Question 6

(a) From the question, we need to get minimum variance from the data, and then use δ_{BS} and $VEGA_{BS}$ to predict a linear model which is shown in part(a).

There are many unknown that we don't really know, the strike price K, the interest free rate r, the real volatility σ. Since fro the question, we need to get the implied volatility first and then plot into equation A4 from the paper.

Therefore, we decide to use
```
blsprice(STrain,K(i),0.003,TauTrain,0.25)
```

which we choose the risk free rate r = 0.003 and volatility $\sigma = 0.25$. We choose option price = call + put. After get the option price, we use the

```
blsimpv(STrain,K(i),0.003,TauTrain,option)
```

to get the implied volatility value. From A4 we know that we need to have the expected value of change of implied volatility. Based on the volatility smile, we choose a range of strike price K based on the value I see from the raw data
```
K = linspace(700,1200,50);
```
And get expected implied volatility from these 50 different strike price. And then minus each other to get the change of volatility. At the end plot into the equation (A4) to train the model.

lm =

Linear regression model:
y ~ 1 + x1^2 + x2^2 + x1*x2

Estimated Coefficients:

	Estimate	SE	tStat	pValue
(Intercept)	0.026692	0.0044286	6.0271	1.6792e-09
x1	0.32899	0.02466	13.341	1.5349e-40
x1^2	-0.33913	0.024708	-13.725	8.3726e-43
x2	-0.00069901	3.7294e-05	-18.743	3.8152e-78
x2^2	1.0662e-06	5.5286e-08	19.285	1.3332e-82
x1:x2	-0.00019383	5.3406e-05	-3.6293	0.00028442

Number of observations: 56402, Error degrees of freedom: 56396
Root Mean Squared Error: 0.265
R-squared: 0.0106, Adjusted R-Squared: 0.0105
F-statistic vs. constant model: 120, p-value = 3.83e-127

```
% -------------------------------------------------------------------
%  Generated by MATLAB on 2-Mar-2020 09:13:27
%  MATLAB version: 9.6.0.1150989 (R2019a) Update 4
% -------------------------------------------------------------------
load RawData

K = linspace(700,1200,50);    %initialize the strike price K
CVtrain_pose = CVTrain';         %tranpose the matrix
%initialization
imp_vol = zeros(length(DeltaTrain),50);
Option_price = zeros(length(DeltaTrain),1);

%used to calculate the option price and then get the implied volatility for
%50 different strike price
for i = 1: 50
    [call,put] = blsprice(STrain,K(i),0.003,TauTrain,0.25);
    option = call + put;
    del_imp_vol(:,i) = blsimpv(STrain,K(i),0.003,TauTrain,option);
end

% get the mean of 50 implied volatility values
expected_vol = zeros(length(DeltaTrain),1);
for j = 1 : length(DeltaTrain)
    expected_vol(j) = mean(del_imp_vol(j,:));
end

%get the minimum variance
delta_vol = zeros(length(DeltaTrain),1);
delta_vol(1) = 0;
delta_vol(2:end) = expected_vol(2:end) - expected_vol(1:end-1);

%fit to the model and get the coefficient values
mimum_variance = DeltaTrain + VegaTrain.* delta_vol./CVtrain_pose;
T = ...
  [0 0 0;
   1 0 0;
   2 0 0;
   0 1 0;
   0 2 0;
   1 1 0];

x1 = DeltaTrain;
x2 = VegaTrain;
X = [x1,x2];
lm = fitlm(X,delta_vol,T);
```

```
%%(b)
% use the model generated from the part (a) to predict and calculate the
% mean, std, var and cvar
delta = 0.026692 + 0.32899.*x1 -0.33913.*x1.^2 -0.00069901.*x2 +1.0662e-
6.*x2.^2-...
    0.00019383.*x1.*x2;
delta = delta';
delta_V = CVTrain;
delta_S = CSTrain;

% calculate the portfolio error (P&L for real data)
error = delta_V - delta.*delta_S;
histogram(error,50)
mean_1 = mean(error)
sd_1 = std(error)
[var_1,cvar_1] = dVaRCVaR(error,0.95)

% calculate the portfolio error (P&L for predicted data)
delta_2 = DeltaTrain';
error2 = delta_V - delta_2.*delta_S;
mean_2 = mean(error2)
sd_2 = std(error2)
[var_2,cvar_2] = dVaRCVaR(error2,0.95)

%summarize into a table
mea = [mean_1;mean_2];
sd = [sd_1;sd_2];
var = [var_1;var_2];
cvar = [cvar_1;cvar_2];
Table = table(mea,sd,var,cvar);
Table.Properties.VariableNames = {'mean','sd','VaR','CVaR'};
Table.Properties.RowNames = {'estimated','real'};
```

(b)

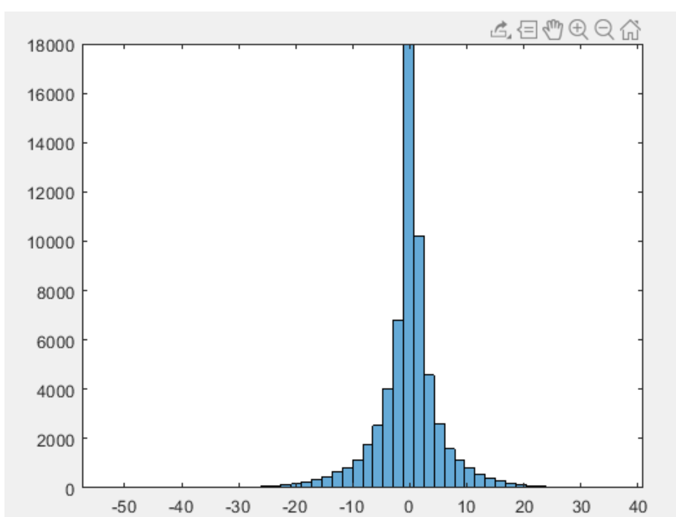

Figure. Histogram of the profit and loss by predicted minimum variance

	mean	sd	VaR	CVaR
estimated	-0.2221	6.0017	-10.294	-16.062
real	0.17121	1.6783	-2.4483	-4.2003

Table: Summary of mean, sd, VaR and CVaR for predicted and real minimum variance

As we can see from the histogram, it relative performs reasonable because most of the portfolio has a P&L = 0. However, from the table, it relatively has a larger sd, and it predicts the VaR and CVaR has a more negative value, usually 4~5 times bigger than the real value, which means it actually hedging more portfolios to have an extreme value.