Dynamic Asset Allocation with Risk Factors

Sihang Liu, Meihui Zhang

[1]

Abstract: Finding the optimal portfolio allocation is a hard process for investors. In this paper, we introduced a way to dynamically change the portfolio each year for investors to accumulate their wealth in a safer way base on their risk aversion level and the market environment. In our model, we included two risk factors to represent the market fluctuation. During the time that the market is really fluctuate, investors are able to increase the frequency of weight adjustment in the portfolio and introduce more risk factors to the model to better fit the market.

1. Motivation:

All investors have the goal to gain higher return with lower risk for the lowest cost, it becomes an essential challenge for investors to pick a good investment strategy. As human beings, we can easily be influenced by emotion while we are making investment decisions. For instance, the prospect theory introduced by Daniel Kahneman and Amos Tversky in 1979, says that people would have an asymmetric manner while facing the equal amount of loss and gain. In this case, although we assume people are all rational in the economic theory, we can usually observe investors making irrational decisions. Here dynamic asset allocation becomes an answer to this puzzle.

Dynamic asset allocation as the core content of the modern portfolio theory. It is based on the different performance of broad categories of assets in different economic cycles and market environments, in order to effectively reduce the volatility of the entire asset portfolio or significantly enhance the expected return level overtime for investors with any utility preferences based on market valuation.

In our project, we have three goals. First, we want to test whether a leading indicator is a good risk factor for us to predict the market, and make asset allocation decisions base on that. Secondly, we want to show how does the risk aversion level influence our model. Lastly, we also want to see how does the interest rate influence the investment opportunity set.

2. Model and Settings

To accomplish our first goal to test the validity of the leading indicator in the asset allocation problem, we first pick a leading indicator we are interested in. We choose the seasonally adjusted New Privately Owned Housing Units Started(NPOHUS) as the indicator. It is a good indicator for economists to predict the future movement of the economy. While the market will be in a up trend, the NPOHUS tends to goes up before that. Besides that, we pick Consumer Price Index(CPI) as the second risk factor.

In order to find the optimal utility choice, we use the martingale approach. Since we are interested in the dynamic problem, now we have:

$$\begin{aligned} &\max_{\pi} E\left[U\left(X_T\right)\right] \\ &s\cdot t \quad dX_t = x_t r_t dt + x_t \pi_t' \sigma_t \left(\theta_t dt + dw_t\right) \\ &x_0 = x \end{aligned} \tag{1}$$

In this case, we want to maximize our terminal utility with given initial wealth. We can simplify the above problem to the equivalent static problem.

$$\begin{aligned} &\max_{X} E\left[U\left(X_T\right)\right] \\ &s\cdot t \quad E\left[\xi_T X_T\right] \leq x_0 = x \end{aligned} \tag{2}$$

ξ_T is the state price density of martingale measure. It is also called the risk neutral density. It is a measure of financial risk and a sufficient statistical quantity for the pricing of derivative securities. It gathers all the information related to pricing and business conditions. that $\xi_T = \frac{Z_T}{b_T}$. And we define $Z_T = \epsilon(-\int_0^{\cdot} \theta_v' dw_v)_T = exp(-\int_0^T \theta_v' dw_v - \frac{1}{2}\int_0^T \|\theta_v\|^2 dv)$ by stochastic exponential, and $b_T = exp(\int_0^T r_v dv)$.

In order to solve the static problem, we need to introduce a new Lagrange multiplier y. It is also the shadow price of the wealth. It represents the maximum price that the management is willing to pay for an additional unit's established resources.

Then we can set the Lagrange function as:

$$\mathbf{L} = E\left[U(X_T\right] - y(E\left[\xi_T X_T\right] - X) \tag{3}$$

Next, we need to find the y such that the constraint holds. And we get:

$$\begin{aligned}
&\max_{X_T} \mathbf{L} = \max_{X_T} E\left[U(X_T) - y(\xi_T X_T - X)\right] \leq E\left[maxU(X_T) - y(\xi_T X_T - X)\right] \\
&F.O.C \quad U'(X_T^*) = y\xi_T \\
&X_T^* = I(y\xi_T)
\end{aligned} \tag{4}$$

In order to solve this problem, we need to assume the utility function is strictly increasing, and the second order derivative of utility is negative, which means the marginal utility is decreasing. In this case, U' would has an inverse $I(y)$ such that $U'(I(y)) = y$. We could finally get find $y^*(x)$ such that $E\left[\xi_T I(y^*(x)\xi_T)\right] = x$. By plug it in equation (4), we could get the optimal terminal wealth as:

$$X_T^* = I(y^*(x)\xi_T) \tag{5}$$

By using the risk neutral valuation, we could get:

$$\begin{aligned}
X_t^* &= E\left[\frac{\xi_T}{\xi_t} I(y^*(x)\xi_T)|\mathbf{F}_t\right] \\
&= E\left[\xi_{t,T} I(y^*(x)\xi_t\xi_{t,T})|\mathbf{F}_t\right] \\
&= E\left[\xi_{t,T} X_T^*|\mathbf{F}_t\right] \qquad by(5)
\end{aligned}$$

$\xi_{t,T}$ measures the future impact, and ξ_t measures the immediate impact. Then we need to introduce the Mallivian Derivative which shows the volatility of wealth: $D_t X_t^* = X_t^* \pi_t' \sigma_t$.

$$D_t X_t^* = D_t E\left[\xi_{t,T} X_T^* | \mathbf{F}_t\right] = E\left[D_t \xi_{t,T} X_T^* | \mathbf{F}_t\right] \tag{6}$$

$$X_t^* \pi_t' \sigma_t = E\left[D_t \xi_{t,T} X_T^* | \mathbf{F}_t\right] = E\left[D_t \xi_{t,T} I(y^*(x)\xi_t \xi_{t,T}) | \mathbf{F}_t\right] \tag{7}$$

$$let \qquad D_t \xi_{t,T} I(y^*(x)\xi_t \xi_{t,T}) = D_t g(\xi_t, \xi_{t,T}) = \frac{\partial g(\xi_t, \xi_{t,T})}{\partial \xi_t} D_t \xi_t + \frac{\partial g(\xi_t, \xi_{t,T})}{\partial \xi_{t,T}} D_t \xi_{t,T} \tag{8}$$

$$D_t \xi_t = -\xi_t \theta_t' \qquad and \qquad D_t \xi_{t,T} = \xi_{t,T} D_t log \xi_{t,T} \tag{9}$$

$$\frac{\partial g(\xi_t, \xi_{t,T})}{\partial \xi_t} = \xi_{t,T} I'(y^*(x)\xi_t \xi_{t,T}) y^*(x) \xi_{t,T} \tag{10}$$

Note: Here $I'(y^*(x)\xi_t \xi_{t,T})$ is the derivative with respect to ξ_t.

$$\frac{\partial g(\xi_t, \xi_{t,T})}{\partial \xi_{t,T}} = I(y^*(x)\xi_t \xi_{t,T}) + \xi_{t,T} I'(y^*(x)\xi_t \xi_{t,T}) y^*(x) \xi_t \tag{11}$$

Note: Here $I'(y^*(x)\xi_t \xi_{t,T})$ is the derivative with respect to $\xi_{t,T}$. At last we can plug in the partial derivative we just got to equation (7), and get:

$$\begin{aligned} X_t^* \pi_t' \sigma_t &= -E\left[\xi_{t,T} I'(y^*(x)\xi_t \xi_{t,T}) y^*(x) \xi_t \xi_{t,T} | \mathbf{F}_t\right] \theta_t' \\ &+ E\left[\xi_{t,T} (I(y^*(x)\xi_t \xi_{t,T}) + I'(y^*(x)\xi_t \xi_{t,T}) y^*(x) \xi_t \xi_{t,T}) D_t log \xi_{t,T} | \mathbf{F}_t\right] \end{aligned} \tag{12}$$

Next, we need to introduce $\gamma(x)$ absolute risk tolerance into the model. $\gamma(x)$ is defined as $-\frac{U'(x)}{U"(x)}$. It is the inverse of absolute risk aversion. Now we can represent $I'(y^*(x)\xi_t \xi_{t,T}) y^*(x) \xi_t \xi_{t,T}$ as $-\gamma(X_T^*)$. Then we represent $D_t log \xi_{t,T}$ as $-H_{t,T}$. Then we can rewrite equation (12) in a more meaningful format.

$$X_t^* \pi_t' \sigma_t = E\left[\xi_{t,T} \gamma(X_T^*) | \mathbf{F}_t\right] \theta_t' + E\left[\xi_{t,T} (\gamma(X_T^*) - X_T^*) H_{t,T} | \mathbf{F}_t\right] \tag{13}$$

$$X_t^* \pi_t' = E\left[\xi_{t,T} \gamma(X_T^*) | \mathbf{F}_t\right] \pi_t^{m\prime} + E\left[\xi_{t,T} (\gamma(X_T^*) - X_T^*) H_{t,T} | \mathbf{F}_t\right] \sigma_t^{-1} \tag{14}$$

For the second half of equation (14), we can divide it further to two parts of interest rate hedge and the market price of risk hedge.

$$Given \quad -log\xi_{t,T} = \int_t^T r_s ds + \frac{1}{2}\int_t^T \|\theta_s\|^2 + \int_t^T \theta_s' dw_s \tag{15}$$

$$Get \quad H_{t,T} = -D_t log\xi_{t,T} = \int_t^T D_t r_s ds + \int_t^T \theta_s' D_t \theta_s ds + \int_t^T dw_s D_t \theta_s \tag{16}$$

$$Let \quad H_{r,T} = H_{t,T}^r + H_{t,T}^\theta \tag{17}$$

$$Get \quad H_{t,T}^r = \int_t^T D_t r_s ds \quad and \quad H_{t,T}^\theta = \int_t^T (dw_s + \theta_s' ds)' D_t \theta_s \tag{18}$$

Now we get the closed form for the interest rate hedge and market price or risk hedge part of H. At last, we can plug these two values back to equation (14) to get the explicit form of interest rate hedge and market price of risk hedge of the optimal portfolio.

$$Interest\ rate\ hedge \quad X_t^* \pi_t^{r\prime} = E\left[\xi_{t,T}(\gamma(X_T^*) - X_T^*)H_{t,T}^r|\mathbf{F}_t\right]\sigma_t^{-1} \tag{19}$$

$$Market\ price\ of\ risk\ hedge \quad X_t^* \pi_t^{\theta\prime} = E\left[\xi_{t,T}(\gamma(X_T^*) - X_T^*)H_{t,T}^\theta|\mathbf{F}_t\right]\sigma_t^{-1} \tag{20}$$

Finally, the optimal portfolio would be:

$$X_t \pi_t^* = E\left[\xi_{t,T}\gamma(X_T^*)|\mathbf{F}_t\right]\pi_t^{m\prime} + X_t \pi_t^r + X_t \pi_t^\theta \tag{21}$$

During our research, we assume the investor has the constant relative risk aversion(CRRA) utility function, $U(X) = \frac{X^{1-R}}{1-R}$.

$$\gamma(X) = -\frac{U'(X)}{U"(X)} = \frac{1}{R}X \tag{22}$$

$$X_t^* \pi_t^* = E\left[\xi_{t,T}X_T^*|\mathbf{F}_t\right]\frac{\pi_t^{m\prime}}{R} + E\left[\xi_{t,T}X_T^* H_{t,T}^r|\mathbf{F}_t\right]\sigma_t^{-1}(\frac{1}{R} - 1) + E\left[\xi_{t,T}X_T^* H_{t,T}^\theta|\mathbf{F}_t\right]\sigma_t^{-1}(\frac{1}{R} - 1) \tag{23}$$

$$\pi_t^* = \frac{\pi_t^{m\prime}}{R} + E\left[\frac{\xi_{t,T}X_T^* H_{t,T}^r}{E_t\left[\xi_{t,T}X_T^*\right]}|\mathbf{F}_t\right]\sigma_t^{-1}(\frac{1}{R} - 1) + E\left[\frac{\xi_{t,T}X_T^* H_{t,T}^\theta}{E_t\left[\xi_{t,T}X_T^*\right]}|\mathbf{F}_t\right]\sigma_t^{-1}(\frac{1}{R} - 1) \tag{24}$$

As we know, when $R = 1$, $U(X) = log(X)$, and in this case, $\pi_t^* = \pi_t^m$. So, we would then produce the mean variance portfolio of our chosen asset classes.

3. Data

In order to accomplish our goals of this project, we decide to include two asset classes in our research - the US stock, and the US small-cap market. We used the NASDAQ index to represent US stock market, and use Russel 2000 to represent the US small-cap market. We also picked two risk factors. As we know, the new privately owned housing unit started in each month could be a good leading indicator of the market. Using this risk factor could provide a good prediction of the market of future. Besides, we pick consumer price index which measure the level of inflation as the second risk factor to indicate the market. Because of the availability of each data, we choose the monthly data from 1991/1/1-2020/10/1. Then we get the following equation of excess return with respect to two risk factors:

$$\begin{aligned}\frac{dS_t^j}{S_t^j} - r_t dt &= \beta_1^j + \beta_2^j \frac{dX_t^{Real\,Estate}}{X_t^{Real\,Estate}} + \beta_3^j \frac{dX_t^{CPI}}{X_t^{CPI}} + d\epsilon_{t,j} \\ &= \sigma_{t,j}' \theta_t dt + dw_t\end{aligned} \tag{25}$$

$$\rightarrow \beta_1^j + \beta_2^j \frac{dX_t^{Real\,Estate}}{X_t^{Real\,Estate}} + \beta_3^j \frac{dX_t^{CPI}}{X_t^{CPI}} = \sigma_{t,j}' \theta_t dt \tag{26}$$

In this case, we could calculate the value for θ_t. In the research, we assume annually interest rate to be constant as 2%.

4. Results

All investors follows a two steps allocation processes in the optimal allocation between risky and risk free asset. First, we make the investment decision by identify the mean variance portfolio which is the same for every investor. Secondly, we make the financing decision by splitting between the risk free asset and the mean variance portfolio base on investor's risk aversion level which has been shown in part 2: Model and Settings.

In order to find the mean variance portfolio, we find the log-return of each risk factor, we use the Skit-learn package to run the linear regression of equation (25), and get:

	NASDAQ	Russel
Intercept	0.08264585464772126	0.07250765706532748
β_1	0.08046760204113716	0.24286242321373963
β_2	-6.001230334634639	-6.446614831318217

Then, we need to calculate the mean-variance portfolio weights without the risk-free asset at September every year. During the whole calculation, we set the risk aversion equaling to R = 1.5. Take year 2011 as an example. The expected returns of NASDAQ and Russel were based on the historical price data in 2010. We set them as the average of annually returns calculated by monthly returns. And the covariance matrix was derived by the annualized monthly returns.

Date	var1	var2	cov
2011/9/1	0.34149277	0.72828689	0.47556895
2012/9/1	0.28123779	0.19470988	0.21803767
2013/9/1	0.08643388	0.14115883	0.09471448
2014/9/1	0.12342784	0.29869273	0.16639596
2015/9/1	0.2814945	0.21718719	0.19918666
2016/9/1	0.24707508	0.32257691	0.25586398
2017/9/1	0.02456283	0.16805477	-0.008764
2018/9/1	0.28603318	0.30069356	0.25071477
2019/9/1	0.46608274	0.27129283	0.28051597
2020/9/1	0.75106415	1.29480157	0.87763908

Figure 1. Covariance matrix data

Repeat the procedures several times. By plugging these pairs of two inputs, we can get the following table of mean-variance weight without the risk-free asset every year for NASDAQ and Russel.

Mean Variance Weight 2011-2020					
	NASDAQ	Russel		NASDAQ	Russel
2011	2.1301	-1.1301	2016	1.1517	-0.15170
2012	-0.5851	1.5851	2017	0.8414	0.1586
2013	1.2170	-0.2170	2018	0.5859	0.4141
2014	1.4810	-0.4810	2019	-0.0523	1.0523
2015	0.1795	0.8205	2019	1.4356	-0.4356

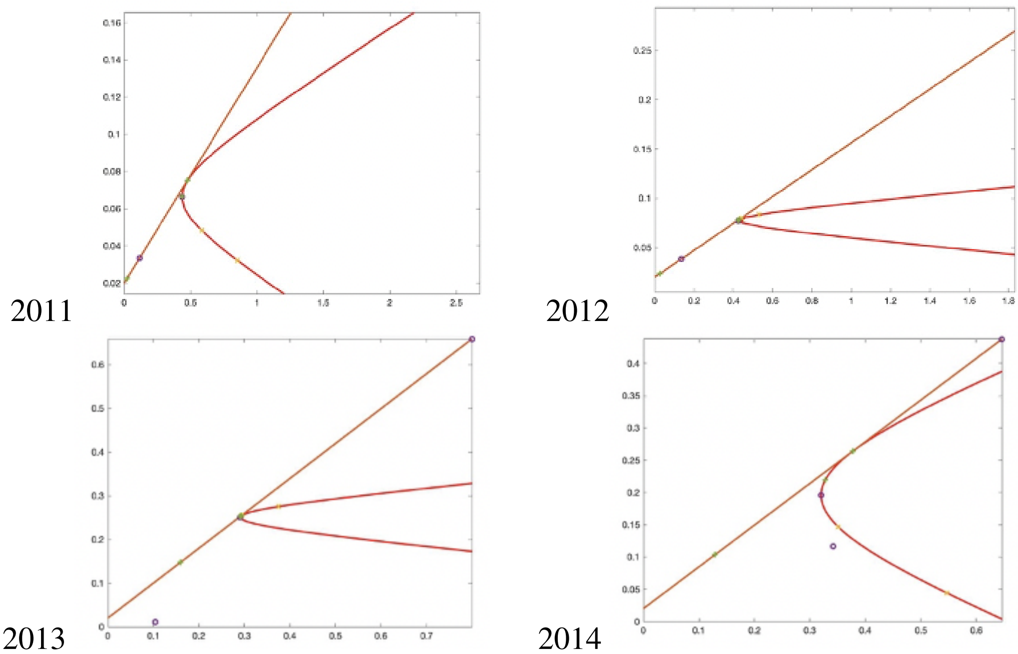

Figure 2. Efficient Frontiers r = 2%

Since the risk-free rate in our model is a constant rather than a stochastic process, we only need to take care about the market price of risk hedge and eliminate the interest rate hedge. The assumption in our model indicates that there is no intermediate consumption. We used Monte Carlo simulation method to calculate the expectation parts in the market price of risk hedge formula. By changing the weight once each year, we get the following dynamic weight on risk free asset, NASDAQ, and Russel from 2011 to 2020.

Date	rf	Nasdaq	Russel
2011/9/1	0.4852	1.09657548	-0.5817755
2012/9/1	0.3785	-0.3636397	0.98513965
2013/9/1	0.3699	0.7668317	-0.1367317
2014/9/1	0.3789	0.9198491	-0.2987491
2015/9/1	0.3023	0.12523715	0.57246285
2016/9/1	0.4125	0.67662375	-0.0891238
2017/9/1	0.4148	0.49238728	0.09281272
2018/9/1	0.4262	0.33618942	0.23761058
2019/9/1	0.3989	-0.0314375	0.63253753
2020/9/1	0.3728	0.90040832	-0.2732083

Figure 3. Dynamic Weight R=1.5

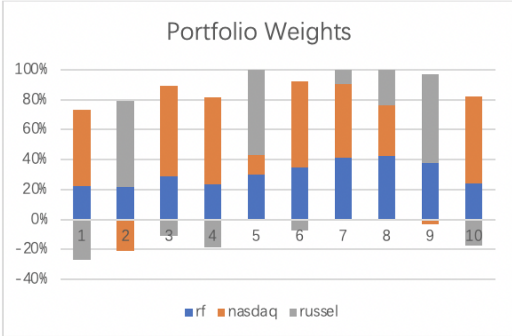

Figure 4. Portfolio Weight Bar Diagram

5. Conclusions

By doing this project, we can see, using New Privately Owned Housing Units Started and CPI as the risk factor to predict the expected return of NASDAQ and Russel in our model performs pretty well. After considering about the market price of risk hedge part in our portfolio, and include the interest risk hedge if the interest rate is not deterministic, we could get a really good return over all. The wealth increased by 70% through 10 years. The graphs down below, shows the accumulation of wealth through 2011 to 2020 annually and monthly without intermediate consumption.

Figure 5. Annually Wealth Curve

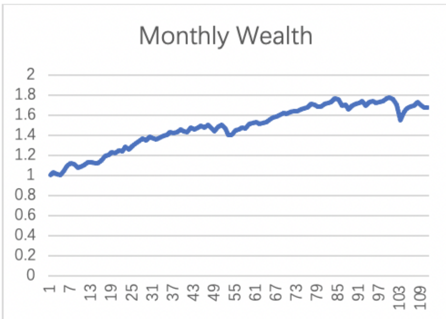

Figure 6. Monthly Wealth Curve 2011-2020

Next, we reset the risk aversion level from 1.5 to 1.6. We want to see does the weight set follows our intuition that when the risk aversion level rises, investors should put more weight in the risk free asset rather than risky asset. And we get the following result. Compare them to Figure 3 and 4, it verifies our hypothesis. A small increase in the risk aversion value would boost weight in risk free asset by a lot, because investors are more sensitive to the losses now.

Date	rf	Nasdaq	Russel
9/1/11	0.6083	0.83436017	-0.4426602
9/1/12	0.5289	-0.2756406	0.74674061
9/1/13	0.5191	0.5852553	-0.1043553
9/1/14	0.5262	0.7016978	-0.2278978
9/1/15	0.4667	0.09572735	0.43757265
9/1/16	0.552	0.5159616	-0.0679616
9/1/17	0.5539	0.37534854	0.07075146
9/1/18	0.5628	0.25615548	0.18104452
9/1/19	0.5416	-0.0239743	0.48237432
9/1/20	0.5214	0.68707816	-0.2084782

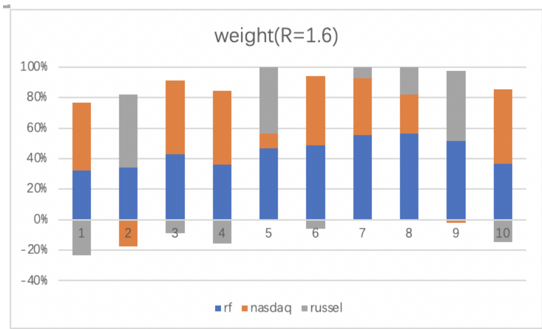

Figure 7. Dynamic Weight R=1.6

Figure 8. Portfolio Weight Bar Diagram

At last, we change the constant interest rate from 2% to 5%, in order to see how does the interest level influence the investment opportunity set. As we know, the intercept of the security market line is equal to r_t, which is constant in our model. So, we could expect the security market line has higher intercept. After we increase the interest rate, the efficient frontier changes by a little bit. For instance, as show in figure 9, the intercept moves from 0.02 to 0.05. And we observe the slope of market security line decrease, which means the sharpe ratio decreases when the interest rate increases. The sharpe ratio measures the how much excess return can investor get while facing to each unit of volatility. So, in this case, when the interest rate increase, the excess return on numerator decreases more than the decrease in the volatility on denominator.

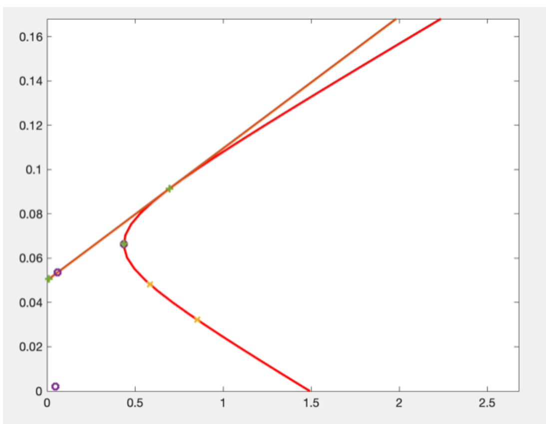

Figure 9. 2011 Efficient Frontier r = 5%

6. Appendix

Matlab Code

```
                r=0.02;
% 2011/9/1
mu=[0.048210619;0.032226991];
cov_mat=[
        0.341492774    0.475568947;
        0.475568947    0.728286891];
% 2012/9/1
mu=[0.0835324;0.079672744];
cov_mat=[
        0.281237793    0.218037671;
        0.218037671    0.194709877];
% 2013/9/1
mu=[0.25502433;0.2754475];
cov_mat=[
        0.086433883    0.094714478;
        0.094714478    0.141158828];
% 2014/9/1
mu=[0.146699072;0.04455271];
cov_mat=[
        0.12342784    0.166395955;
        0.166395955    0.298692731];
% 2015/9/1
mu=[0.067413809;0.02997711];
cov_mat=[
        0.281494505    0.199186658;
```

```
        0.199186658    0.217187185];
% 2016/9/1
mu=[0.006437465;0.005098549];
cov_mat=[
        0.247075076    0.25586398;
        0.25586398    0.32257691];
% 2017/9/1
mu=[0.239660855;0.212029668];
cov_mat=[
        0.024562833    -0.008764038;
        -0.008764038    0.168054768];
% 2018/9/1
mu=[0.062453366;0.014107361];
cov_mat=[
        0.286033181    0.250714767;
        0.250714767    0.300693559];
% 2019/9/1
mu=[0.106652424;0.013212058];
cov_mat=[
        0.466082735    0.280515974;
        0.280515974    0.271292831];
% 2020/9/1
mu=[0.254490861;0.035460168];
cov_mat=[
        0.751064152    0.877639079;
        0.877639079    1.294801571];

A=5;
```

```
points_graph=200;
mu_max=1;
[mv_port, port_tangency, m_port_rf, m_port_norf, opt_port_norf,
    opt_port_rf, mu_port_vec, sigma_port_vec]=mean_variance(r,
    cov_mat, mu, A, points_graph, mu_max)


gamma = 2;
S0 = 10; %b S_0^gamma-1=0.2
b = 0.2/S0.^(gamma-1);
r = 0.02;
R=1.5;
rho=1-1/R;
T = 1; %investment horizon
N = 2^6; %#discretization points
h = T/(N+1); %discretization step
M = 500000   ;
a=0.072058
logS = log(S0).*ones(M,1); %log stock price
logZ= zeros(M,1); %log density process
sigma0=b.*S0.^(gamma-1); % initial vol.
D_logS=sigma0.*ones(M,1); % initial Malliavin derivative
theta0 = a./sigma0; % initial market price of risk (MPR)
D_logZ=-theta0.*ones(M,1); % inital Malliavin derivative
    log density
cum_sum_xirho_D_logxi=zeros(M,1); % initial value of int_t^
    T xi_t,v^rho D_t log xi_t,v dv
cum_sum_xirho=zeros(M,1); % initial value of int_t^T xi_t,v
```

```
^rho dv
%drift
muS=a;
%Simulation
    for i = 2:N
      dW=sqrt(h).*randn(M,1); %BM increment
      dW=dW-mean(dW); %Stabilizing martingale difference transform
      xirho=exp(-r.*(i-1).*rho.*h).*exp(rho.*logZ); %xi_t,v^rho
      cum_sum_xirho=cum_sum_xirho+xirho.*h; %int_t^T xi_t,v}^rho dv
      cum_sum_xirho_D_logxi=cum_sum_xirho_D_logxi+ xirho.*D_logZ.*h; % int_t^T xi_t,v^rho D_t log xi_t,v dv
      sigmaS=b.*exp(logS).^(gamma-1); % stock volatility
      D_sigmaS=(gamma-1).*sigmaS.*D_logS; %MD volatility
      theta = a./sigmaS; %MPR
      D_theta=-D_sigmaS.*a./(sigmaS.*sigmaS); %MD MPR
      logS = logS+(muS-0.5.*sigmaS.*sigmaS).*h+sigmaS.*dW; %log stock price
      logZ = logZ-0.5*theta.*theta.*h - theta.*dW; % log density
      D_logS = D_logS -(sigmaS.*h-dW).*D_sigmaS; %MD log stock price
      D_logZ=D_logZ-(dW+theta.*h).*D_theta;  %MD log density
    end
   %Density
```

```
    xirho=exp(-r.*rho.*T).*exp(rho.*logZ);
    S=exp(logS);
    %Mean variance portfolio weight
    pim=theta0./(R.*sigma0);
    %Hedging demand bequest motive (no intermediate
        consumption)
    pib_noc=rho.*mean(xirho.*D_logZ)./mean(xirho)./sigma0;
    pi_noc=pim+pib_noc;
```

Python Code

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
from sklearn.linear_model import LinearRegression
import pandas as pd
import numpy as np

data = pd.read_excel("regression data.xlsx")
y1 = data["NASDAQ Annually excess"][1:]
y2 = data["Russel Annually excess"][1:]
x1 = data["change in NPOHUS"][1:]
x2 = data["change in CPI"][1:]

df = pd.DataFrame([x1,x2])
X = pd.DataFrame(df.values.T, index=df.columns, columns=df.
    index)
model = LinearRegression()
model = model.fit(X,y1)
beta2,beta3 = model.coef_
```

```
beta1 = model.intercept_
pred = pd.read_excel("asset allocation.xlsx")
y1_pred = beta1+pred['x1']*beta2+pred['x2']*beta3

model = LinearRegression()
model = model.fit(X,y2)
beta2,beta3 = model.coef_
beta1 = model.intercept_
y2_pred = beta1+pred['x1']*beta2+pred['x2']*beta3

df = pd.DataFrame([y1_pred,y2_pred])
pr = pd.DataFrame(df.values.T, index=df.columns, columns=df
    .index)
```