods in practice (pp. 197-223). Springer New York.

[14] Patil, N., Das, D., & Pecht, M. (2012). A prognostic approach for non-punch through and field stop IGBTs. Microelectronics Reliability, 52(3), 482-488.

[15] Pitt, M. K., & Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters. Journal of the American statistical association, 94(446), 590-599.

[16] Rios, M. P., & Lopes, H. F. (2013). The extended Liu and West filter: Parameter learning in Markov switching stochastic volatility models. In State-Space Models (pp. 23-61). Springer New York.

[17] Taylor, S. J. (1986). Modelling Financial Time Series. New York: Wiley.

[18] West, M. (1993a). Approximating posterior distributions by mixture. Journal of the Royal Statistical Society. Series B (Methodological), 409-422.

[19] West, M. (1993b). Mixture models, Monte Carlo, Bayesian updating, and dynamic models. Computing Science and Statistics, 325-325.

[20] Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies, 21(4), 1455-1508.

[21] Zhou, X., Nakajima, J., & West, M.(2014) "Bayesian forecasting and portfolio decisions using dynamic dependent sparse factor models." International Journal of Forecasting 30.4: 963-980.