Market Microstructure and High Frequency Finance Project

Małgorzata Konopacka, Artur Simon

3 02 2022

Based on paper: "Dynamic Trading with Predictable Returns and Transaction Costs".

Theoretical part

Model and Solution We consider an economy with S securities traded at each time $t \in \{\dots, -1, 0, 1, \dots\}$. The securities' price changes between times t and $t+1$ in excess of the risk-free rate return, $p_{t+1} - (1+r^f)p_t$, are collected in an $S \times 1$ vector r_{t+1} given by

$$r_{t+1} = Bf_t + u_{t+1}.$$

Here, f_t is a $K \times 1$ vector of factors that predict returns, B is an $S \times K$ matrix of factor loadings, and u_{t+1} is the unpredictable zero-mean noise term with variance $var_t(u_{t+1}) = \Sigma$.

The return-predicting factor f_t is known to the investor already at time t and it evolves according to

$$\Delta f_{t+1} = -\Phi f_t + \epsilon_{t+1}.$$

where $\Delta f_{t+1} = f_{t+1} - f_t$ is the change in the factors, Φ is a $K \times K$ matrix of mean reversion coefficients for the factors, and ϵ_{t+1} is the shock affecting predicitors with variance $var_t(\epsilon_{t+1}) = \Omega$.

The interpretation of these assumptions is straightforward: the investor analyzes the securities and his analysis results in forecasts of excess returns. The most direct interpretation is that the investor regresses the return of security s on the factors f that could be past returns over various horizons, valuation ratios, and other return-predicting variables, and thus estimates each variable's ability to predict returns as given by β^{sk} (collected in the matrix B).

ASSUMPTION 1: *Transaction costs are proportional to the amount of risk,* $\Lambda = \lambda\Sigma$.

The investor's objective is to choose the dynamic trading strategy $(x_0, x_1, \dots)$ to maximize the present value of all future expected excess returns, penalized for risks and trading costs,

$$\max_{x_0, x_1, \dots} E_0\Big[\sum_t (1-\rho)^{t+1}(x_t^T r_{t+1} - \frac{\gamma}{2} x_t^T \Sigma x_t) - \frac{(1-\rho)^t}{2} \Delta x_t^T \Lambda \Delta x_t\Big]$$

where γ is risk-aversion coefficient, and ρ is a discount rate.

Under Assumption 1, the optimal trading rate is the scalar $\alpha/\gamma < 1$, where

$$\alpha = \frac{-(\gamma(1-\rho) + \lambda\rho) + \sqrt{(\gamma(1-\rho) + \lambda\rho)^2 + 4\gamma\lambda(1-\rho)^2}}{2(1-\rho)}.$$

The trading rate is decreasing in transaction costs (λ) *and increasing in risk aversion* (γ).

The optimal portfolio is a weighted average of the existing portfolio x_{t-1} and the aim portfolio:

$$x_t = (1 - \frac{\alpha}{\lambda})x_{t-1} + \frac{\alpha}{\lambda} aim_t.$$

Proposition (Weight Signals Based on Alpha Decay) *Under Assumption 1, if the matrix* Φ *is diagonal,* $\Phi = diag(\phi^1, \ldots, \phi^K)$*, then the aim portfolio simplifies as the Markowitz portfolio with each factor* f_t^k *scaled down based on its own alpha decay* ϕ^k *:*

$$aim_t = (\gamma\Sigma)^{-1} B\Big(\frac{f_t^1}{1+\phi^1\alpha/\gamma}, \ldots, \frac{f_t^K}{1+\phi^K\alpha/\gamma}\Big)^T.$$

Practical part

We consider WIG20 Futures, a derivative product based on index of 20 Polish stocks which are listed on the main market. It is the most liquid derivative product on the Polish stock exchange.

First of all, we have to load the data and select only those columns, which are relevant for us, namely Date Close price and Volume. We consider the sample period January 3, 2000 to December 30, 2021.

```
WIG20 = fread("fw20_d.csv")[, .(Data, Zamkniecie, Wolumen)]
colnames(WIG20) = c("Date", "Close", "Volume")

WIG20
```

```
##             Date Close Volume
##    1: 03.01.2000  1920   4070
##    2: 04.01.2000  1847   4255
##    3: 05.01.2000  1811   5172
##    4: 06.01.2000  1848   5220
##    5: 07.01.2000  1982   5671
##   ---
## 5504: 23.12.2021  2231   9282
## 5505: 27.12.2021  2242  13797
## 5506: 28.12.2021  2269  16971
## 5507: 29.12.2021  2273  14454
## 5508: 30.12.2021  2279  11366
```

Then, we calculate the price changes.

```
WIG20[, Return := c(NA, diff(WIG20$Close))]
```

Summary Statistics are shown below.

Average.Price	Standard.Deviation.of.Price.Changes	Contract.Multiplyer	Daily.Trading.Volume
43589.02	663.76	20.00	25682.09

To predict returns, we use the characteristic-based model, where our derivative characteristic is its own past return at various horizons. Hence, to predict returns, we run a linear regression:

$$r_{t+1} = \beta_0 + \beta_1 \times f_t^{5D} + \beta_2 \times f_t^{1Y} + \beta_3 \times f_t^{5Y} + u_{t+1},$$

where the left-hand side is the daily future price change and the right-hand side contains the return predictors: f^{5D} is the average 5 days' price change divided by the past 5 days' standard deviation of daily price changes, f^{1Y} is the past year's average daily price change divided by the past year's standard deviation, and f^{5Y} is the analogous quantity for a 5-year window. Hence, the predictors are rolling Sharpe ratios over three different horizons.

The return predictors are chosen so that they have very different mean-reversion:

$$\begin{aligned}
\Delta f_{t+1}^{5D} &= -0.2519 f_t^{5D} + \epsilon_{t+1}^{5D}, \\
\Delta f_{t+1}^{1Y} &= -0.0034 f_t^{1Y} + \epsilon_{t+1}^{1Y}, \\
\Delta f_{t+1}^{5Y} &= -0.0001 f_t^{5Y} + \epsilon_{t+1}^{5Y}.
\end{aligned} \tag{1}$$

In order to calculate rolling Sharpe ratios we have to compute rolling averages and stanard deviations.

```
Returns_SD = data.table(
  "MA_5D" = frollmean(WIG20$Return, 5),
  "SD_5D" = frollapply(WIG20$Return, n = 5, FUN =  sd),
  "MA_1Y" = frollmean(WIG20$Return, 250),
  "SD_1Y" = frollapply(WIG20$Return, 250, sd),
  "MA_5Y" = frollmean(WIG20$Return, 5 * 250),
  "SD_5Y" = frollapply(WIG20$Return, n = 5 * 250, sd)
)

Reg = data.table(
  "REG_5D" = Returns_SD$MA_5D / Returns_SD$SD_5D,
  "REG_1Y" = Returns_SD$MA_1Y / Returns_SD$SD_1Y,
  "REG_5Y" = Returns_SD$MA_5Y / Returns_SD$SD_5Y
)

WIG20 = cbind(WIG20, Returns_SD, Reg)
```

The paper does not put any constrains on the distribution of shocks, so we assumed that shock to the sharpe ratios are normally distributed, i.e:

$$\begin{aligned}
\epsilon_t^{5D} &\sim \mathcal{N}(\mu^{5D}, \sigma^{5D}) \\
\epsilon_t^{1Y} &\sim \mathcal{N}(\mu^{1Y}, \sigma^{1Y}) \\
\epsilon_t^{5Y} &\sim \mathcal{N}(\mu^{5Y}, \sigma^{5Y})
\end{aligned}$$

Parameters of the distribution are estimated based on the whole data set. We use *fitdist* function from the package *fitdistrplus*.

```
fit_5D = fitdist(Reg$REG_5D[!is.na(Reg$REG_5D)], distr = "norm")$estimate
fit_1Y = fitdist(Reg$REG_1Y[!is.na(Reg$REG_1Y)], distr = "norm")$estimate
fit_5Y = fitdist(Reg$REG_5Y[!is.na(Reg$REG_5Y)], distr = "norm")$estimate
```

Having parameters, we can create a matrix of shocks to predictors.

```
shocks = data.frame(
    "5D" = rnorm(4257, fit_5D[1] * 1, fit_5D[2]),
    "1Y" = rnorm(4257, fit_1Y[1] * 1, fit_1Y[2]),
    "5Y" = rnorm(4257, fit_5Y[1] * 1, fit_5Y[2])
  )
```

Our first day based on which we will predict returns is the earliest possible day having calculated 5 Year rolling sharpe ratio, i.e. 27.12.2004. Next, we calculate predictors according to formula (1).

```r
WIG20 = na.omit(WIG20)


predictors = data.frame(
    "5D" = WIG20$REG_5D[1],
    "1Y" = WIG20$REG_1Y[1],
    "5Y" = WIG20$REG_5Y[1]
)

decay = c(0.2519, 0.0034, 0.0010)


for (i in 2:4258) {
  predictors[i, ] = predictors[(i-1), ] * (1 - decay) + shocks[(i-1), ]
}
```

We run linear regression and obtain:

```r
model_df = cbind(WIG20$Date, WIG20$Return, predictors)


model = lm(WIG20$Return[2:4258] ~ X5D[1:4257] + X1Y[1:4257] + X5Y[1:4257], model_df)
```

$$r_{t+1} = 35.115 - 21.262 f_t^{5D} - 14.097 f_t^{1Y} - 2.337 f^{5Y} + u_{t+1}.$$

We estimate variance Σ using daily price changes over the full sample. We set the absolute risk aversion to $\gamma = 10^{-6}$. The time discount rate is equal to $\rho = 1 - \exp(-0.02/260)$, corresponding to 2% annualized rate.

```r
gamma = 10 ^ (-6)

rho = 1 - exp(-0.02 / 260)
```

The scalar of transaction costs λ is basen on the paper *Measuring and modeling execution cost and risk*. According to Engle, Ferstanberg, and Russell traders amounting to 1.59 % of the daily volume in a stock have a price impact of about 0.1 %. We can write it as equation to solve:

$$1.59\% \times Volume \times \frac{\lambda}{2} \times \sigma^2 = 0.1\% \times Price.$$

Using summary statistics (average volume, standard deviation and average price) we obtain

```r
lambda = 2 * ((0.001 * avg_price) / (0.0159 * avg_vol * sd_return ^ 2))

lambda
```

```
## [1] 4.84576e-07
```

With everything we need, we prepare trading strategies according to Examples.

Example 1 (Timing a Single Security): A simple case is when there is only one security. This occurs when an investor is timing his long or short view of a particular security or market. In this case, Assumption 1

$(\Lambda = \lambda\Sigma)$ is without loss of generality since all parameters are scalars, and we use the notation $\sigma^2 = \Sigma$ and $B = (\beta^1, \ldots, \beta^K)$. Assuming that Φ is diagonal, we can apply Proposition directly to get the optimal timing portfolio:

$$x_t = (1 - \frac{\alpha}{\lambda})x_{t-1} + \frac{\alpha}{\lambda}\frac{1}{\gamma\sigma^2}\sum_{i=1}^{K}\frac{\beta^i}{1+\phi^i\alpha/\gamma}f_t^i$$

```
alpha = (-(gamma * (1 - rho) + lambda * rho)
         + sqrt((gamma * (1 - rho) + lambda * rho) ^ 2 + 4 * gamma *
                                    lambda * (1 - rho) ^ 2) ) / (2 * (1 - rho))
```

```
predictors = cbind(1, predictors)

# Optimal portfolio
X_vec = 0

decay = c(0, decay)
for (i in 2:4258) {
  X_vec[i] = (1 - alpha / lambda) * X_vec[(i - 1)] + alpha / lambda *
   (1 / (gamma * sd_return ^ 2)) *
    sum(predictors[(i - 1), ] * model$coefficients / (1 + decay * alpha / gamma))
}
```

Example 2 (Markovitz model) The aim portfolio in their dynamic setting turns out to be closely related to the optimal portfolio in a static model without transaction costs $(\Lambda = 0)$, which we call the Markowitz portfolio.

$$Markovitz_t = (\gamma\Sigma)^{-1}Bf_t$$

```
# Markovitz portfolio
Markovitz = (gamma * sd_return ^ 2) ^ (-1) * apply(predictors * model$coefficients, 1, sum)
```

Example 3 (Static Model): Consider an investor who performs a static optimization involving current expected returns, risk, and transaction costs. Such an investor simply solves

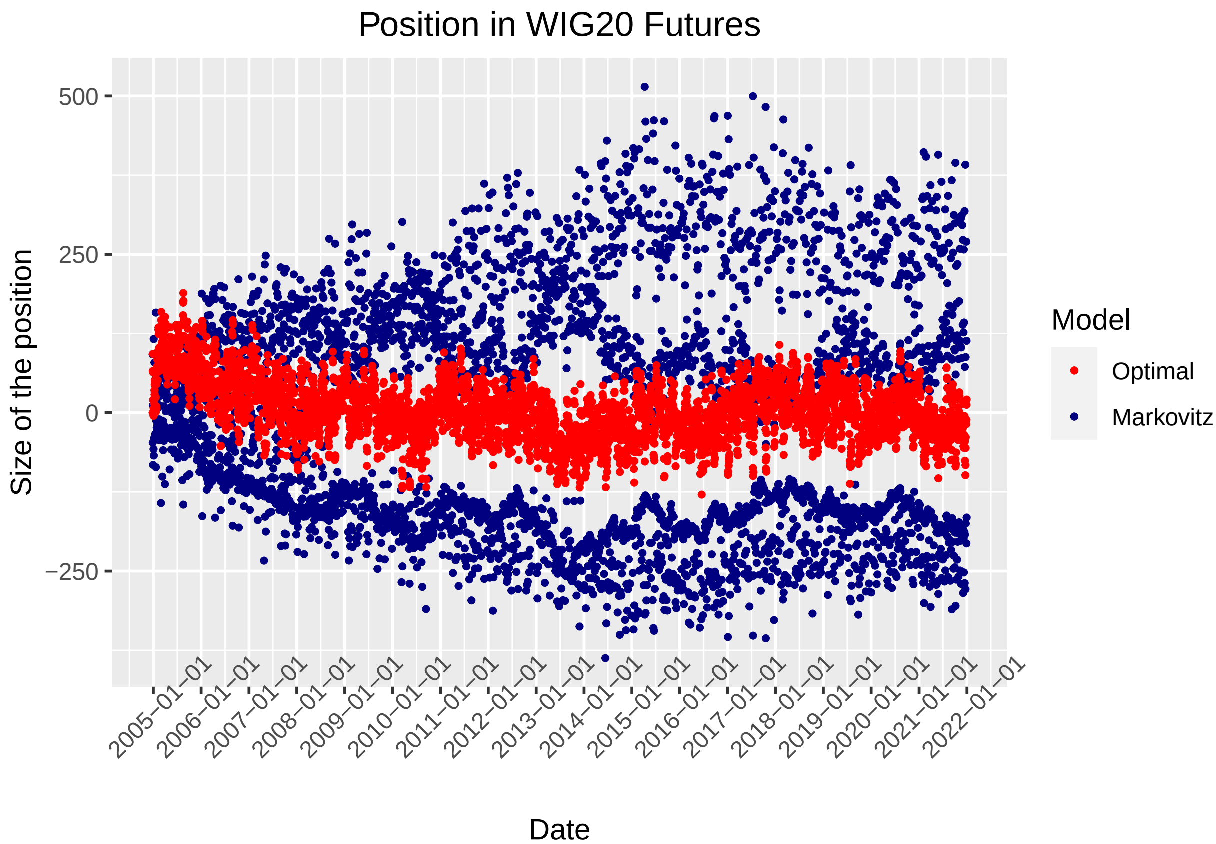
$$\max_{x_t} x_t^T E_t(r_{t+1}) - \frac{\gamma}{2}x_t^T\Sigma x_t - \frac{\lambda}{2}\Delta x_t^T\Sigma\Delta x_t,$$

with solution

$$x_t = \frac{\lambda}{\gamma+\lambda}x_{t-1} + \frac{\lambda}{\gamma+\lambda}(\gamma\Sigma)^{-1}E_t(r_{t+1}) = x_{t-1} + \frac{\gamma}{\gamma+\lambda}(Markovitz_t - x_{t-1}).$$

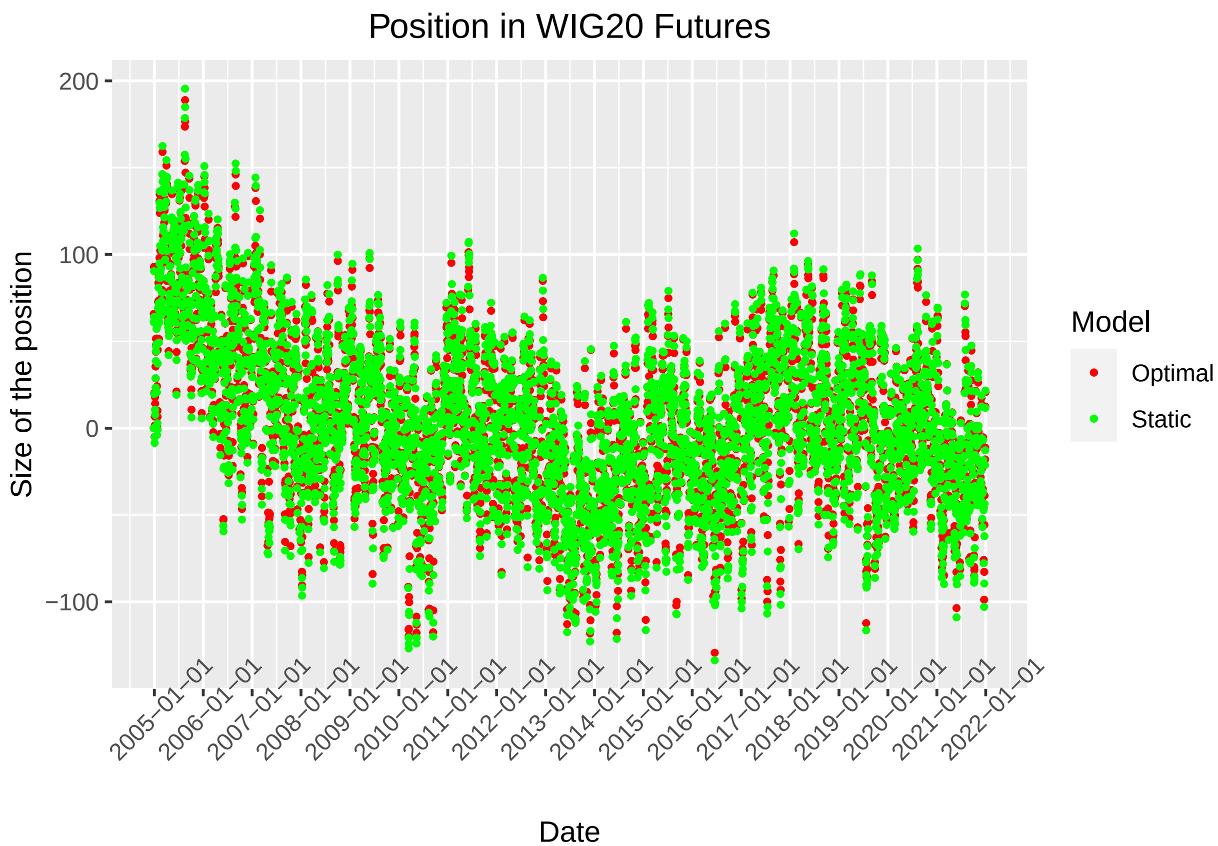
```
Static = 0
for (i in 2:4258) {
  Static[i] = Static[(i - 1)] +
    gamma / (gamma + lambda) * ( (gamma * sd_return ^ 2) ^ (-1)
                                * sum(predictors[(i-1), ] * model$coefficients)
                                - Static[(i-1)])
}

results_df = data.frame(
    Date = as.Date(WIG20$Date, tryFormats = c("%d.%m.%Y")),
    X = 1:length(X_vec),
    Optimal = X_vec,
    Markovitz = Markovitz,
    Static = Static
  )
```

Results To ilustrate the difference in the possitions of the different strategies, we present figures with the position over time.



The optimal portfolio is a smoother version of the Markovitz strategy. Hence, it reduces trading cost while at the same time capture most of the exscess return. The optimal position tends to be on the same side (long/short) as Markovitz portfolio. Long positions tend to be bigger than short, and as the time passes, Markovitz portfolio has larger and larger net positions, while Optimal strategy is relatively stable when it comes to size of the position.

Now we can compare static and optimal portfolio.



They behave almost identical. The differenes are very little.

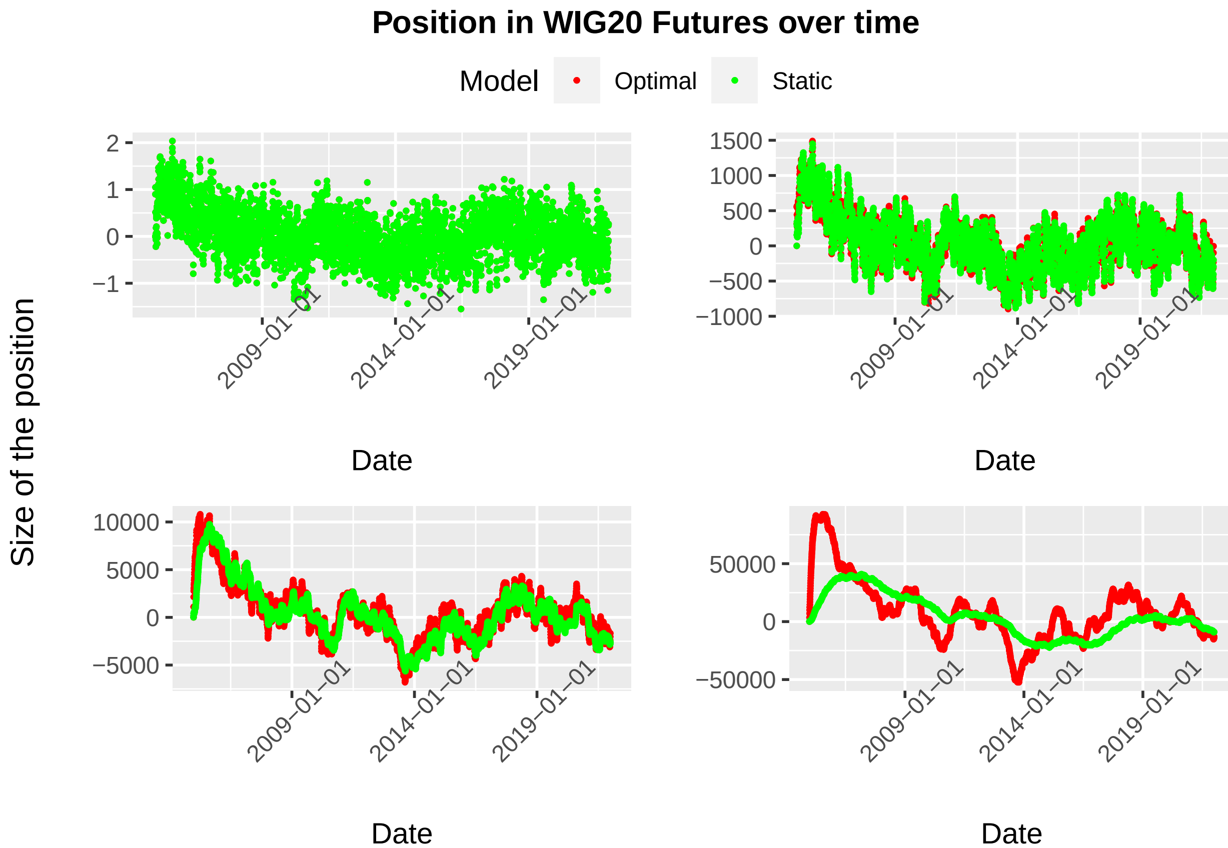
In the table below, they are presented Sharpe ratios of the portfolios.

	Daily	Annualized
Optimal	0.031	0.496
Static	0.030	0.482
Markovitz	0.017	0.274

Optimal portfolio has the biggest sharpe ratio. However, static strategy is not far behind. Markovitz portfolio performs the worst. It is significantly worse than the other two.

We can compare the behaviour of the portfolios for different values of γ.We will do this for $\gamma = \{10^{-4}, 10^{-7}, 10^{-8}, 10^{-9}\}$

Increase of risk-aversion causes that the size of the position over time in optimal portfolio is much smoother. The behaviour of the Markovitz and Optimal portoflio does not change significantly.



For the first two scenarios there is a little difference between Optimal and Static strategies. When portfolio is becomming more risk-averse, divergence is more visible. For $\gamma = 10^{-8}$ they move always in the same direction, but Optimal portfolio has bigger net positions. For $\gamma = 10^{-9}$ there is a visible difference. Static portfolio is much smoother with less changes between days.

The table below presents the sharpe ratios of portfolios for different value of γ. The value in column *gamma* is the power to which 10 has been raised.

gamma	Optimal	Static	Markovitz
-4	0.496	0.496	0.274
-7	0.346	0.347	0.274
-8	0.365	0.448	0.274
-9	0.339	0.542	0.274

We can observe that the Sharpe ratio of Markovitz portfolio is not affected by changing risk aversion. It has the same value for different γ. Optimal portfolio perform worse when we increase risk aversion. However, for every scenario it is better than Markovitz. Static portfolio is not consistent. Its sharpe ratio is the lowest for $\gamma = 10^{-7}$, but the highest for the most risk-averse scenario. Static portfolio outperforms Optimal, except the most riskiest portfolios.