A MODEL PREDICTIVE CONTROL APPROACH FOR PAIRS TRADING USING MULTIPLE SPREADS

September 27th, 2016

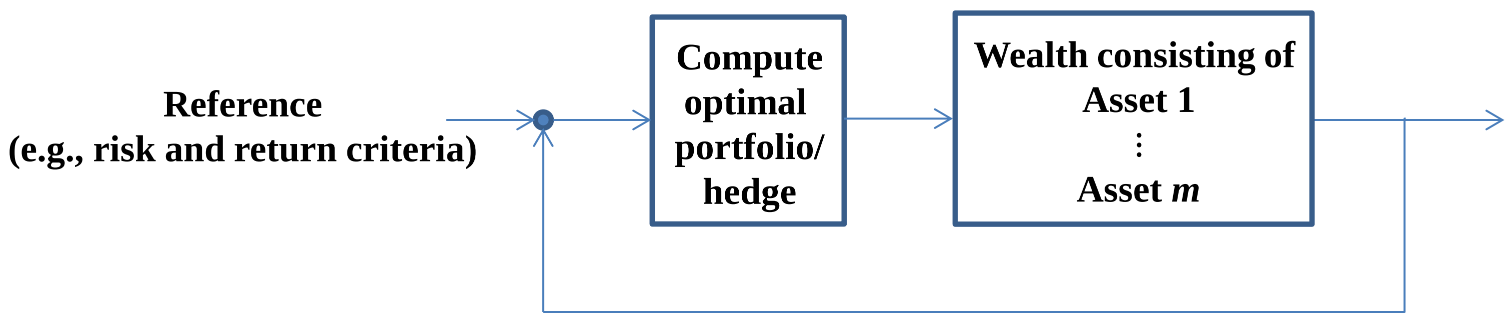
Yuji YAMADA
Faculty of Business Sciences, University of Tsukuba, JAPAN
James A. PRIMBS
Mihaylo College of Business and Economics, Cal State Fullerton, USA

Control & Finance

- In finance, we need to reduce risk of asset value fluctuation and manage/construct optimal investment portfolios of, e.g., stocks.

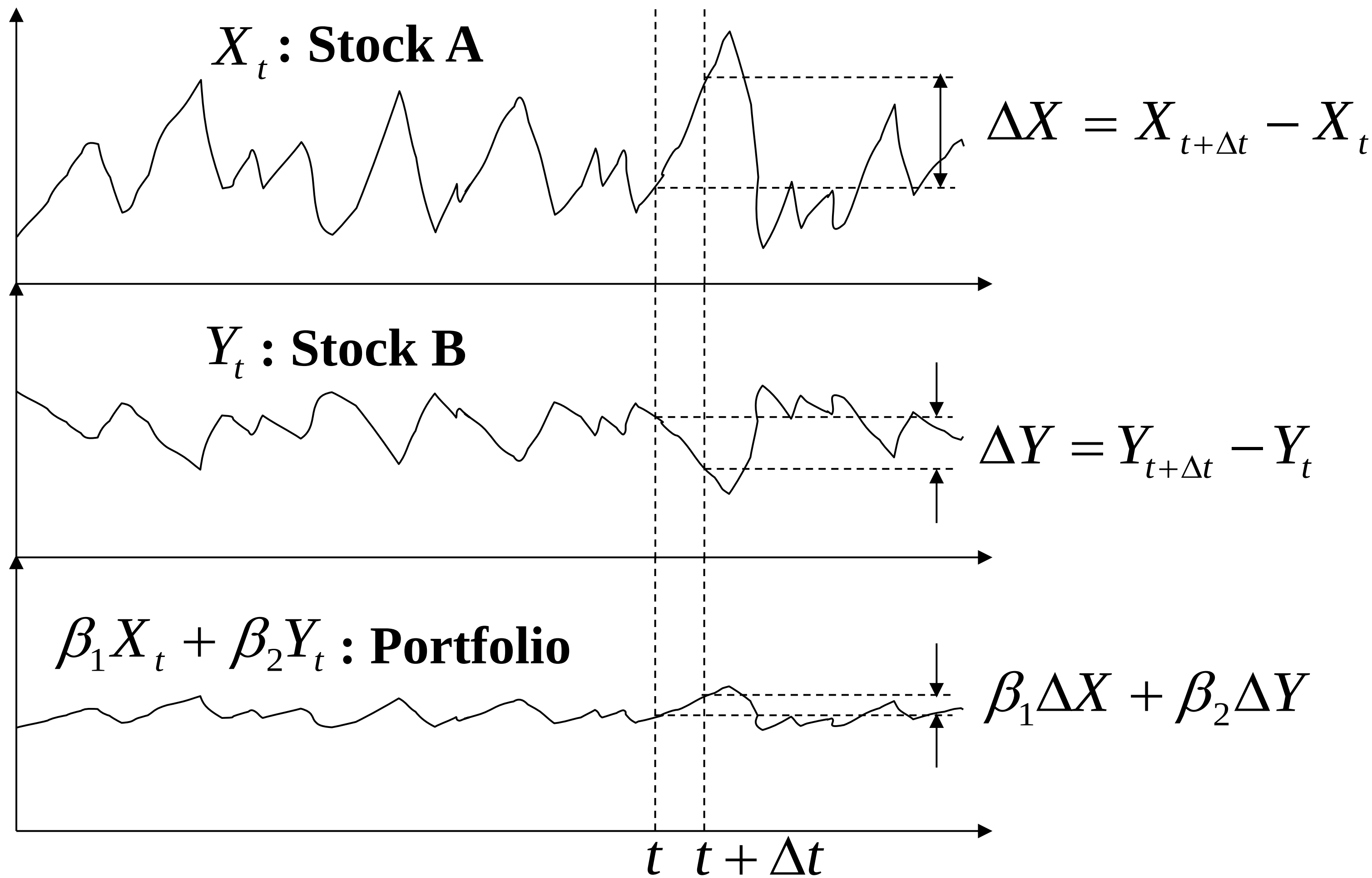
- These problems may be thought of control problems to find optimal decision variables (e.g., shares of stocks) given optimization criteria.

  - Reduce risk of asset value fluctuation ➡ Hedging problem
  - Construct optimal investment policy ➡ Portfolio optimization



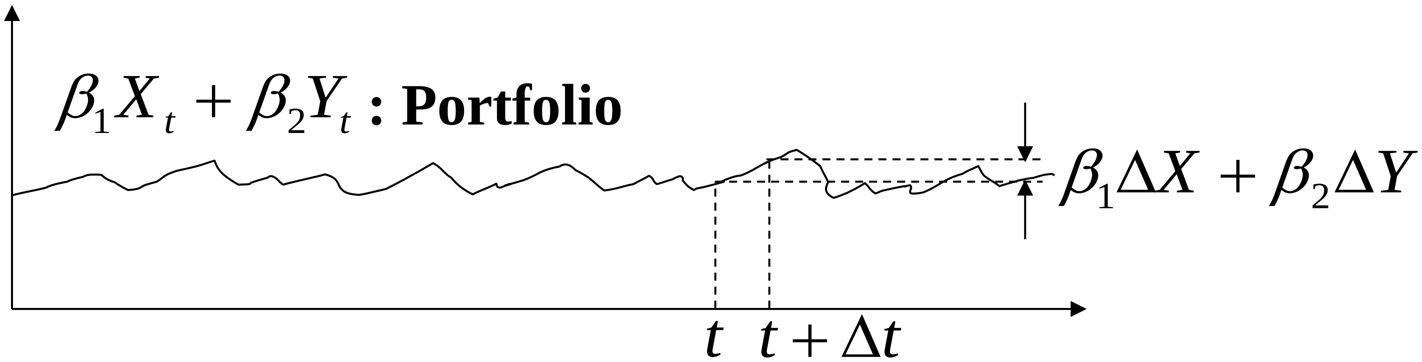
Portfolio optimization

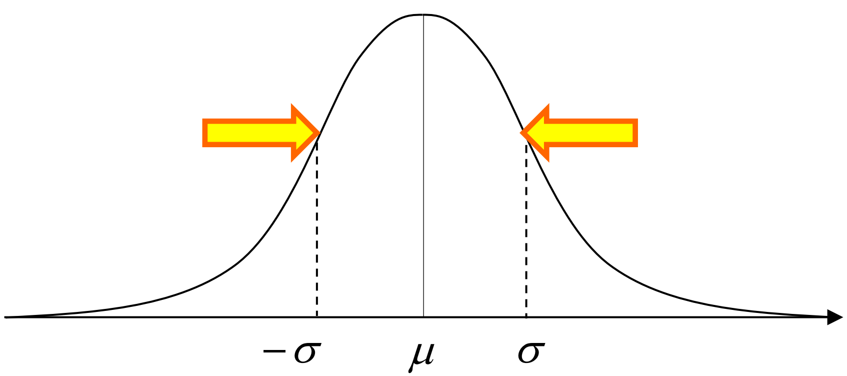
Find optimal shares of stocks, β_1, β_2, to minimize risk under admissible expected return constraint



Portfolio optimization

Find optimal shares of stocks, $\boldsymbol{\beta_1, \beta_2}$, to minimize risk under admissible expected return constraint





Distribution of $\boldsymbol{\beta_1 \Delta X + \beta_2 \Delta Y}$

$$\mu := \mathrm{E}\left[\beta_1 \Delta X + \beta_2 \Delta Y\right]$$

➡ **Fixed expected return**

$$\sigma^2 := Var\left[\beta_1 \Delta X + \beta_2 \Delta Y\right]$$

➡ **Minimize variance (risk)**

Usually difficult to control expected returns, because of random walks!

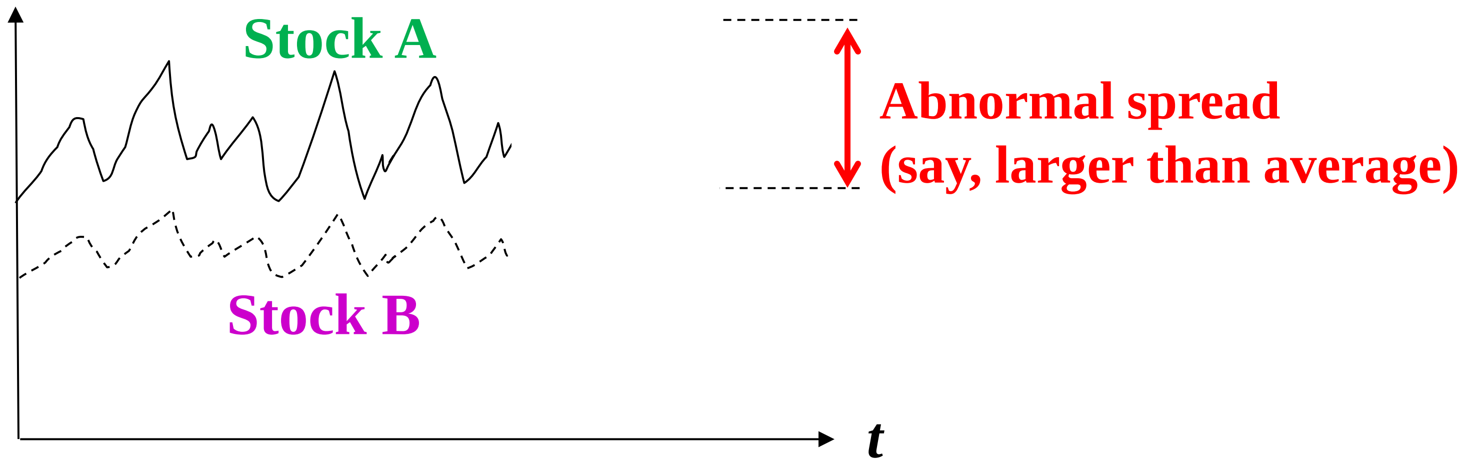
Model predictive control for spread portfolio optimization

Outline

- Basic idea of pairs trading & spread portfolio
- Optimal vs. Myopic portfolios
- MPC approach for spread portfolio optimization
- Concluding remarks

Pairs trading

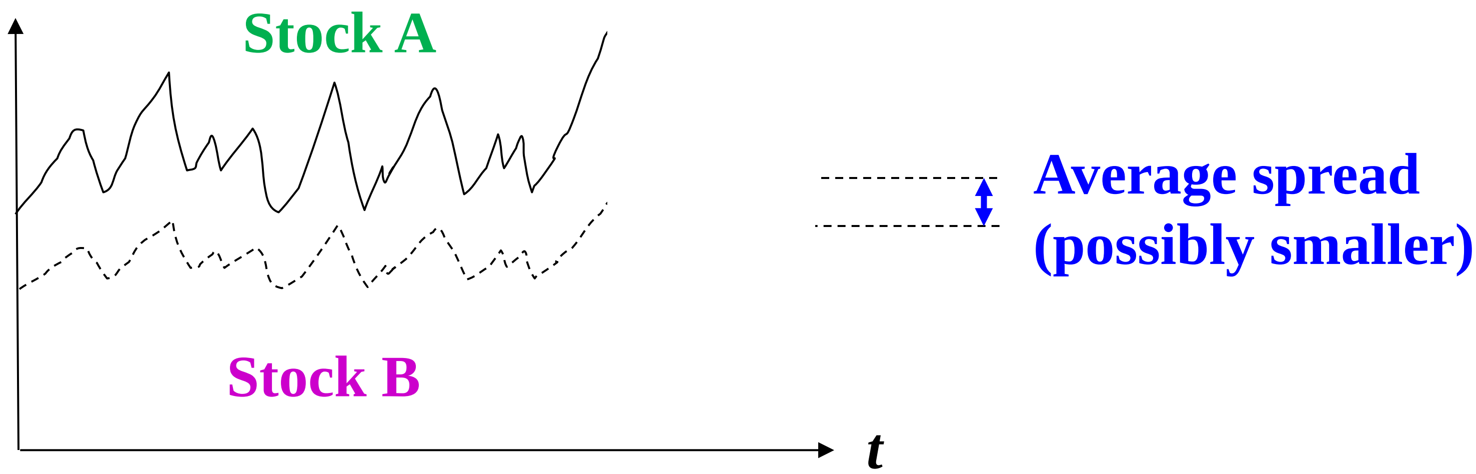
Step 1) Find a pair of stocks whose prices tend to move together (say Stock A and Stock B).



Spread = Price of Stock A – Price of Stock B

Step 2) If an abnormal spread (say, larger spread than average) is observed, buy one stock (say, Stock B) and short sell the other (say, Stock A), expecting that the spread will converge to the average in a future period.

Pairs trading



Step 3) When the spread converges to the average, sell Stock B, buy Stock A, and clear the position.

Profit of

|Abnormal spread – Average spread|

Why pairs trading?

From the practical point of view, it always leads to a positive profit if the spread converges. It does not matter whether the market is going up or down.

From the economic theory point of view, stock market prices are believed to evolve according to a random walk and thus the prices of the stock market cannot be predicted. In spite of this random walk belief, we often observe that there are pairs of stocks whose spread processes may exhibit predictable mean reversions.

"Cointegration" of pairs of stocks

What is cointegration?

Random walk: **Assume that you are walking when you are really drunk. You are not sure which way you are walking. Then the location you have at time t just depends on the location at time t-1 plus "random noise"**

$$X_t = X_{t-1} + \text{"random noise"}$$

In this case, $\{X_t\}$ has a unit root and is nonstationary $\Rightarrow$ Random walk

Cointegration: **Assume that there are two random walk processes, X_t and Y_t. If there exists a constant $\beta\,(\neq 0)$ s.t. $X_t - \beta Y_t$ is stationary, then X_t and Y_t are said to be cointegrated.**

⬇ ⬇ ⬇ ⬇

- **The spread of cointegrated stocks has predictable mean-reverting property.**
- **Construct an optimal portfolio on multiple spreads of cointegrated stocks.**

Spread process

We should find cointegrated pairs of stocks for pairs trading, because their spreads are expected to converge on average.

That is, find pairs of stocks whose prices are $X_t \sim I(1)$ **and** $Y_t \sim I(1)$, **but** $S_t := X_t - \beta Y_t$ **is stationary expressed as, e.g.,**

$$\textbf{AR(q):}\quad S_t = \phi_1 S_{t-1} + \phi_2 S_{t-2} + \cdots + \phi_q S_{t-q} + \varepsilon_t$$

In the simplest case of q = 1, it holds that

$$S_t - S_{t-1} = (\phi_1 - 1) S_{t-1} + \varepsilon_t$$

Using continuous time notation, the above equation corresponds to the following Ornstein-Uhlenbeck (OU) process:

$$dS_t = -\kappa (S_t - \theta) dt + \sigma dZ_t$$

Model predictive control for spread portfolio optimization

Outline

- Basic idea of pairs trading & spread portfolio
- Optimal vs. Myopic portfolios
- MPC approach for spread portfolio optimization
- Concluding remarks

Spread portfolio optimization problem

◆ **Basic idea:**

✓ **If a pair of stocks is cointegrated, their price difference (or spread) may be modeled as a stationary process**

✓ **The spread of each pair of stocks can be traded directory by taking a long-short position on the two stocks**

◆ **Assumption:**

✓ ***m* pairs of stock prices,** $X_t^{(i)}$ **and** $Y_t^{(i)}$, **whose spreads are modelled as**

$$S_t^{(i)} := X_t^{(i)} - \beta^{(i)} Y_t^{(i)}, \quad S_t^{(i)} = \phi_1^{(i)} S_{t-1}^{(i)} + \cdots + \phi_q^{(i)} S_{t-q}^{(i)} + \sigma^{(i)} \varepsilon_t^{(i)}, \quad i = 1, \ldots, m$$

$u_t := \left[u_t^{(1)}, \ldots, u_t^{(m)}\right]^T$ **: Shares invested in each entry of** $S_t := \left[S_t^{(1)}, \ldots, S_t^{(m)}\right]^T$

☐ **Wealth dynamics:** $\Delta W_t = u_t^T \Delta S_t + r_f \left(W_t - u_t^T S_t\right)$ **Risk free interest rate**

$$= r_f W_t + W_t v_t^T \left(\Delta S_t - r S_t\right), \quad v_t := u_t / W_t$$

Method 1: Myopic portfolio

- **Conditional mean & variance on the wealth return**

$$\mathrm{E}_t\left[\frac{\Delta W_t}{W_t}\right] = r_f + v_t^T\left(\mathrm{E}_t[\Delta S_t] - rS_t\right), \quad \mathrm{V}_t\left[\frac{\Delta W_t}{W_t}\right] = v_t^T \Sigma v_t$$

$\Sigma \in \Re^{m\times m}$: **Covariance matrix of** $\sigma^{(i)}\varepsilon_t^{(i)},\ i = 1,\dots,m$

- **Conditional mean-variance optimization on the wealth return:**

$$\max_{v_t}\left\{\mathrm{E}_t\left[\frac{\Delta W_t}{W_t}\right] - \frac{\gamma}{2}\cdot \mathrm{V}_t\left[\frac{\Delta W_t}{W_t}\right]\right\}$$

γ : **Risk aversion coefficient**

- **Mean-variance optimal portfolio: (Myopic portfolio)**

$$v_t^* = \frac{1}{\gamma}\Sigma^{-1}\left(\mathrm{E}_t[\Delta S_t] - rS_t\right), \quad u_t^* = v_t^* W_t$$

Method 2: Long term optimal portfolio

- **Long term optimal portfolio for spreads [Kim, Primbs and Boyd 2008]:**

$$\max_{v_t} \frac{1}{1-\gamma} E\left[W_T^{1-\gamma}\right]$$

$$\text{s.t. } dS_t = -K(S_t - \theta)dt + BdZ_t$$

$$\frac{dW_t}{W_t} = rdt + v_t^T\left((-K(X_t - \theta) - rS_t)dt + BdZ_t\right)$$

$$S_t = \begin{bmatrix} S_t^{(1)} \\ \vdots \\ S_t^{(N)} \end{bmatrix}, \ \theta = \begin{bmatrix} \theta^{(1)} \\ \vdots \\ \theta^{(N)} \end{bmatrix}, \ K = \begin{bmatrix} \kappa^{(1)} & & 0 \\ & \ddots & \\ 0 & & \kappa^{(N)} \end{bmatrix}, \ B = \begin{bmatrix} \sigma^{(1)} & & 0 \\ & \ddots & \\ 0 & & \sigma^{(N)} \end{bmatrix}$$

- The solution depends on a Riccati equation and linear differential equations [Herzog et al. ‘04, Campbell et al. ’03, Campbell and Viceira ’02] and is dynamic.

Search for Cointegrated Pairs

- **Apply the following procedures to search for cointegrated pairs**

Screening procedure) Compute the DF statistics and the correlation coefficient of each pair. If the absolute value of correlation coefficient is below 0.8 or the DF statistics is above certain CV, then remove the pair.

Selection procedure) Sort the pairs following the order of smaller DF statistics. Apply the Engle-Granger cointegration test [Engle and Granger 1987] from the top of the list to select cointegrated pairs without overlap of a company.

Search for Cointegrated Pairs

DF statistics of the spread: $S_t := X_t - \beta_{X|Y} Y_t$

$$\Delta S_{t-1} = b S_{t-1} + \varepsilon_t, \ \ \Delta S_{t-1} = S_t - S_{t-1}$$

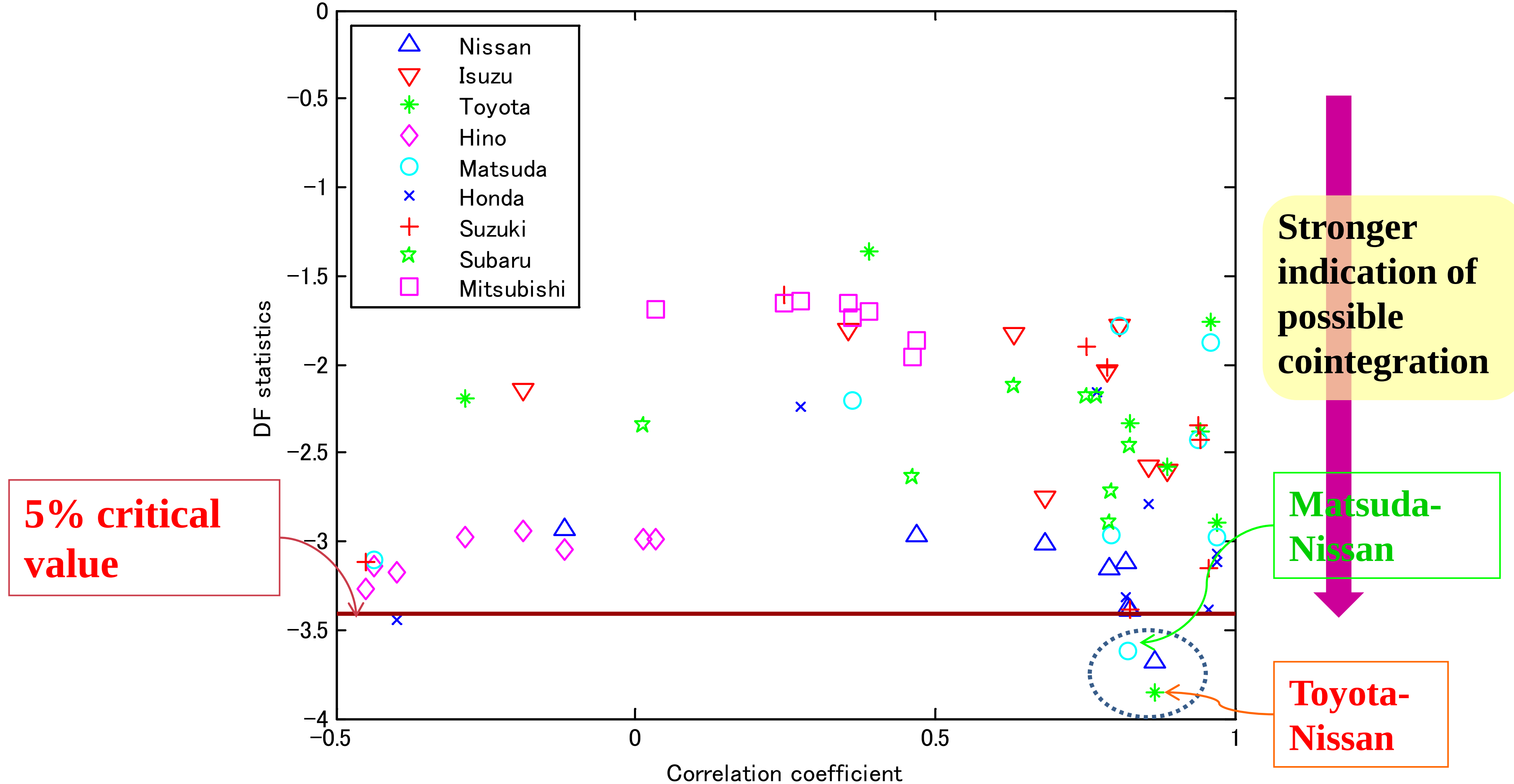
$$\textbf{DF statistics} := \frac{\hat{b}}{Standard\ Error}, \quad \hat{b}\textbf{: OLS estimator}$$

If the DF statistics is less than a critical value with a given significance level, the null hypothesis of "unit root" is rejected.

The smaller DF statistics is, the stronger the rejection of the hypothesis.

Automotive industries

Correlation coefficients vs. DF statistics: Nissan, Isuzu, Toyota, Hino, Matsuda, Honda, Suzuki, Subaru, Mitsubishi (2004-2006)



Empirical simulation

1st Analysis: Find cointegrated pairs of stocks in Nikkei 225 using data period of 2004-2006 (3 years), construct the optimal and myopic portfolios, and simulate the out-of-sample performance of 2007 (or later period)

Possible combinations of pairs are 225×(225-1)/2 = 25,200, but in fact we need to check the double of this number by switching dependent and independent variables in the linear regression.

➡ **# of pairs is 50,400**

$$X_t = \beta_{X|Y} Y_t + const + residuals$$

$$Y_t = \beta_{Y|X} X_t + const + residuals$$

Nikkei 225 has 35 categories of industries, so choose some and find cointegrated pairs in each category.

➡ **10 categories consisting of 84 industries**

Search for Cointegrated Pairs

By applying the screening and selection procedure, the following 13 pairs were selected using the data in the period of 2004-2006:

Japan Tobacco vs. Asahi Breweries (Foods)

Showa Shell Sekiyu vs. Nippon Oil (Oil & coal products)

Tokyo Electron vs. Sharp (Electric machinery)

TDK vs. Fuji Electric (Electric machinery)

Panasonic vs. Meidensha (Electric machinery)

Panasonic Electric Works vs. Sony (Electric machinery)

Minebea vs. Kyocera (Electric machinery)

Toyota vs. Nissan (Automotive)

Olympus vs. Nikon (Precision instruments)

Tokyu vs. Tobu (Railway/Bus)

KDDI vs. NTT (Communications)

Chubu vs. Kansai (Electric power)

CSK vs. Dentsu (Services)

Out-of-Sample Simulations

Objective: **Apply the optimal and myopic portfolios for pairs trading to evaluate the out-of-sample performance**

- ✓ **Number of pairs?**
- ✓ **Data periods?**

Simulation 1)

- ➢ **Data period: Fixed**
  - **- Parameter estimation period; 2004—2006**
  - **- Simulation peirod (out-of-sample); 2007**
- ➢ **Number of pairs: Vary from 3 to 13**

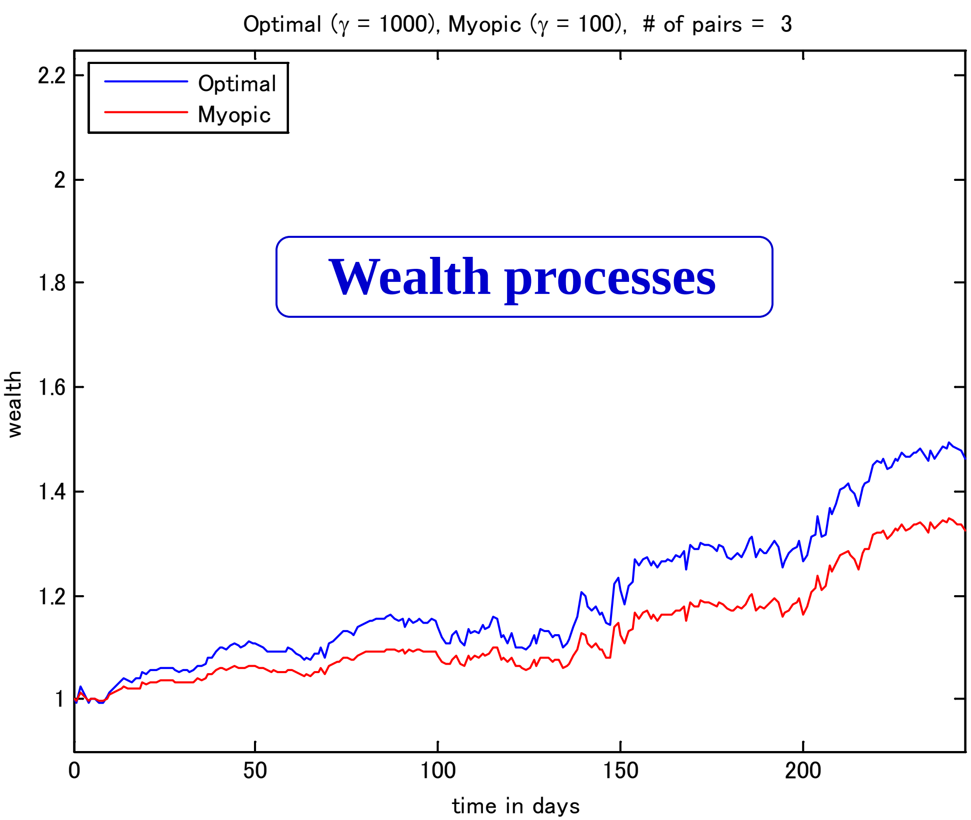
Simulation 2)

- ➢ **Number of pairs: Fixed (=13)**
- ➢ **Simulation period: Vary from 2007 to 2009**

Simulation 1) # of pairs?

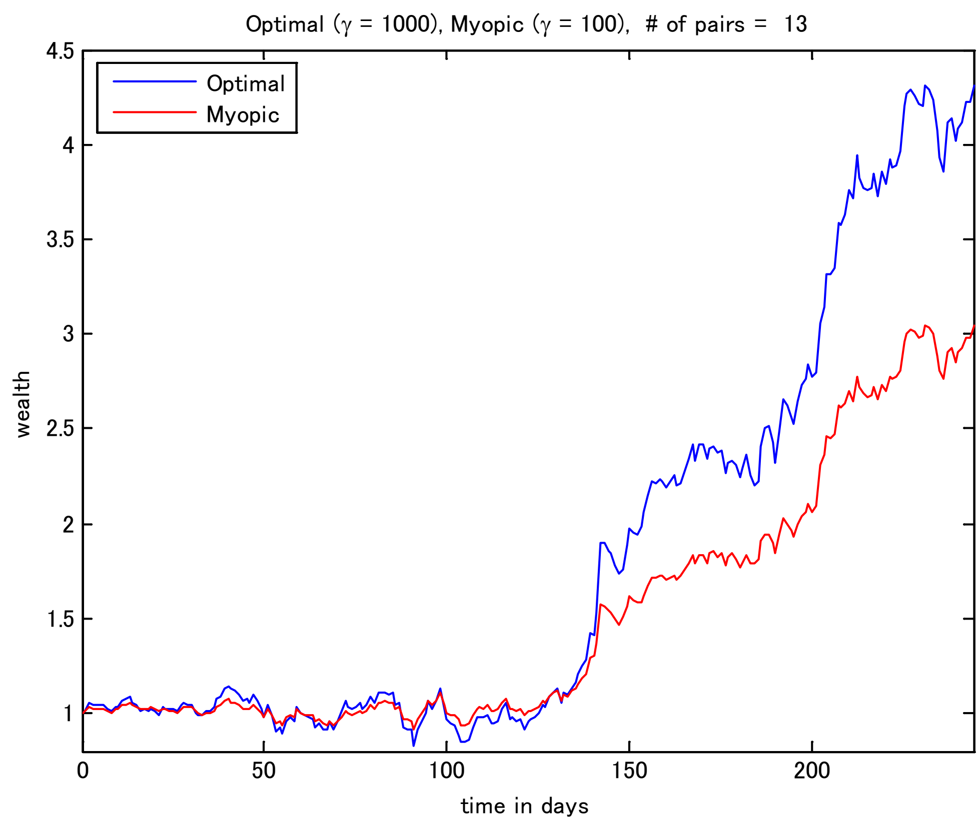
- **Parameter estimation period; 2004—2006**
- **Simulation peirod (out-of-sample); 2007**

Number of pairs: 3



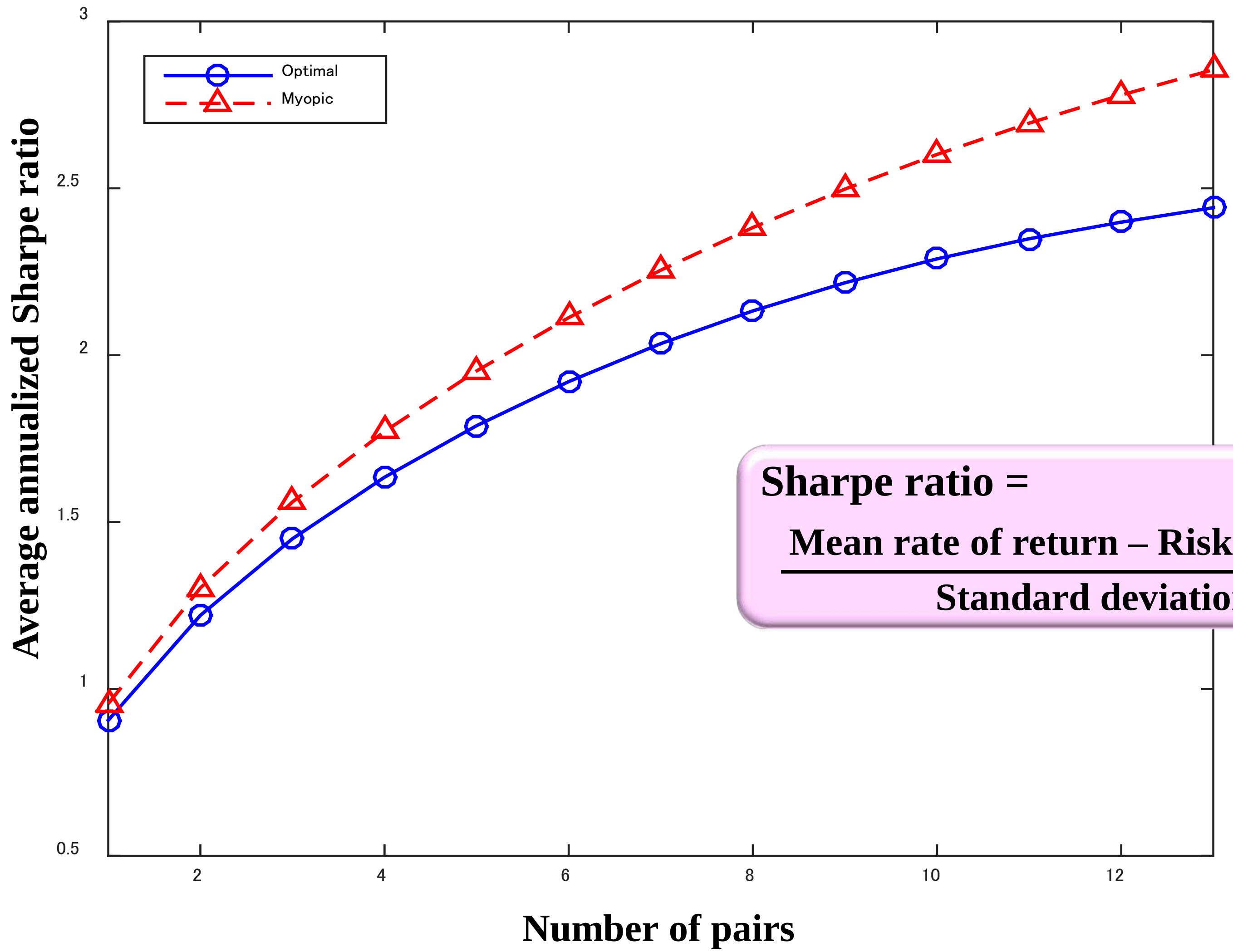
Wealth processes

Number of pairs: 13



Simulation 1) # of pairs?

- Number of pairs vs. Average annualized Sharpe ratio



Sharpe ratio =

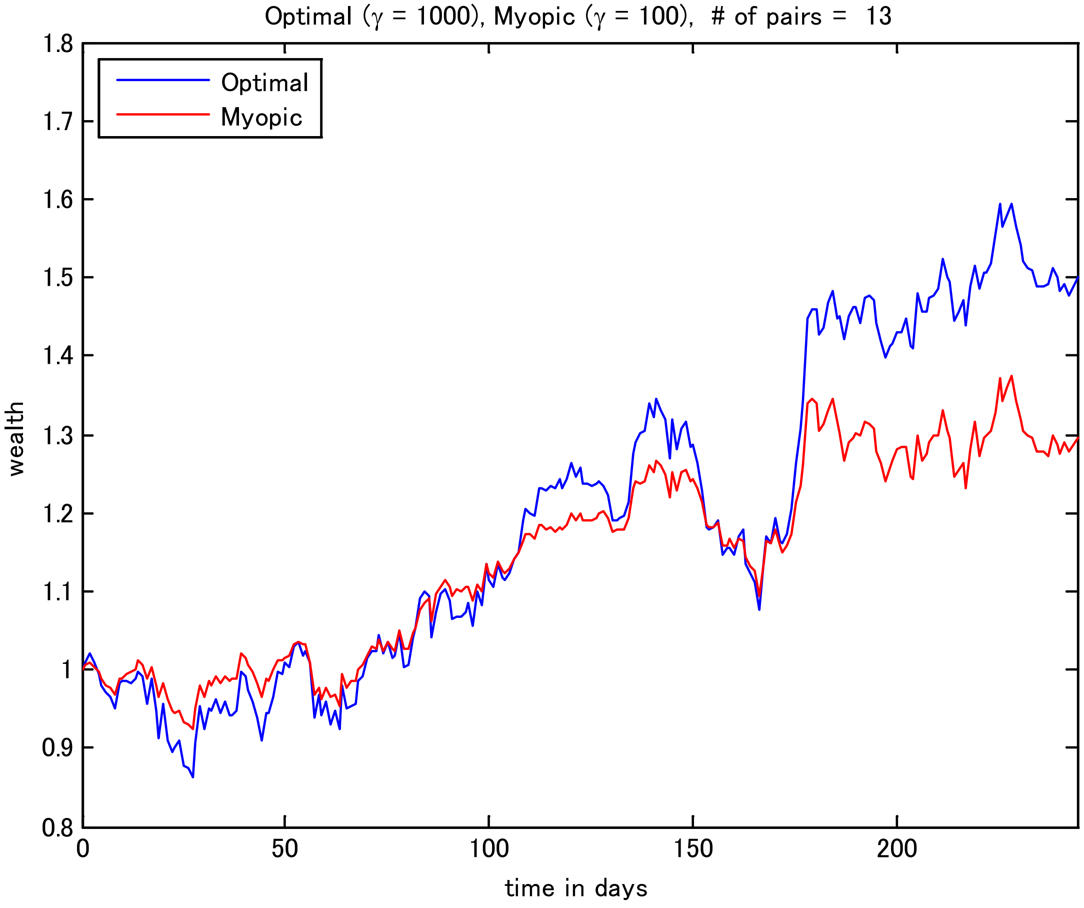
$$\frac{\text{Mean rate of return} - \text{Risk free rate}}{\text{Standard deviation}}$$

Simulation 2) Extend simulation period

- **Number of pairs: Fixed (=13)**

  - Parameter estimation period; 2005—2007

  - Simulation peirod (out-of-sample); 2008

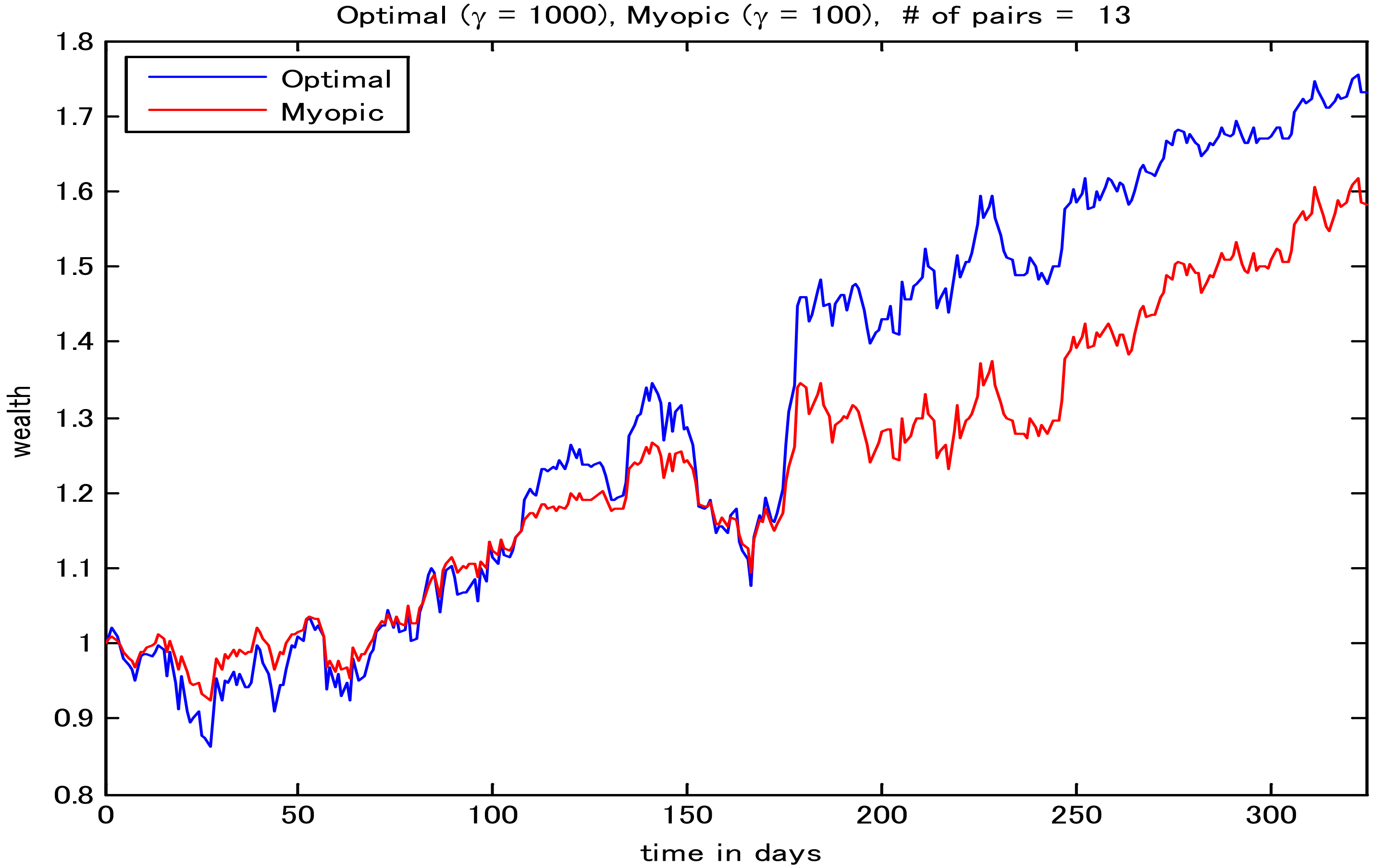


Simulation 2) Extend simulation period

- **Number of pairs: Fixed (=13)**

  - Parameter estimation period; 2005—2007

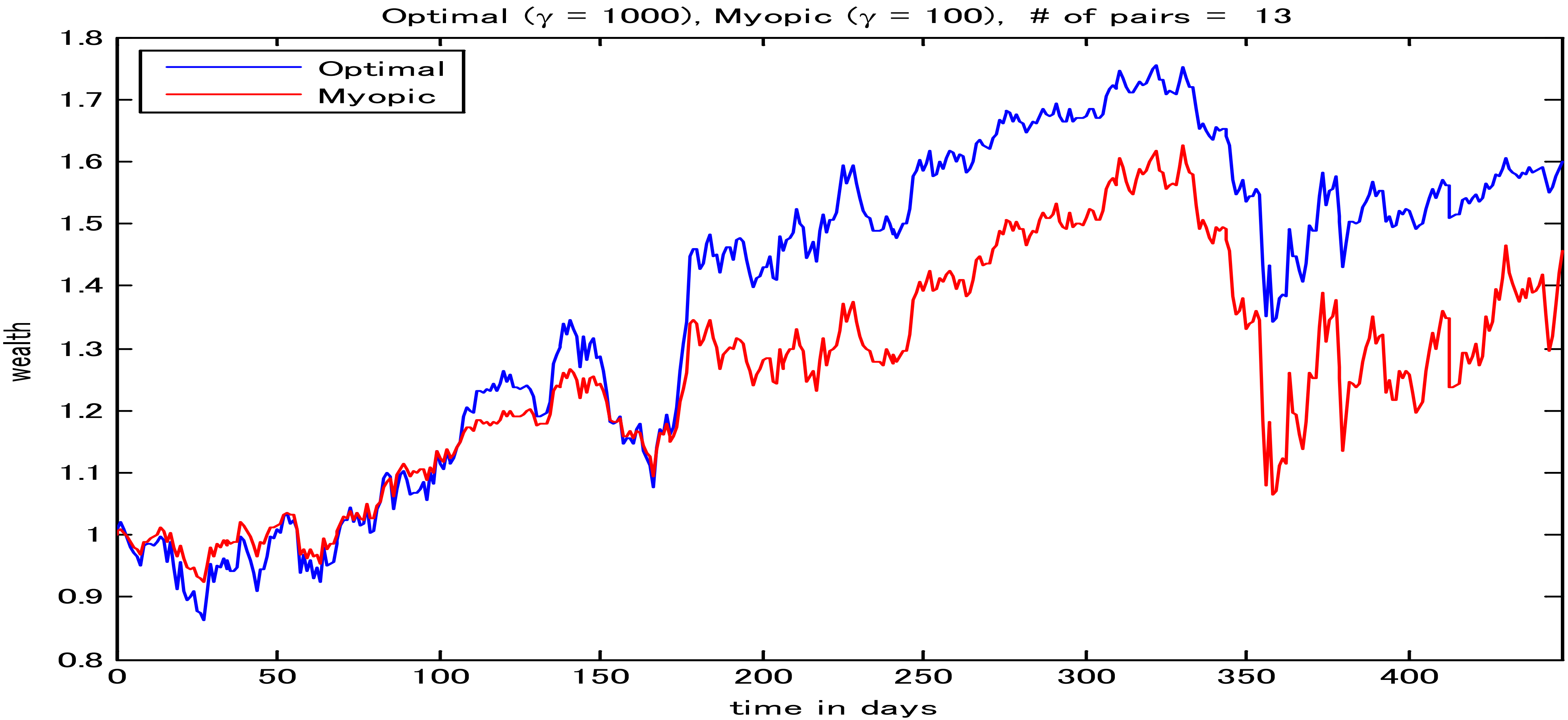
  - Simulation peirod (out-of-sample); 2008-2009.4



Simulation 2) Extend simulation period

- Number of pairs: Fixed (=13)

  - Parameter estimation period; 2005—2007
  - Simulation peirod (out-of-sample); 2008-2009.10



Summary so far

- Myopic vs. Long term optimal portfolios
  - Long term optimal portfolio provides a better wealth level.
  - In terms of Sharpe ratio, myopic portfolio seems to be better.

- Long term optimal portfolio is not as good as expected…
  - May be inflexible, due to fixed horizon and control policy.

- Extend myopic portfolio to take longer horizon & flexible control period into account.

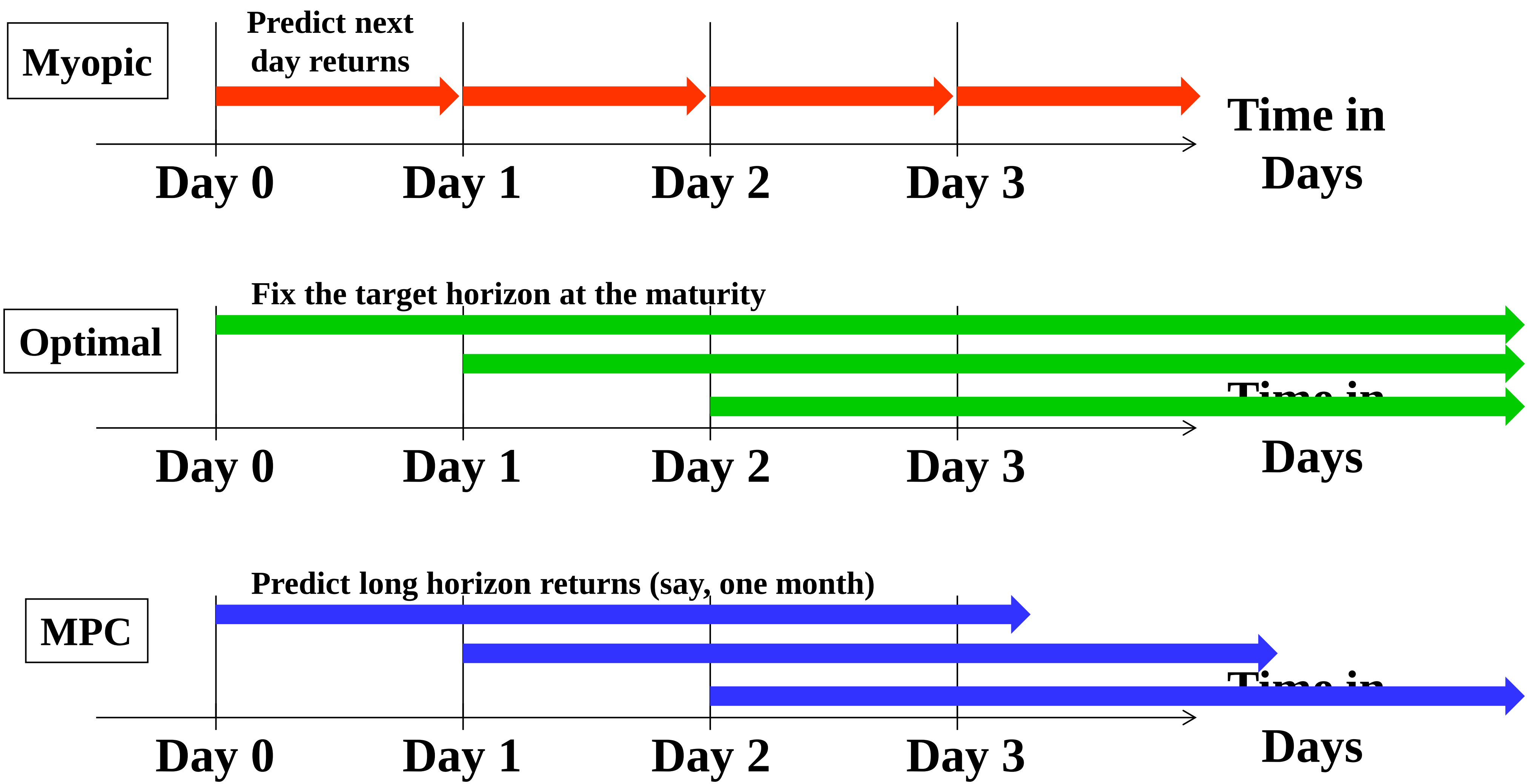
Model predictive control (MPC) for spread trading

Model predictive control for spread portfolio optimization

Outline

- Basic idea of pairs trading & spread portfolio
- Optimal vs. Myopic portfolios
- MPC approach for spread portfolio optimization
- Concluding remarks

Myopic vs. MPC vs. Long term optimal



Model predictive control for spread trading

- **Basic idea**
  - **Express multivariate spread process based on VARMA model.**
  - **Solve conditional MV problem for arbitrarily long prediction horizon using the prediction of future spread.**
- **Once the conditional MV problem is solved, we can calculate an optimal portfolio that gives a static feedback control law.**
- **Update the control law dynamically with the current state variables for the spreads at each rebalance period.**

Construction of MPC

$$VAR(q): S_t = \Phi_1 S_{t-1} + \cdots + \Phi_q S_{t-q} + c + e_t,\ \Phi_i \in \Re^{m\times m},\ c \in \Re^m$$

$$S_t := \left[S_t^{(1)}, \ldots, S_t^{(m)}\right]^T, \quad S_t^{(i)} := X_t^{(i)} - \beta^{(i)} Y_t^{(i)}, \quad i = 1, \ldots, m$$

$$e_t \sim N(0, \Sigma),\ \Sigma \in \Re^{m\times m} : \text{covariance matrix}$$

- **Wealth in the future time horizon:**

$$W_{t+\tau} = u_t^T S_{t+\tau} + \left(1 + r_f\right)^\tau \left(W_t - u_t^T S_t\right),\ u_t := \left[u_t^{(1)}, \ldots, u_t^{(m)}\right]^T \in \Re^m$$

- **Total return of the wealth:**

$$R_{t,\tau} := \frac{W_{t+\tau}}{W_t} = (1+r)^\tau + v_t^T \left[S_{t+\tau} - (1+r)^\tau S_t\right], \quad v_t := \frac{u_t}{W_t}$$

- **Conditional mean-variance optimization:**

$$\max_{u_t \in \Re^m} \left\{ \mathrm{E}_t\left[R_{t,\tau}\right] - \frac{\gamma}{2} \cdot \mathrm{V}_t\left[R_{t,\tau}\right] \right\},\ \mathrm{V}_t\left[R_{t,\tau}\right] = \mathrm{V}_t\left[v_t^T S_{t+\tau}\right] : \text{conditional variance}$$

Construction of MPC

$$VAR(1): S_t = \Phi_1 S_{t-1} + c + e_t,\ \ \Phi_1 \in \Re^{16\times 16}$$

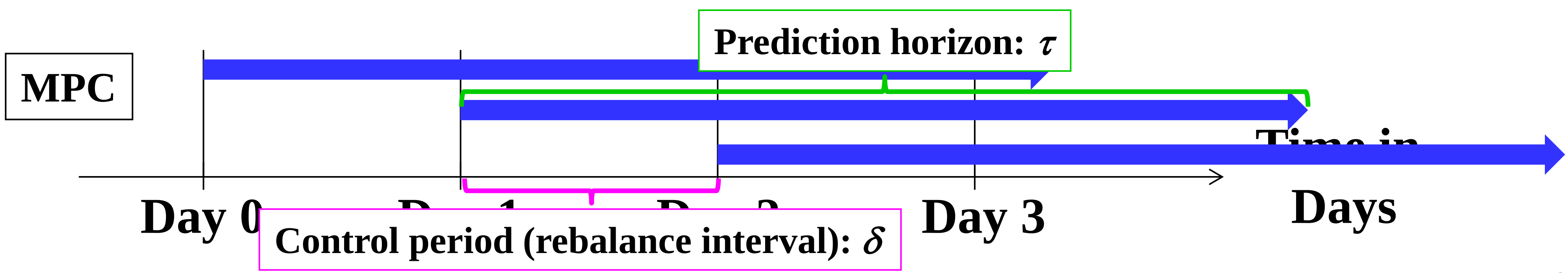
$$S_{t+\tau} = \Phi_1 S_{t+\tau-1} + c + e_{t+\tau} = \Phi_1^2 S_{t+\tau-2} + (\Phi_1 + 1)c + (\Phi_1 e_{t+\tau-1} + e_{t+\tau}) = \cdots \ \cdots \ \cdots$$

$$= \Phi_1^{\tau} S_t + \left(\Phi_1^{\tau-1} + \cdots + \Phi_1 + I\right)c + \left(\Phi_1^{\tau-1} e_{t+1} + \cdots + \Phi_1 e_{t+\tau-1} + e_{t+\tau}\right)$$

- **Predicted value:** $\mathrm{E}_t(S_{t+\tau}) = \Phi_1^{\tau} S_t + \left(\Phi_1^{\tau-1} + \cdots + \Phi_1 + I\right)c$

- **Optimal control input (share unit vector on the spread):**

$$u_t^* = \frac{W_t}{\gamma}\left(\Sigma + \Phi_1 \Sigma \Phi_1^T + \cdots \Phi_1^{\tau-1}\Sigma\left(\Phi_1^{\tau-1}\right)^T\right)^{-1} \times \left[\mathrm{E}_t(S_{t+\tau}) - (1 + r_\tau)S_t\right] \in \Re^{16}$$



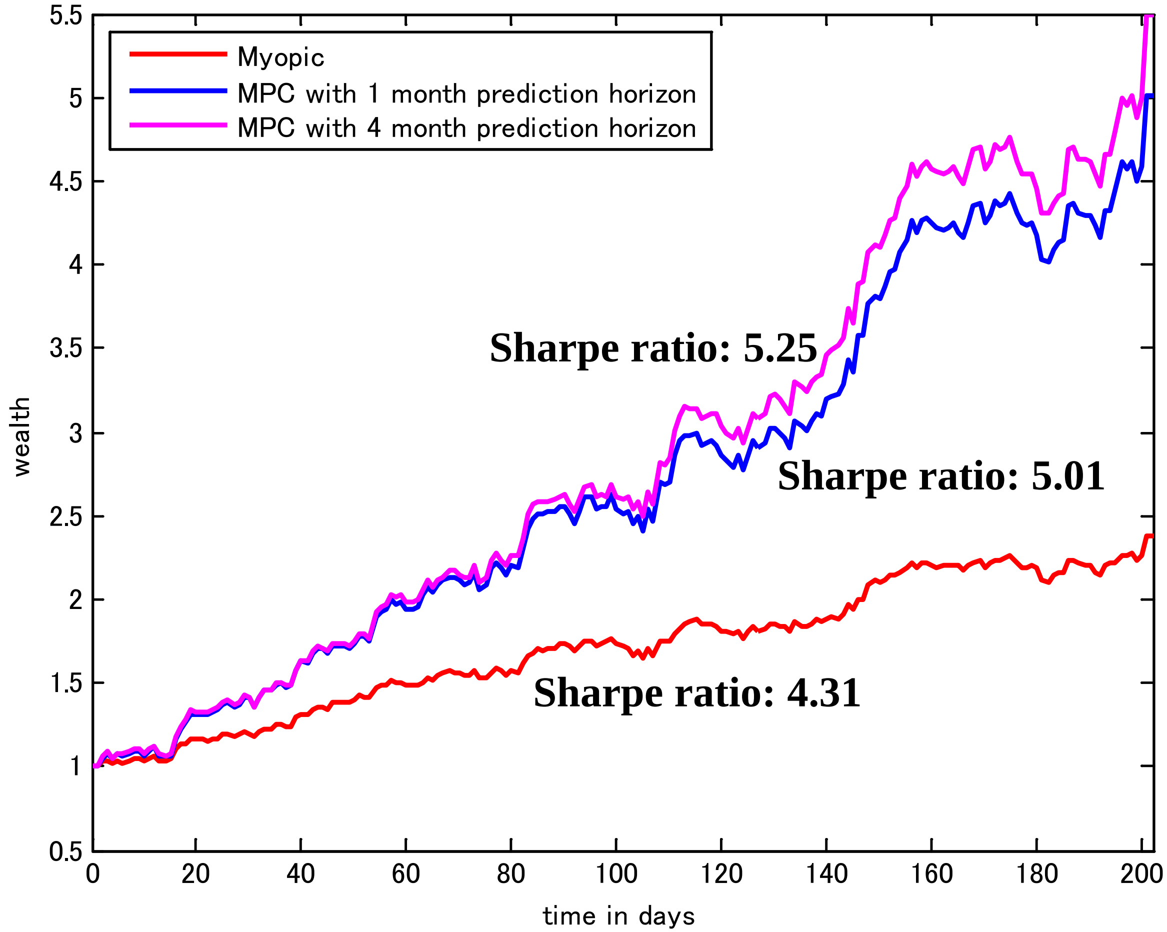
New simulations

- List of 16 pairs based on the data period of 2007—2009:

Sapporo vs. Asahi Breweries (Foods)
Ajinomoto vs. Kikkoman (Foods)
Nippon Oil vs. Nippon Mining Holdings (Oil & coal products)
Advantest vs. Taiyo Yuden (Electric machinery)
TDK vs. Tokyo Electron (Electric machinery)
Canon vs. Kyocera (Electric machinery)
Minebea vs. Denso (Electric machinery)
Fanuc vs. Toshiba (Electric machinery)
Fuji Electric vs. Sony (Electric machinery)
Honda vs. Nissan (Automotive)
Matsuda vs. Fuji Heavy Industries (Automotive)
Konica Minolta Holdings vs. Nikon (Precision instruments)
West Japan Railway vs. East Japan Railway (Railway/Bus)
NTT Data vs. NTT (Communications)
Tokyo Gas vs. Chubu Electric Power (Gas/Electric power)
Yahoo vs. Konami (Services)

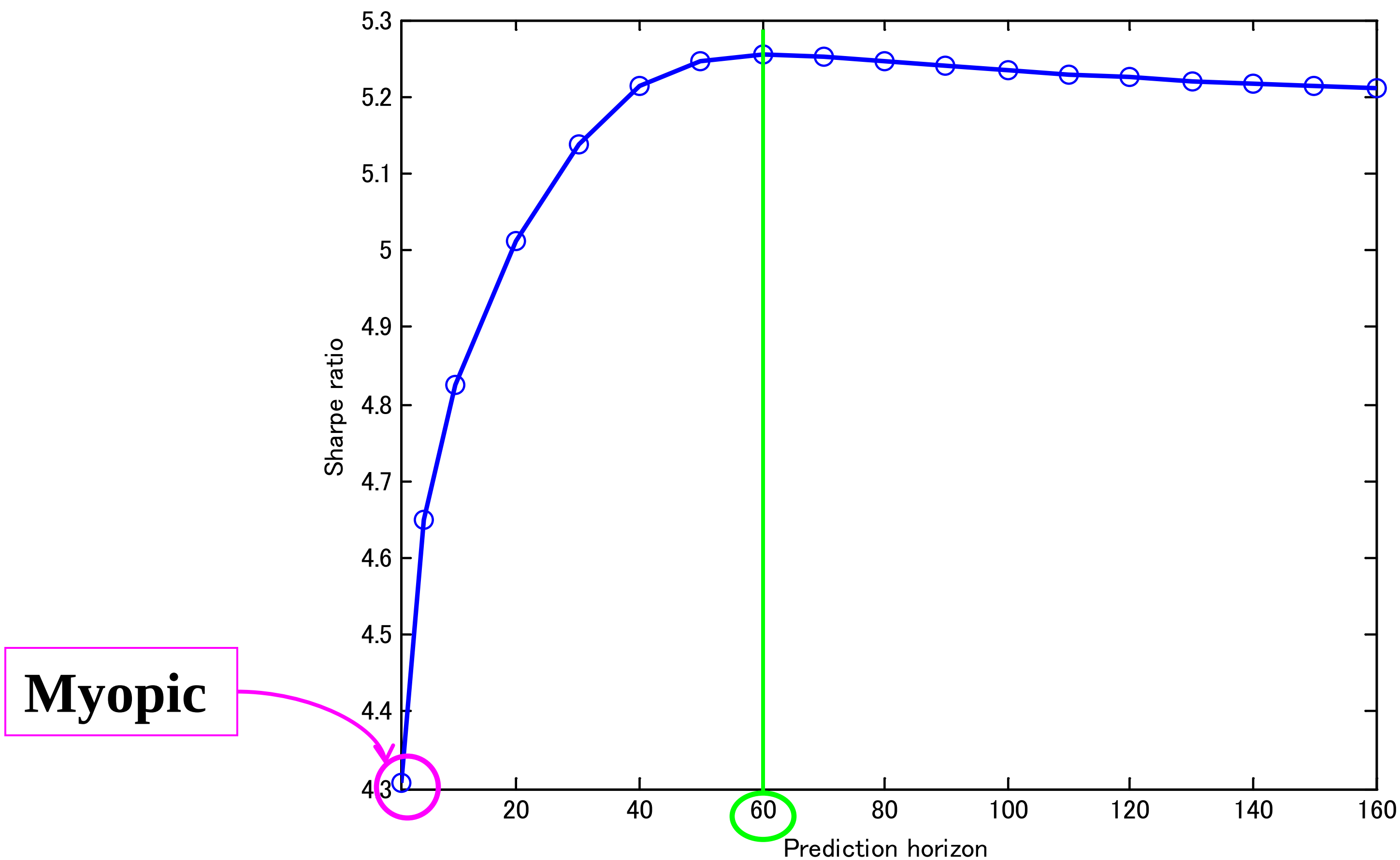
Myopic vs. MPC with different prediction horizons

- **Myopic; τ = 1 day (prediction horizon) & δ = 1 day (rebalance interval)**
- **MPC; τ = 20 days and τ = 80 days with δ = 1 day**



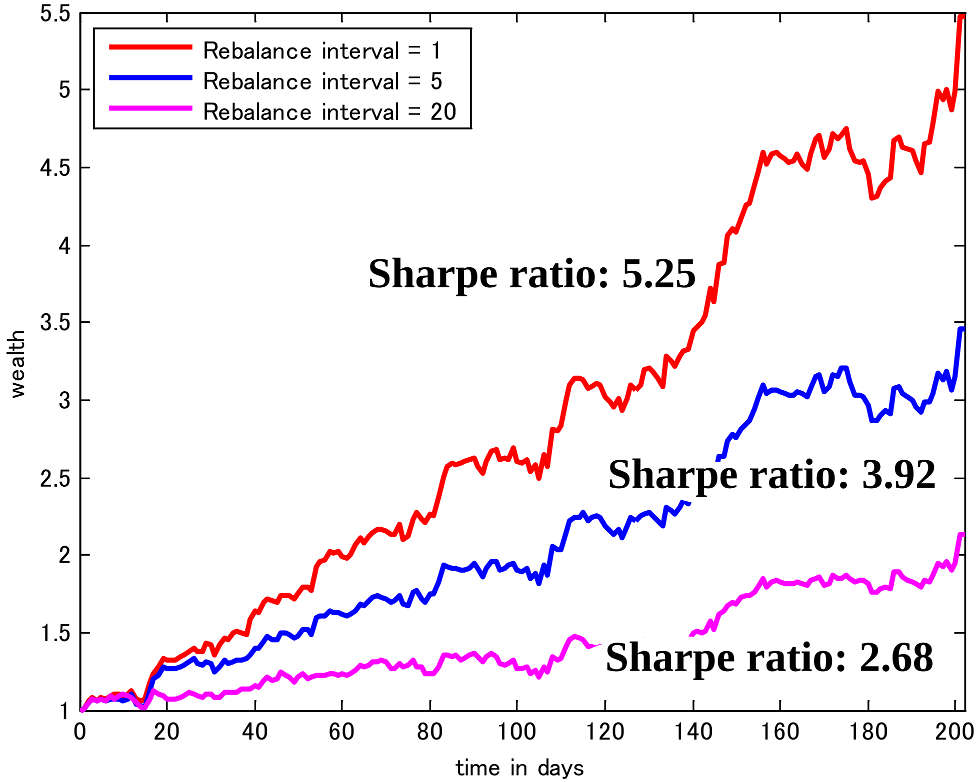
Prediction horizon vs. Sharpe ratio

- **Rebalance interval【δ】; 1 day**
- **Prediction horizon【τ】; 1day (myopic) – 160 days (8 months)**

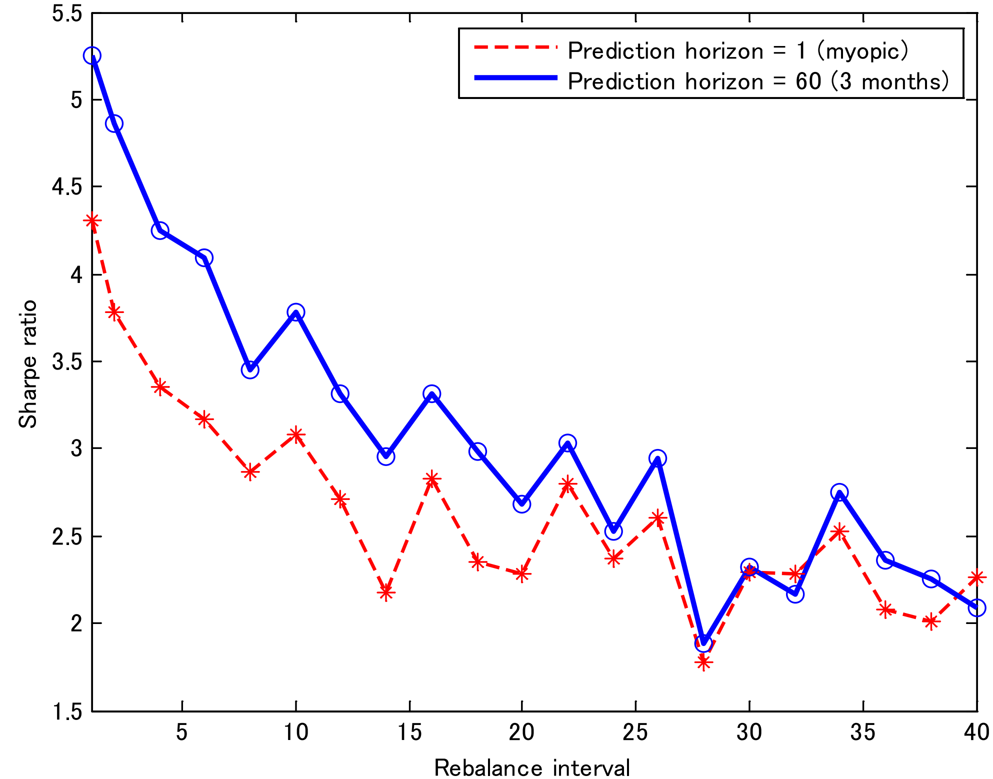


MPC with different rebalance interval

δ = 1 day, 5 days, and 20 days (τ = 60 days)



δ vs. Sharpe ratio (τ = 1 day and 60 days)

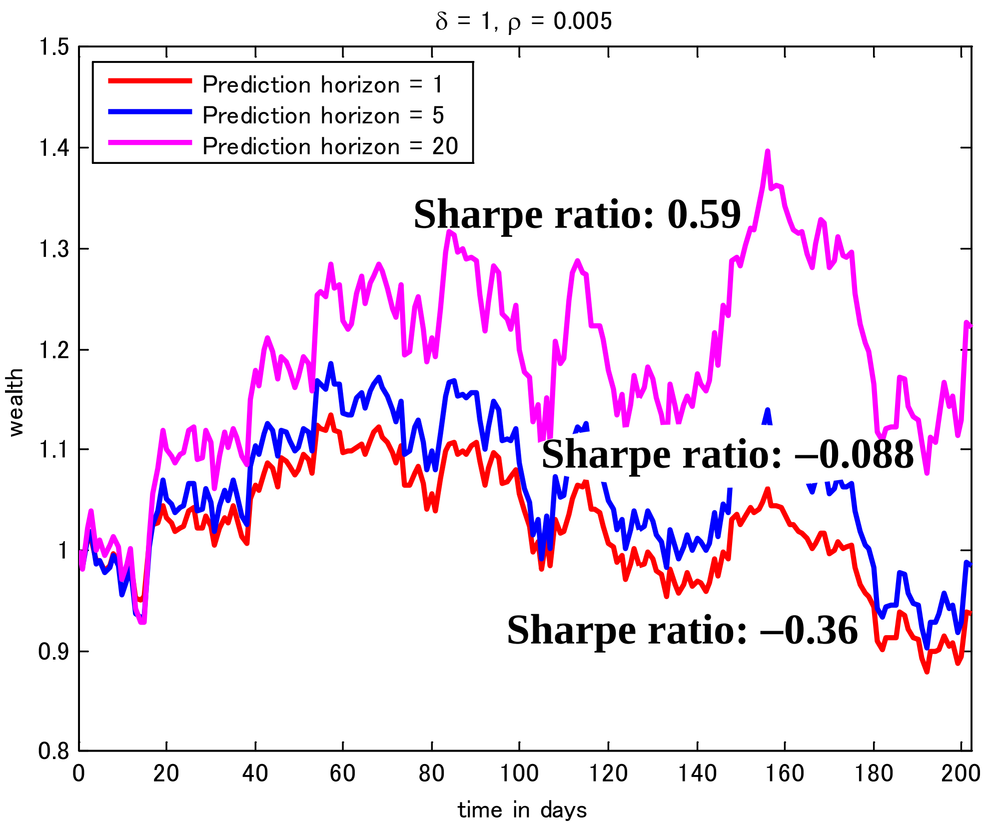


- A larger δ (less frequent rebalance) tends to provide a lower level of Sharpe ratio
- MPC is almost always better for rebalance interval of $\delta \leq 40$ days

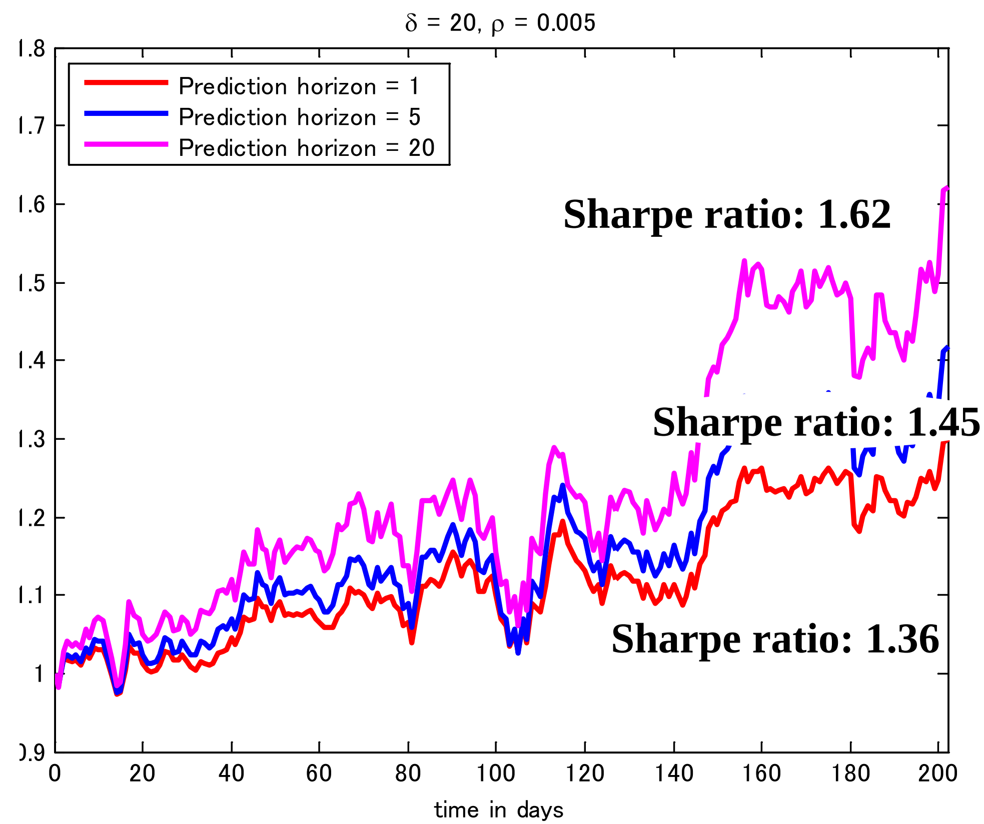
Effect of transaction costs (1)

- Proportional transaction cost of 0.5%
- Prediction horizon; 1 day, 5 days, and 20 days

Rebalance interval; 1 day



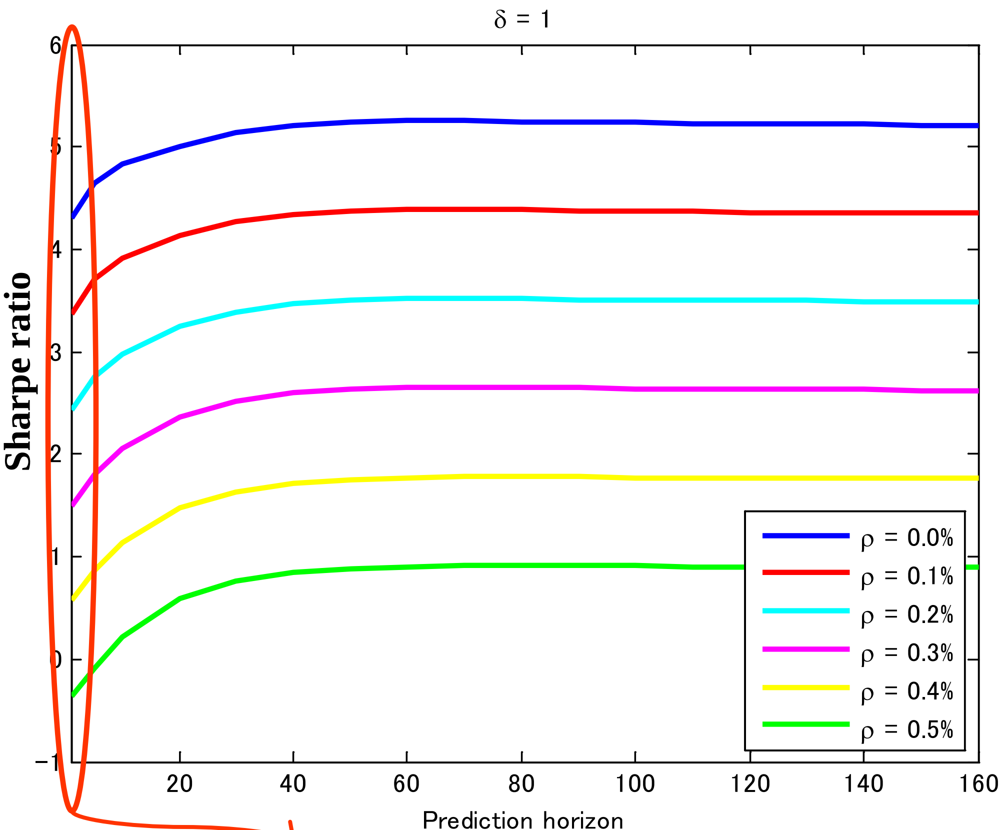
Rebalance interval; 20 days



Effect of transaction costs (2)

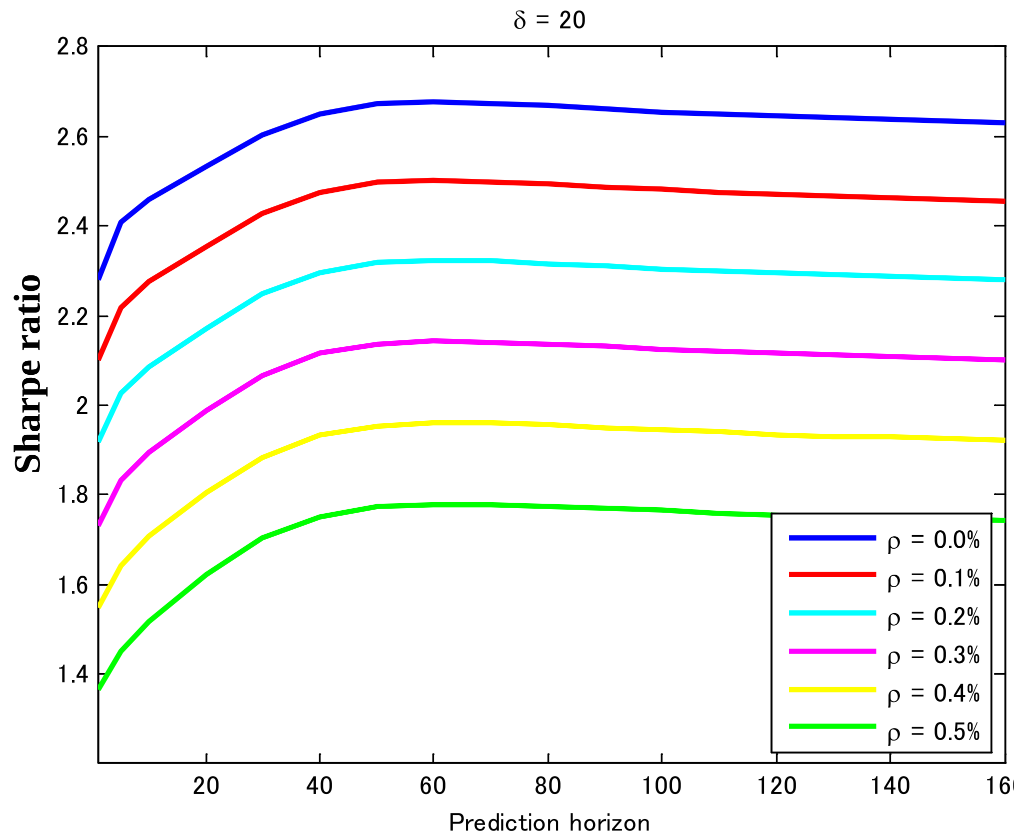
- **Prediction horizon vs. Sharpe ratio for different transaction cost rates**

Rebalance interval; 1 day



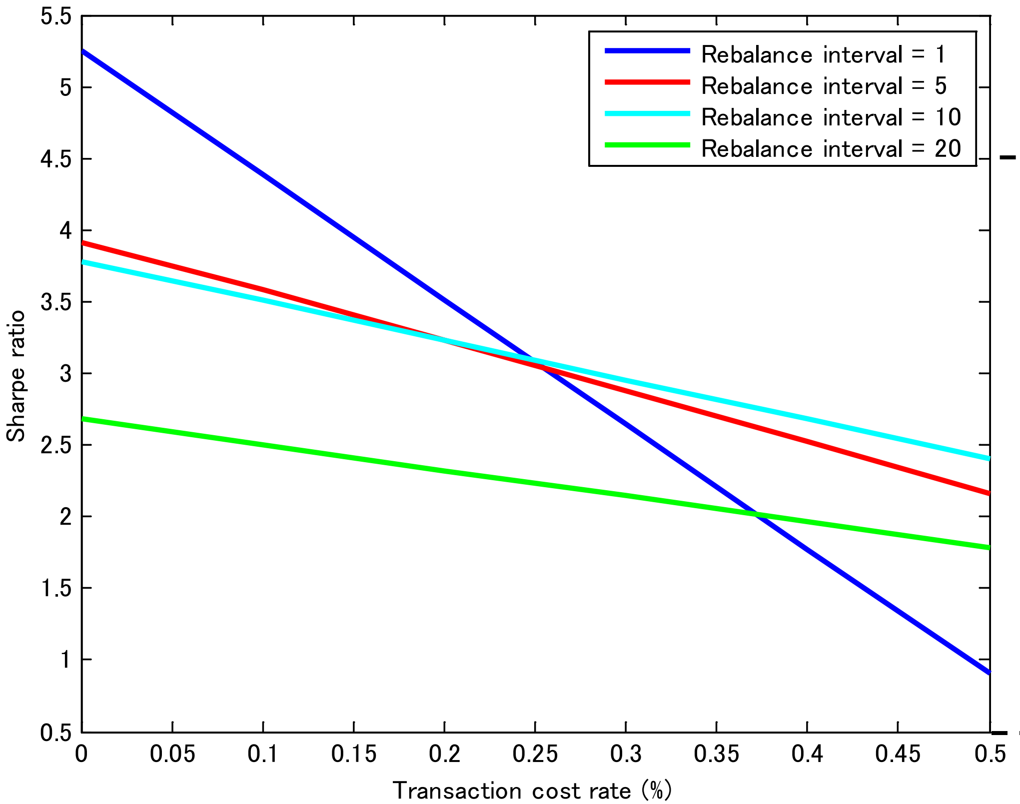
Myopic

Rebalance interval; 20 days



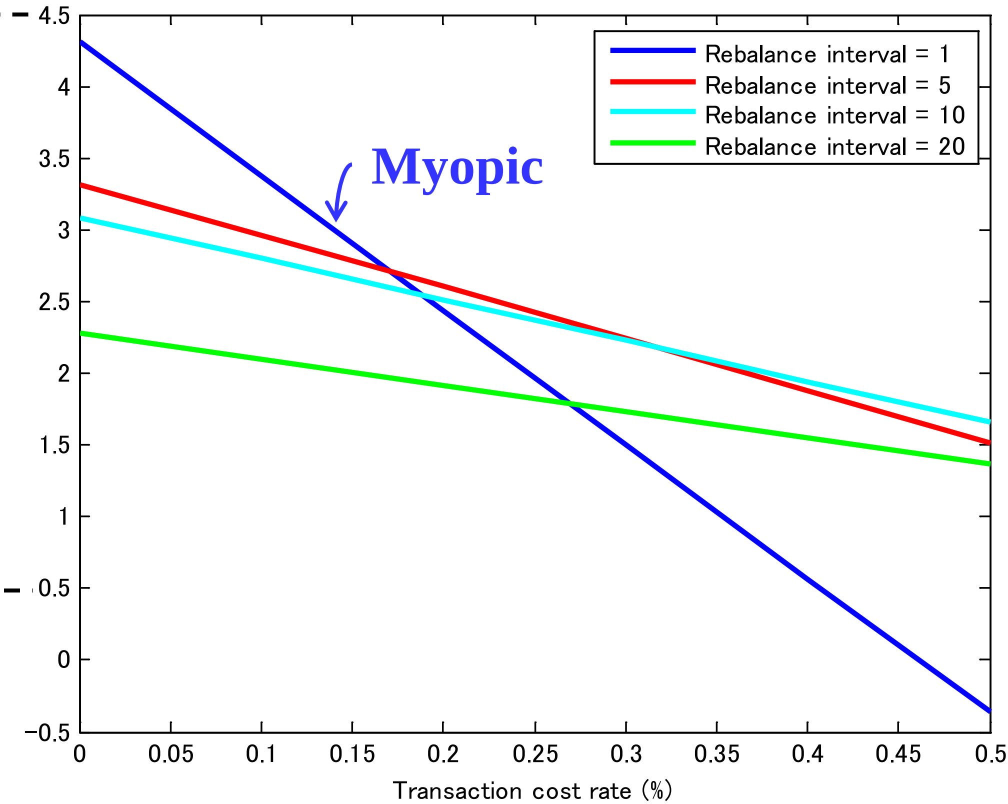
Effect of transaction costs (3)

- Transaction cost rates vs. Sharpe ratio for different rebalance interval, δ



MPC with different rebalance interval【τ = 60】

【τ = 1】



Conclusion

- **Summary**
  - **MPC with adequate prediction horizon may provide a better performance than Myopic strategy for both wealth level and Sharpe ratio.**
  - **Although the wealth drops significantly with the increase of transaction costs, their performance may be improved by adjusting rebalance interval.**
- **Remarks**
  - **We have tested recent data period for TOPIX500, S&P500, and HSI, and obtained a similar performance.**
  - **Proportional transaction cost constraint may be added in the objective function and is solved by quadratic optimization.**