AMS518 Project

Chi-Sheng Lo

Portfolio replication of Taiwan 50 index using stochastic volatility with heavy-tailed distribution as rebalancing signal

Keywords: Mixed-integer linear programming, index tracking, state-space, Bayesian method, stochastic volatility, heavy-tailed distribution, high-dimensional time series

JEL classification: C11, C22, C61

Introduction: research objectives and overview of TW50

Two research questions

1. **Fundamental:** How well I can track the index given the objective and constraints I set?
2. **Advanced:** How can I design a more active strategy? Better?

Approaches to answer my research questions

1. Index tracking formulation with PSG -> In-sample -> Out-of-Sample -> Compare tracking error (TE) and excess return (ER)
2. Volatility model -> indicator -> signal -> portfolio rebalancing backtest

Contribution and where am I different?

1. Different approach/model/sampling in doing TW50 index tracking
2. Combine the cutting-edge of both econometrics and operations research

Snapshot of TW50 index

Index Universe	Listed companies of Taiwan Stock Exchange
Weighting Method	Free Float Market Capitalization
Base Date	2002.04.30
Launch Date	2002.10.29
Base Value	5,000
Calculation Frequency	Every 5 seconds
Number of Constituents	50
Periodic Review	March, June, September, December

Dataset

Sample: 2011.01.03 to 2021.10.4 (2641 days of price data)

Source: CMoney (Financial data company in Taiwan)

Computation

Index tracking: PSG for R

Volatility estimation: MATLAB

Model formulation/introduction for volatility estimation of HTSVM

Evolution of state-spaced stochastic volatility model

State-space by Kalman (1960) -> Stochastic volatility model by Taylor (1986) -> Moving average SVM -> SVM with heavy-tailed

Strengths of HTSVM

1. Time-varying volatility
2. Non-linear equation for observation
3. Serial dependence across time
4. Persistence of measurement
3. Extreme values

Drawback of HTSVM

It is a "black box"!!

Stochastic Volatility Model for Heavy-tailed distribution (HTSVM)

$y_t = \mu + e^{\frac{1}{2}h_t}\lambda_t^{\frac{1}{2}}\varepsilon_t,\ \varepsilon_t \sim N(0, 1)$ (1): "Observation"

$h_t = \mu_h + \phi_h(h_{t-1} - \mu_h) + \varsigma_t,\ \varsigma_t \sim (0, \sigma_h^2),\ h_1 \sim N\left(\mu_h,\ \sigma_h^2/\left(1 - \Phi_h^2\right)\right)$ (2): "State"

The model is subject to $|\phi_h| < 1$. Since HTSVM adopts the student's t distribution where if $(\lambda_t|v) \sim$ IG$(v/2, v/2)$, then $\tilde{\varepsilon}_t = \lambda_t^{\frac{1}{2}}\varepsilon_t$ has a standard student's t distribution.

The HTSVM must also allow for persistence through an MA(1) error process in which:

$y_t = \mu + u_t$ (3)

$u_t = \varepsilon_t + \psi\varepsilon_{t-1},\ \varepsilon_t \sim N\left(0, \lambda_t e^{h_t}\right)$ (4)

Subject to $\varepsilon_0 = 0$ and $|\psi| < 1$

Notations:

u: average daily return
μ_h: unconditional mean (expected value)
y: observation
h_t: log volatility
Φ_h: First-order autoregression coefficient
ψ: Moving average coefficient
σ_h^2: Variance
λ: scale mixture variable
v: degree of freedom parameter

Solution method for volatility (HTSVM) estimation

Solution Method: Chan and Hsiao (2013)

Platform: MATLAB

Numerical:

1. Bayesian

2. MCMC : Metropolis-Hastings algorithm

Prior:
$E(\mu) = 0$
$E(\psi) = 0$
$E(\mu_h) = 0$,
$E(\Phi_h) = 0.95$
$E(\sigma_h^2) = 0.02$

Posterior draws:
1. $p(u|y, h, \lambda, v,\ u_h, \Phi_h, \sigma_h^2) = p(u|y, h, \lambda)$
2. $p(h|y, \lambda, u, v, u_h, \Phi_h, \sigma_h^2) = p(h|y, \lambda, u, u_h, \sigma_h^2)$
3. $p(\lambda|y, h, u,\ u_h, \Phi_h, \sigma_h^2) = \prod_{t=1}^{T} (\lambda_t|y_t, h_t, u, v)$
4. $p(v|y, h, \lambda,\ u, u_h, \Phi_h, \sigma_h^2) = p(v|\lambda)$
5. $p(\sigma_h^2|y, h, \lambda,\ u, v, u_h, \Phi_h) = p(\sigma_h^2|h, u_h, \Phi_h)$
6. $p(u_h|y, h, , \lambda,\ u, v, \Phi_h, \sigma_h^2) = p(u_h|h, \Phi_h, \sigma_h^2)$
7. $p(\Phi_h|y, h, , \lambda,\ u, v, u_h, \sigma_h^2) = p(\Phi_h|h, u_h, \sigma_h^2)$

Metropolis-Hastings Algorithm

Set the initial value of $\delta = \delta^{(0)}$
Iterate over the state for i = 1,…, M:
Draw $\delta^{(*)}$ from $q(\delta|\delta^{(i-1)})$
Compute the acceptance probability:

$$\alpha = \alpha(\delta^{(*)}, \delta^{(i-1)}) = \frac{k(\delta^{(*)})q(\delta^{(i-1)}|\delta^{(*)})}{k(\delta^{(i-1)})q(\delta^{(*)}|\delta^{(i-1)})}$$

Decision:
Generate a draw from the uniform distribution u ~ U[0, 1].
If u $\leq \alpha$, then accept the draw and set $\delta^{(i)} = \delta^{(*)}$.
If u > α, then reject the draw and stay at the previous draw $\delta^{(i)} = \delta^{(i-1)}$.

Model formulation for index tracking

Formulation

Mixed Integer Linear Programming

Objective of the problem

$$\min_{\vec{x}} \varepsilon_{MAX}(\vec{x}) = \min_{\vec{x}} \max_{1 \le t \le T} |L_t(\vec{x})| \qquad \textbf{(1)}$$

Constraints of the problem

Cardinality constraint (restricts the number of assets in the rebalanced portfolio):

$$\sum_{i=1}^{N} \delta(x_i) \le \mathbf{K} \qquad \textbf{(2)}$$

Buy-in constraint (all non-zero positions $\ge \sigma$):

$$\sum_{i=1}^{N} \beta_{\sigma}^{+}(x_i) \le \mathbf{0} \qquad \textbf{(3)}$$

Rebalance portfolio + transaction cost constraint:

$$\sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \partial_i \left|x_i - x_i^0\right| + \mathbf{A}\sum_{i=1}^{N} \delta\left(\left|x_i - x_i^0\right|\right) \le \mathbf{C} \qquad \textbf{(4)}$$

Total transaction cost constraint:

$$\sum_{i=1}^{N} \partial_i \left|x_i - x_i^0\right| + \mathbf{A}\sum_{i=1}^{N} \delta\left(\left|x_i - x_i^0\right|\right) \le \boldsymbol{\gamma}\mathbf{C} \qquad \textbf{(5)}$$

$$x_i \ge \mathbf{0},\ \mathbf{i} = \mathbf{1}, \ldots, \mathbf{N} \qquad \textbf{(6)}$$

Where variable cost: $\sum_{i=1}^{N} \partial_i \left|x_i - x_i^0\right|$**; fixed cost:** $\mathbf{A}\sum_{i=1}^{N} \delta\left(\left|x_i - x_i^0\right|\right)$

Solution method for index tracking

Transaction cost setting

Commission: 0.1424%

Tax: 0.3%

Total transaction cost constraint is set at $76000 (after taking account of slippage cost as well)

Other modification in problem statement for PSG solver

$$\text{kpol} = \frac{1000000}{Entry\ price}\left|x_i - x_i^0\right|$$

KB = $1000000

However, total budget constraint is set at $20000000. (

Dimension change

For example: the full sample

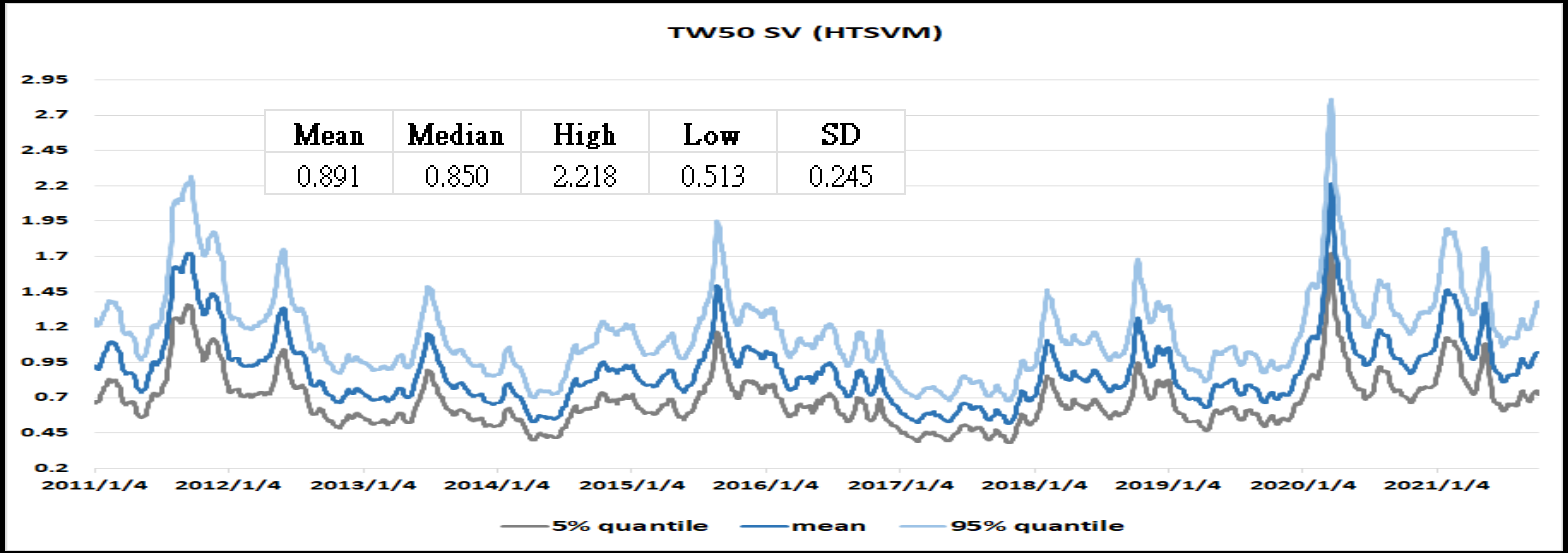
```
length(problem.list$matrix_inmmax)<-134691
dim(problem.list$matrix_inmmax)<-c(51, 2641)
```

Problem statement for PSG solver

```
problem.list$problem_statement <- sprintf (
 "
minimize
 max_risk(matrix_inmmax)
Constraint: <=17
 cardn_pos(0.01, matrix_ksi)
Constraint: <= 0
 buyin_pos(0.01, matrix_ksibuy)
Constraint: <= 20000000
 linear(matrix_ksi)
 +variable(trcost)
Constraint: <= 76000
 variable(trcost)
Constraint: <= 0
 -variable(trcost)
 +0.01*polynom_abs(matrix_ksipol)
 +100*cardn_pos(0.01, matrix_ksipol)
 +100*cardn_neg(0.01, matrix_ksipol)
Box: >= 0
Solver: precision=7, stages =30
  "
)
```

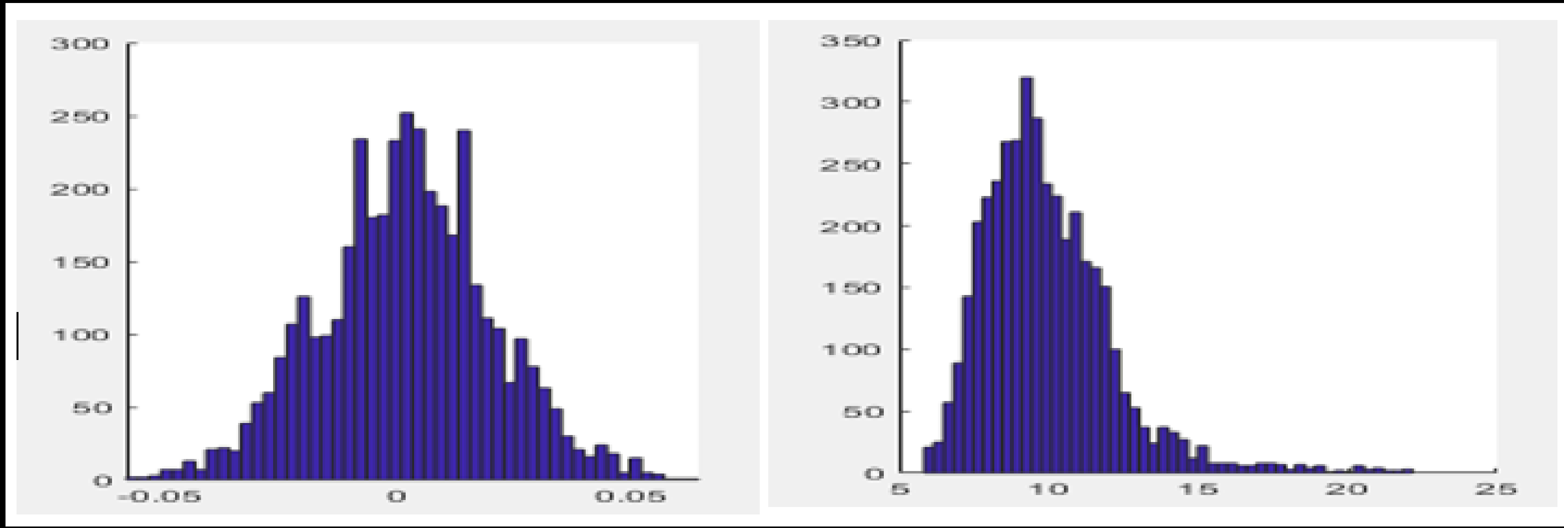
Result of HTSVM: Volatility time series and density

Figure 7.1 Stochastic volatility model with heavy-tailed distribution



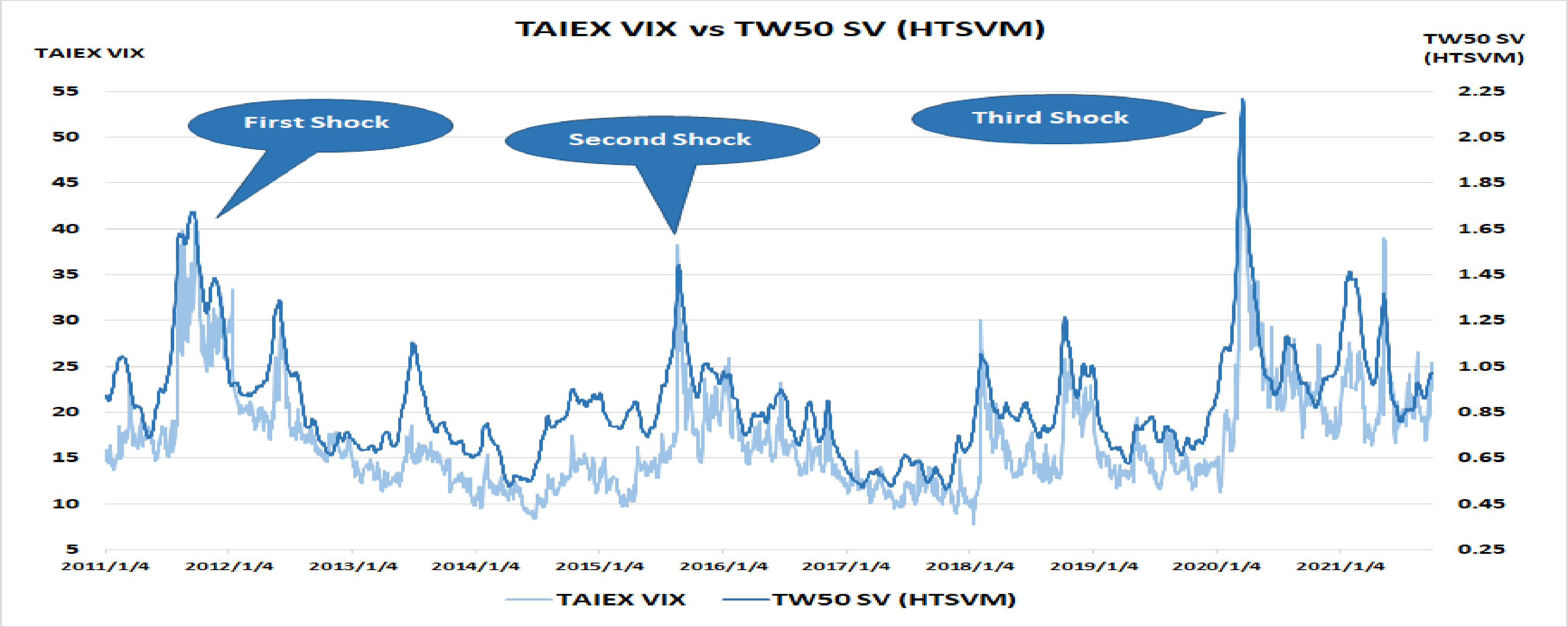
Mean	Median	High	Low	SD
0.891	0.850	2.218	0.513	0.245

Figure 7.2 Density comparison: normal (left) versus heavy tailed (right)



Result of HTSVM: major shock and rebalance signal

Figure 8.1 TAIEX VIX vs TW50 and identification of three major shocks



First shock	September, 2011
Second shock	August, 2015
Third shock	March, 2020

R0	01.03.2011 ~ 08.25.2015
R1	08.26.2015 ~ 03.20.2020
R2	03.23.2020 ~ 10.04.2021

Long run mean: 0.891
SD: 0.245

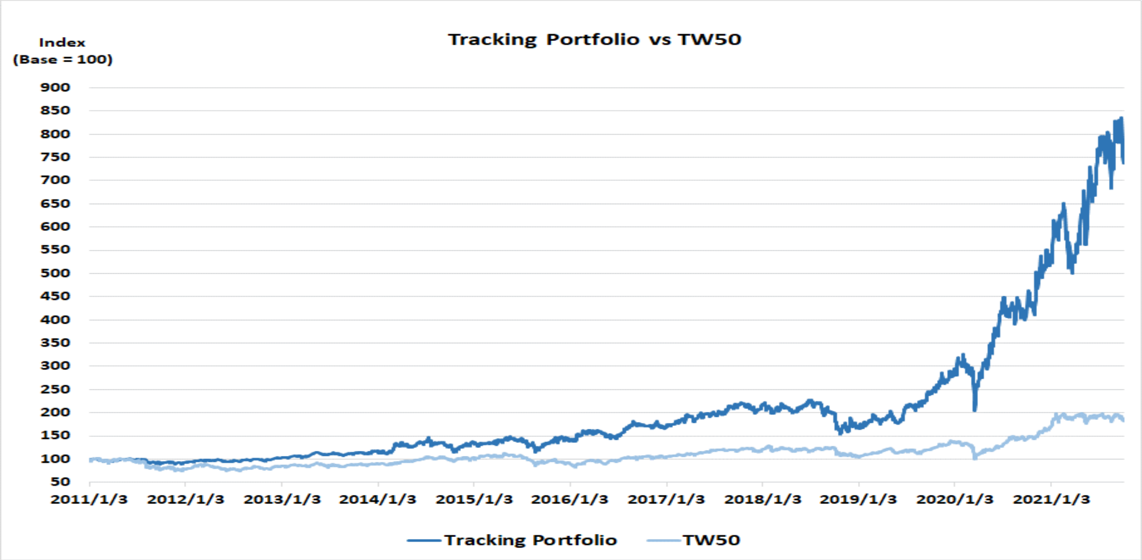
→

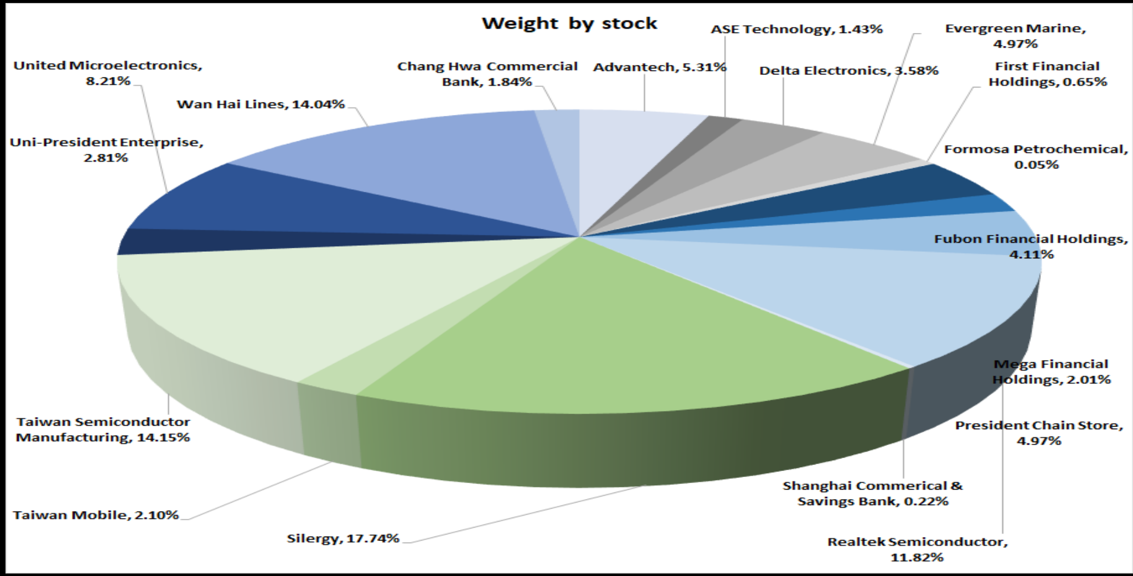
HTSVM @ 2SD above LR mean

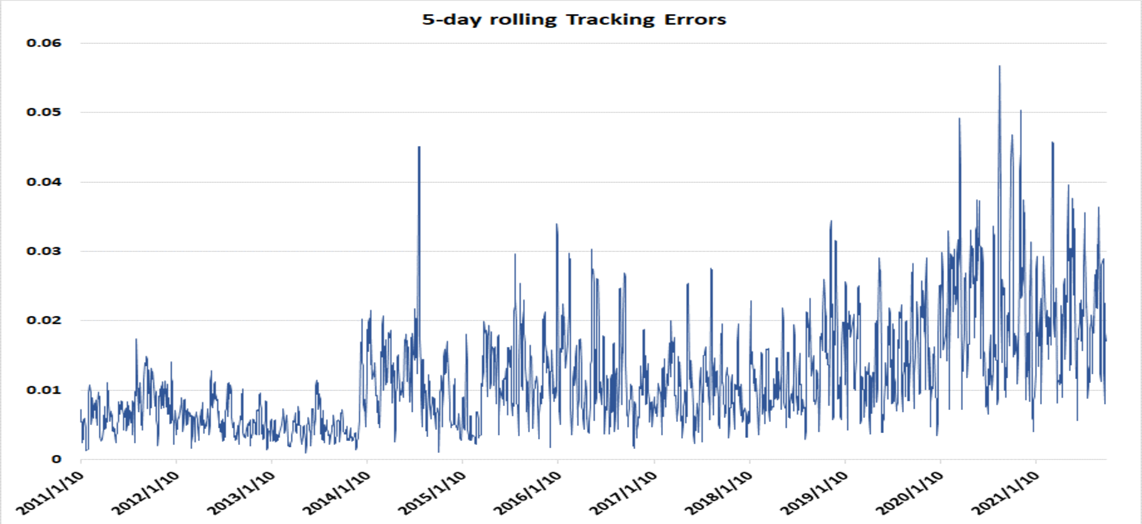
→

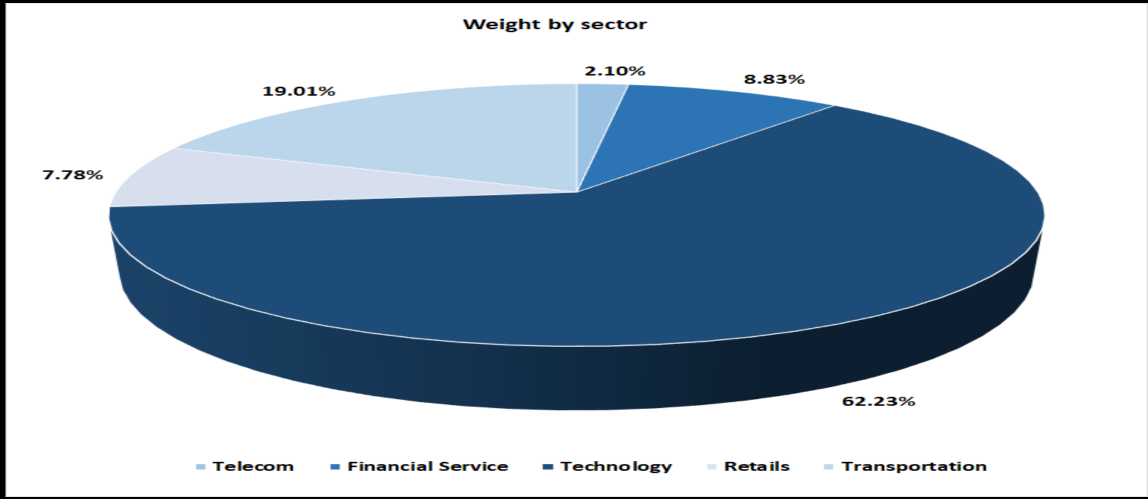
Rebalance

Result: TW50 full-sample performance and weight distribution







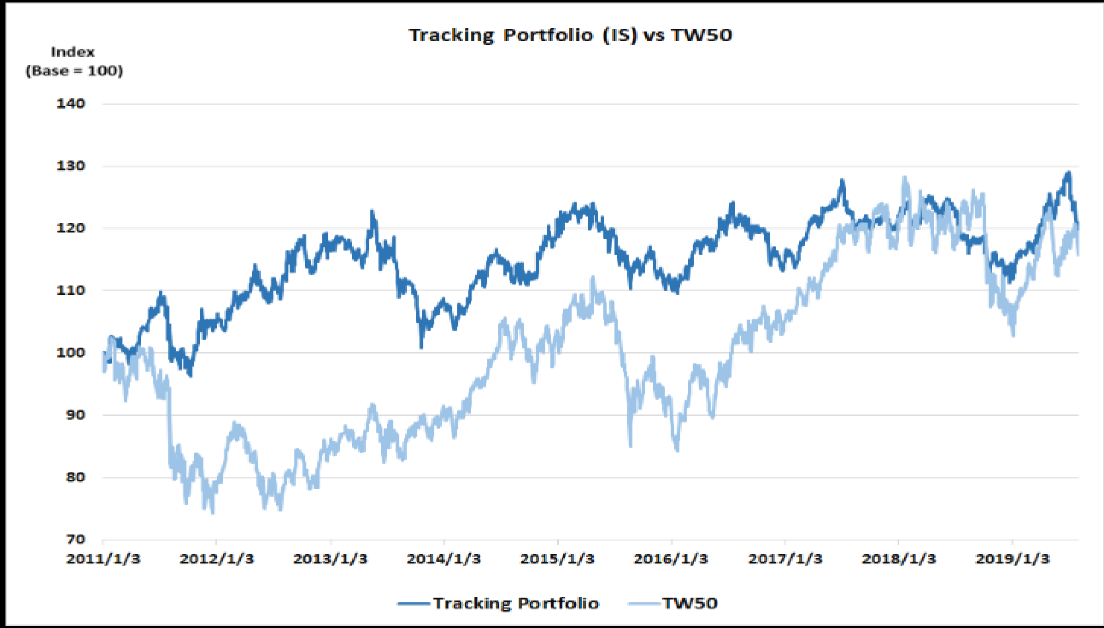


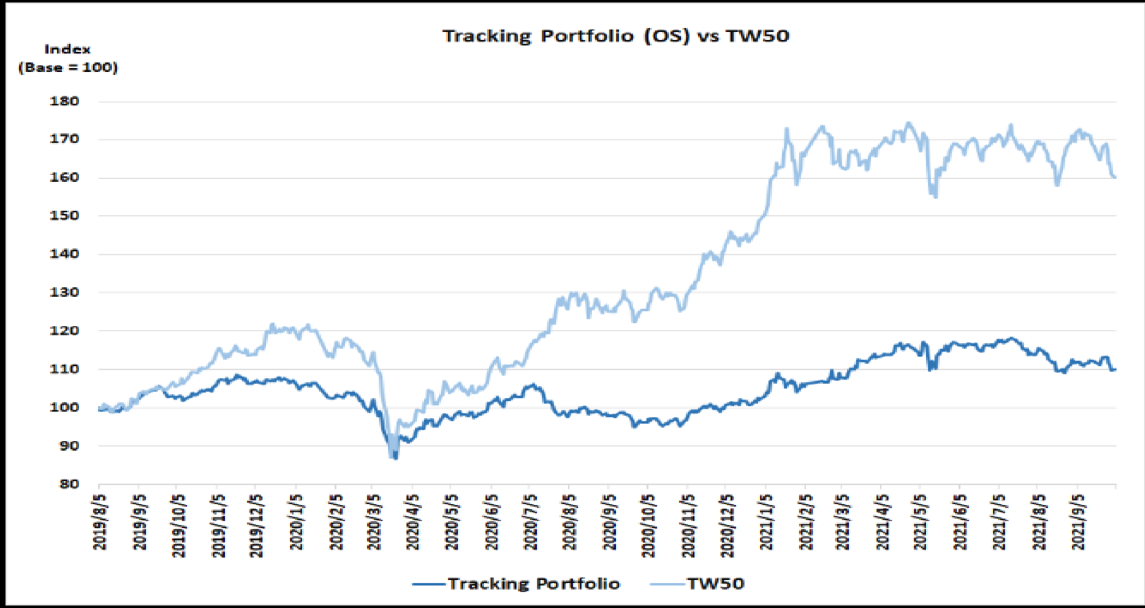
TE: $\sqrt{Var\left(Return_{trackign\ portfolio} - Return_{TW50}\right)}$

ER: $\mathrm{Log}\left(\frac{Price\ of\ trackign\ portfolio_{last\ day}}{Price\ of\ tracking\ portfolio_{first\ day}}\right) - \mathrm{Log}\left(\frac{Price\ of\ TW50_{last\ day}}{Price\ of\ TW50_{first\ day}}\right)$

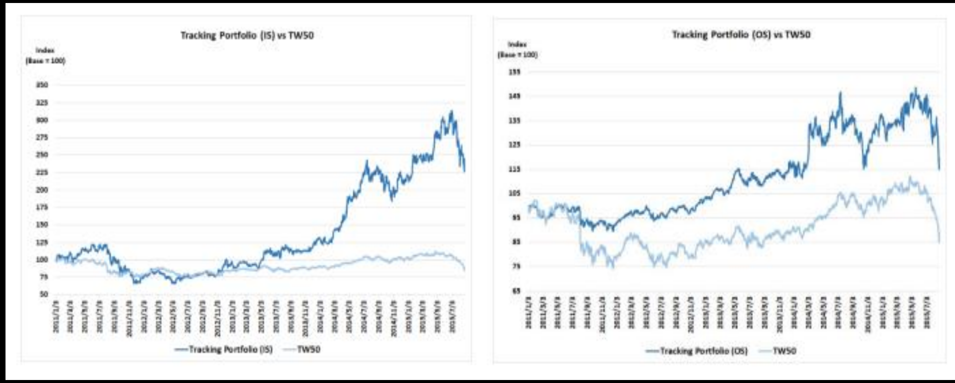
Result: TW50 In-sample (80%) vs Out-Of-Sample (20%) (without looking at volatility shock)

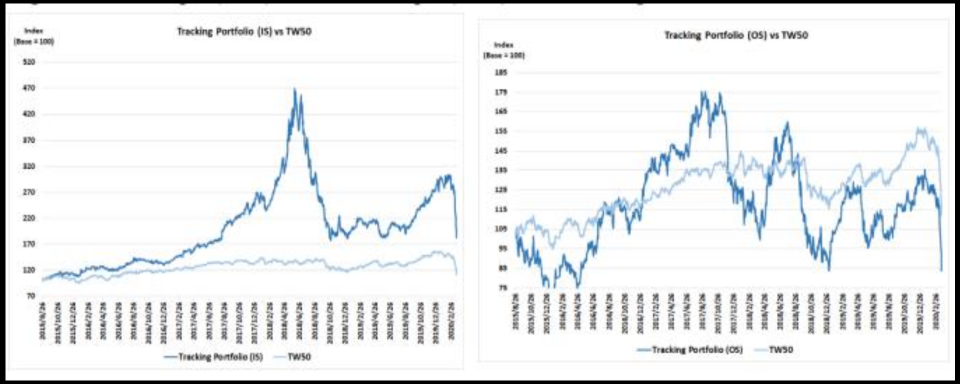
	TE	ER
In-Sample	0.020	0.857
Out-Of-Sample	0.015	−0.881





Result: TW50 sub-sample analysis for rebalancing based on volatility shock





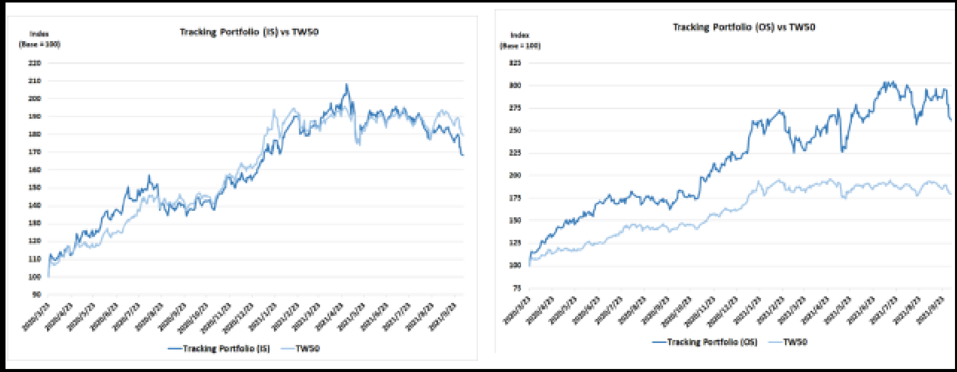


Table 8. Weighted average of TE and ER from both in-sample and out-of-sample

	In-Sample		Out-Of-Sample	
	TE	ER	TE	ER
R0	0.018	1.014	0.009	0.296
R1	0.016	0.498	0.020	−0.274
R2	0.009	−0.065	0.016	0.375
Weighted Average	**0.0159**	**0.6415**	**0.0149**	**0.0664**

Table 9. Pearson correlation

	ER (In-Sample)	TE (In-Sample)
ER (Out-Of-Sample)	0.310	
TE (Out-Of-Sample)		0.612

Concluding Remark

Summary of the finding

1. Tracking errors have been small and consistent.

2. In-sample ER is much stranger than Out-of-Sample ER

3. Active rebalancing after volatility shock can outperform passive strategy.

	In-Sample		Out-Of-Sample	
	TE	ER	TE	ER
Full-Sample (passive)	0.020	0.857	0.015	−0.881
Rebalancing by volatility shock	0.016	0.642	0.015	0.066

Potential extension

1. Multiple objectives

2. Weighted average of multiple objectives

References (for written paper)

Beasley, JE, Meade, N & Chung TJ 2003, ‘An evolutionary heuristic for the index tracking problem’, *European Journal of Operational Research*. vol. 148: 3. pp. 621-643.

Bloom, N 2009, ‘The impact of uncertainty shocks’, *Econometrica*, vol. 77: 3, pp. 623-685.

Chan, JCC & Hsiao, CYL 2013, ‘Estimation on stochastic volatility models with heavy tails and serial dependence’, *Bayesian Inferences in the Social Sciences*, Chichester, UK: Wiley.

Dose, C & Cincotti, S 2005, ‘Clustering of financial time series with application to index and enhanced index tracking portfolio’, *Physica A: Statistical Mechanics and its Applications*, vol. 18: 4, pp. 145-151.

Johansen, S 1991, ‘Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models’, *Econometrica*, vol. 59: 6, pp. 1551-1580.

Kalman, R 1960, ‘A new approach to linear filtering and prediction problems’, *ASME Journal of Basic Engineering*, vol. 92, pp. 35-45.

Li, Q, Sun, L & Bao, L 2011, ‘Enhanced index tracking based on multi-objective immune algorithm’, *Expert Systems with Applications*, vol. 38, pp. 6101-6106.

Mezali, H & Beasley, JE 2014, ‘Index tracking with fixed and variable transaction cost’, *Optimization Letters*, vol. 8, pp. 61-80.

Taylor, SJ 1986, *Modelling financial time series*, Chichester, UK: Wiley.

Watanabe, T & Asai, M 2001, ‘Stochastic volatility models with heavy-tailed distributions: A Bayesian analysis’, Available from: https://www.math.chuo-u.ac.jp/~sugiyama/15/15-02.pdf/ [Accessed 28th September, 2021].