

MFE 230E

Empirical Methods

Assignment 7

Nicolas Corthorn
Robert Sharow
Xian Li
Sherry Xiu

- (a) From our three measures of volatilities, we can see that the annualized volatility with daily returns is essentially the same as the annualized volatility from monthly volatilities calculated with daily returns. Although our results are slightly different, this is due to the fact that there were 13 leap years and the formula to annualize daily volatilities (multiplying by $\sqrt{252}$) does not account for the extra leap year days. The annualized monthly volatilities from monthly returns is slightly different due to the serial correlation from daily returns.

We also plotted annual volatilities calculated with the three methods to see how the volatilities of the three assets changed over time.

	Volatility	SP500	GE	IBM
Annualized Daily		0.16818	0.28592	0.25863
Annualized Monthly		0.15268	0.27312	0.24681
Annualized Monthly (Daily Data)		0.16827	0.28607	0.25877

Table 1: Annualized Volatilities

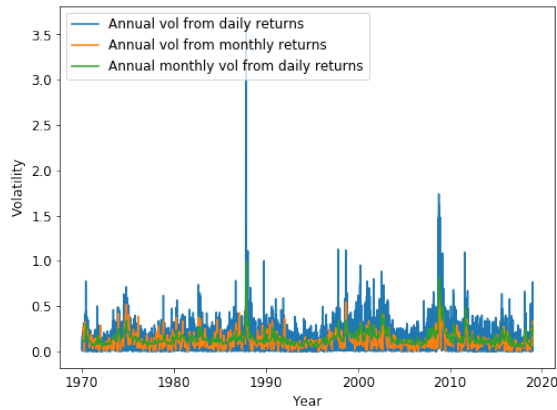


Figure 1: Plot of S&P 500 Volatilities

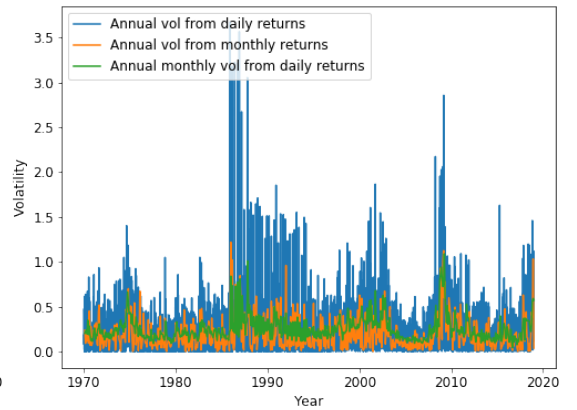


Figure 2: Plot of GE Volatilities

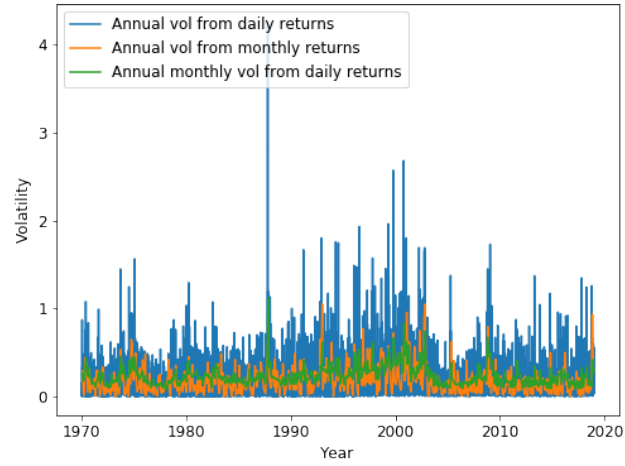


Figure 3: Plot of IBM Volatilities

- (b) We can see that the constants for all the models are close to zero and are small compared to the ϕ coefficient. This means that volatility at time t is affected quite a bit by the volatility at $t-1$, meaning volatility is persistent. This property is also visible from the graphs from part (a) which is independent of the measurement methodology of volatility we choose (calculating annual volatility with daily data, monthly data, or annualizing monthly volatilities calculated with daily data) and applies to S&P 500, GE, and IBM.

OLS Regression Results						
Dep. Variable:	y	R-squared:	0.392			
Model:	OLS	Adj. R-squared:	0.391			
Method:	Least Squares	F-statistic:	377.0			
Date:	Wed, 15 May 2019	Prob (F-statistic):	3.44e-65			
Time:	20:02:41	Log-Likelihood:	1493.4			
No. Observations:	587	AIC:	-2983.			
Df Residuals:	585	BIC:	-2974.			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	0.0157	0.002	10.029	0.000	0.013	0.019
x1	0.6274	0.032	19.416	0.000	0.564	0.691
Omnibus:	609.702	Durbin-Watson:	2.226			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	59397.335			
Skew:	4.444	Prob(JB):	0.00			
Kurtosis:	51.472	Cond. No.	41.2			

Figure 4: AR(1) estimation for S&P 500

OLS Regression Results						
=====						
Dep. Variable:	y	R-squared:	0.308			
Model:	OLS	Adj. R-squared:	0.307			
Method:	Least Squares	F-statistic:	260.4			
Date:	Wed, 15 May 2019	Prob (F-statistic):	1.01e-48			
Time:	20:02:44	Log-Likelihood:	1168.0			
No. Observations:	587	AIC:	-2332.			
Df Residuals:	585	BIC:	-2323.			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
const	0.0321	0.003	11.295	0.000	0.027	0.038
x1	0.5575	0.035	16.136	0.000	0.490	0.625
=====						
Omnibus:	268.443	Durbin-Watson:	2.351			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1855.617			
Skew:	1.897	Prob(JB):	0.00			
Kurtosis:	10.841	Cond. No.	25.4			
=====						

Figure 5: AR(1) estimation for GE

OLS Regression Results						
=====						
Dep. Variable:	y	R-squared:	0.278			
Model:	OLS	Adj. R-squared:	0.277			
Method:	Least Squares	F-statistic:	225.0			
Date:	Wed, 15 May 2019	Prob (F-statistic):	2.84e-43			
Time:	20:02:46	Log-Likelihood:	1270.8			
No. Observations:	587	AIC:	-2538.			
Df Residuals:	585	BIC:	-2529.			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	0.0318	0.003	12.104	0.000	0.027	0.037
x1	0.5272	0.035	15.001	0.000	0.458	0.596
=====						
Omnibus:	354.723	Durbin-Watson:	2.252			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	5329.383			
Skew:	2.377	Prob(JB):	0.00			
Kurtosis:	16.975	Cond. No.	30.8			

Figure 6: AR(1) estimation for IBM

- (c) We can see from the plots below that overall the AR(1) model can predict the volatility for S&P 500, GE, and IBM pretty well. The only concern is that the predicted values are overall smoother than the actual volatilities.

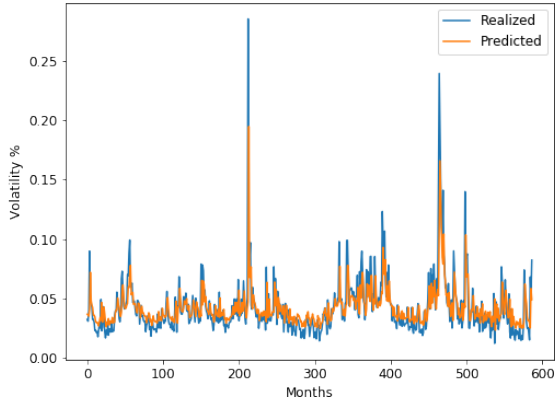


Figure 7: Realized vs Predicted volatility for S&P 500

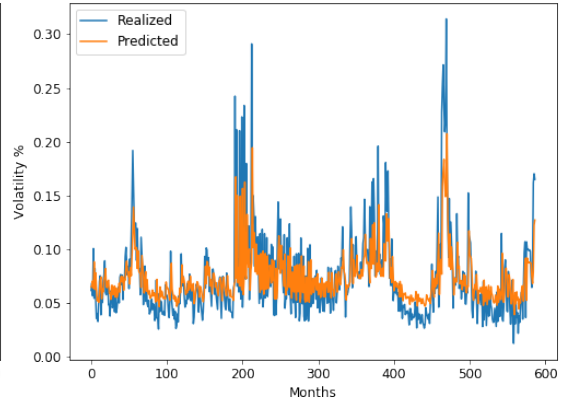


Figure 8: Realized vs Predicted volatility for GE

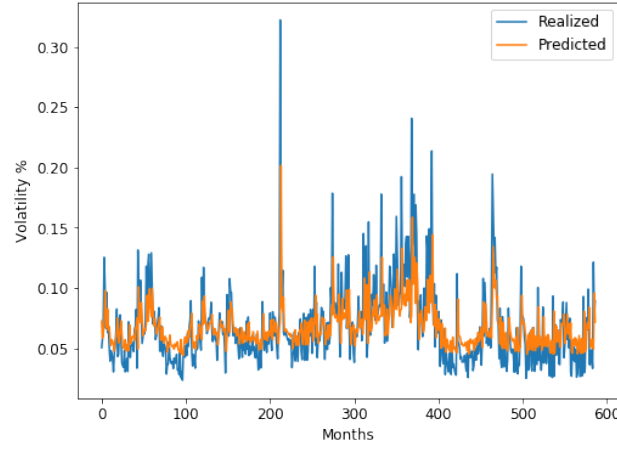


Figure 9: Realized vs Predicted volatility for IBM

(d) For Mean Square Error we have the following

MSE for S&P 500 is 0.00036.

MSE for GE is 0.00109.

MSE for IBM is 0.00077.

Another more intuitive performance measure is RMSE. We have

RMSE for S&P 500 is 0.019.

RMSE for GE is 0.033.

RMSE for IBM is 0.028.

Hence, by comparing MSE, the AR(1) model makes the best prediction for S&P 500 volatility.

It can also be noted from the RMSE, that the error is low in absolute terms. For example, for S&P 500 we have a RMSE of 1.9% for a notional of around 16% which is low.

2. (a) Now we have the following GARCH model

$$\begin{aligned} r &= \mu_t + \epsilon_t \\ \epsilon_t &= \sigma_t e_t \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ e_t &\sim N(0, 1) \end{aligned}$$

For S&P 500, we have $\mu = 6.46e - 03$, $\omega = 1.91e - 04$, $\alpha = 0.10$, and $\beta = 0.80$.
For GE, we have $\mu = 0.01$, $\omega = 2.40e - 04$, $\alpha = 0.14$, and $\beta = 0.83$.
For IBM, we have $\mu = 7.12e - 03$, $\omega = 4.09e - 04$, $\alpha = 0.09$, and $\beta = 0.83$.

Given standard error from the calculations, we know that α 's and β 's are significant, especially the β 's. Notice for all 3 stocks. β 's as big as 0.8 mean that there is a lot of persistence in volatility: when we observe a large volatility, we are very likely to observe another one following it.

(b) We can see that although the conditional volatility from GARCH cannot catch the extreme values, it can predict the persistent volatility after the spikes very well.

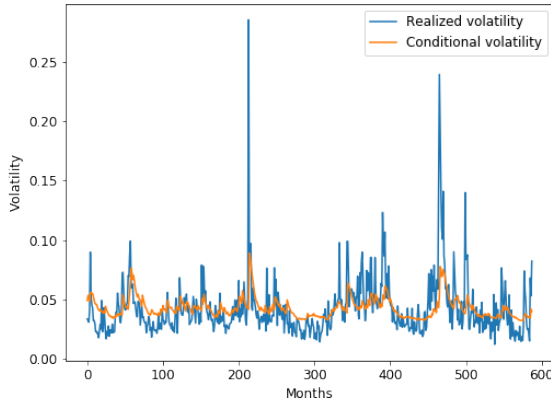


Figure 10: Realized vs Conditional volatility for S&P 500

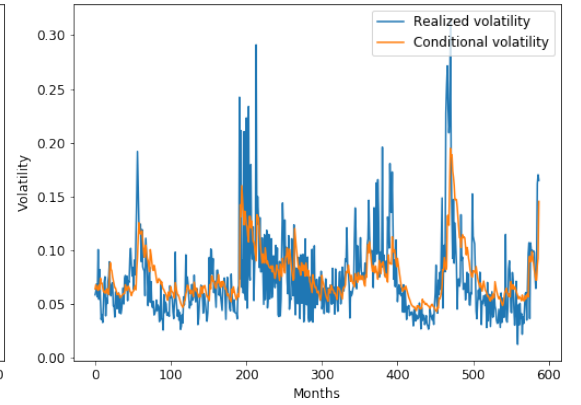


Figure 11: Realized vs Conditional volatility for GE

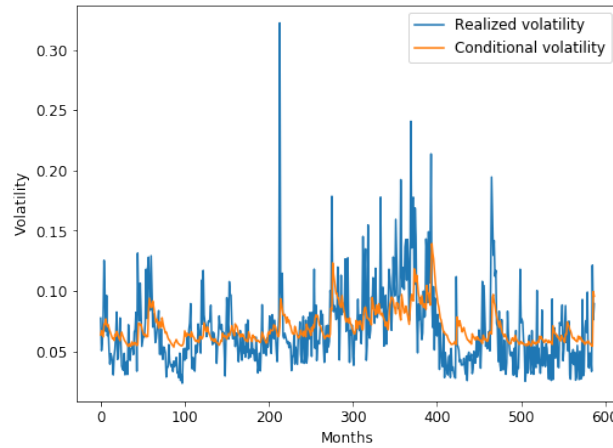


Figure 12: Realized vs Conditional volatility for IBM

```

Regression
                                OLS Regression Results
=====
Dep. Variable:                  y      R-squared:                  0.250
Model:                        OLS      Adj. R-squared:             0.249
Method:                      Least Squares      F-statistic:                195.5
Date:                        Wed, 15 May 2019      Prob (F-statistic):         1.53e-38
Time:                        20:02:46      Log-Likelihood:             1434.8
No. Observations:              588      AIC:                        -2866.
Df Residuals:                  586      BIC:                        -2857.
Df Model:                      1
Covariance Type:              nonrobust
=====
                                coef      std err          t      P>|t|      [0.025      0.975]
-----
const                -0.0189         0.004     -4.256     0.000     -0.028     -0.010
x1                   1.4221         0.102    13.982     0.000         1.222         1.622
=====
Omnibus:                  629.808      Durbin-Watson:              1.170
Prob(Omnibus):             0.000      Jarque-Bera (JB):           46605.888
Skew:                      4.824      Prob(JB):                   0.00
Kurtosis:                  45.535      Cond. No.                   117.
=====

```

Figure 13: Regressing Conditional and Realized Volatilities for S&P 500

```

Regression
                                OLS Regression Results
=====
Dep. Variable:                  y      R-squared:                  0.293
Model:                        OLS      Adj. R-squared:             0.292
Method:                      Least Squares      F-statistic:                243.0
Date:                        Wed, 15 May 2019      Prob (F-statistic):         4.36e-46
Time:                        20:02:47      Log-Likelihood:             1164.2
No. Observations:              588      AIC:                        -2324.
Df Residuals:                  586      BIC:                        -2316.
Df Model:                      1
Covariance Type:              nonrobust
=====
                                coef      std err          t      P>|t|      [0.025      0.975]
-----
const                 0.0040         0.005      0.877     0.381     -0.005     0.013
x1                   0.9064         0.058    15.589     0.000         0.792         1.021
=====
Omnibus:                  250.201      Durbin-Watson:              1.436
Prob(Omnibus):             0.000      Jarque-Bera (JB):           1427.327
Skew:                      1.810      Prob(JB):                   1.15e-310
Kurtosis:                  9.719      Cond. No.                   42.4
=====

```

Figure 14: Regressing Conditional and Realized Volatilities for GE

Regression						
OLS Regression Results						
Dep. Variable:	y	R-squared:	0.258			
Model:	OLS	Adj. R-squared:	0.257			
Method:	Least Squares	F-statistic:	204.2			
Date:	Wed, 15 May 2019	Prob (F-statistic):	5.74e-40			
Time:	20:02:47	Log-Likelihood:	1265.6			
No. Observations:	588	AIC:	-2527.			
Df Residuals:	586	BIC:	-2519.			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-0.0174	0.006	-2.889	0.004	-0.029	-0.006
x1	1.2202	0.085	14.292	0.000	1.053	1.388
Omnibus:	349.718	Durbin-Watson:	1.460			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4941.665			
Skew:	2.348	Prob(JB):	0.00			
Kurtosis:	16.403	Cond. No.	73.9			

Figure 15: Regressing Conditional and Realized Volatilities for IBM

- (c) RMSE for S&P 500 is 0.021.
 RMSE for GE is 0.033.
 RMSE for IBM is 0.028.

Looking at the RMSE, we would say that GARCH performs well in predicting volatility for all 3 assets.

3. (a) Now we have the following GJR-GARCH model

$$\begin{aligned} r &= \mu_t + \epsilon_t \\ \epsilon_t &= \sigma_t e_t \\ \sigma_t^2 &= \omega + (\alpha + \gamma I_{t-1}) \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ e_t &\sim N(0, 1) \end{aligned}$$

For S&P 500, we have $\mu = 5.56e - 03$, $\omega = 1.28e - 04$, $\alpha = 0.056$, $\gamma = 0.1160$, and $\beta = 0.822$.

For GE, we have $\mu = 0.0135$, $\omega = 2.38e - 04$, $\alpha = 0.104$, $\gamma = 0.070$, and $\beta = 0.834$.

For IBM, we have $\mu = 6.60e - 03$, $\omega = 5.10e - 04$, $\alpha = 0.055$, $\gamma = 0.098$, and $\beta = 0.796$.

(b) We can see that although the conditional volatility from GJR-GARCH cannot catch the extreme values, it can predict the persistent volatility spikes very well.

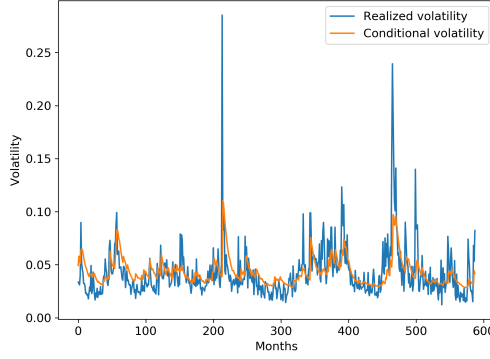


Figure 16: Realized vs Conditional volatility for S&P 500

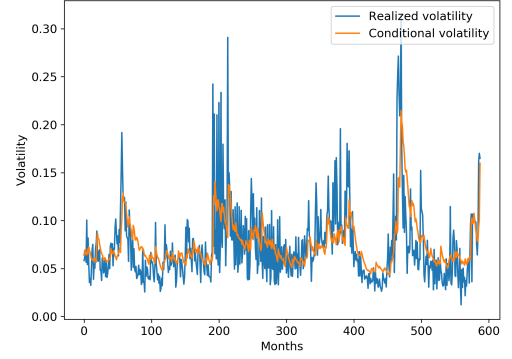


Figure 17: Realized vs Conditional volatility for GE

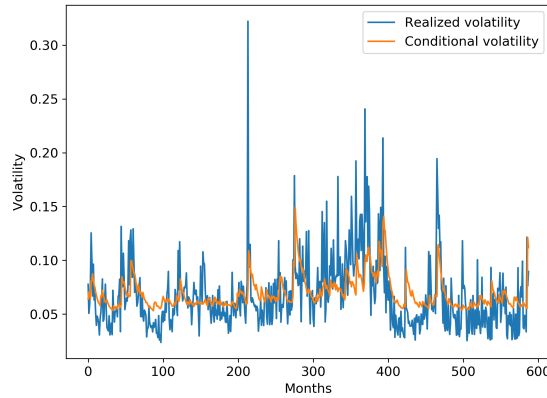


Figure 18: Realized vs Conditional volatility for IBM


```

Regression
      OLS Regression Results
=====
Dep. Variable:          y    R-squared:          0.276
Model:                  OLS    Adj. R-squared:    0.275
Method:                 Least Squares    F-statistic:    223.8
Date:                   Wed, 15 May 2019    Prob (F-statistic):    4.37e-43
Time:                   21:19:31    Log-Likelihood:    1445.2
No. Observations:       588    AIC:            -2886.
Df Residuals:           586    BIC:            -2878.
Df Model:               1
Covariance Type:        nonrobust
=====
               coef    std err          t      P>|t|    [0.025    0.975]
-----
const         -0.0020    0.003       -0.664    0.507    -0.008    0.004
x1             1.0222    0.068     14.959    0.000    0.888    1.156
=====
Omnibus:             644.223    Durbin-Watson:           1.204
Prob(Omnibus):        0.000    Jarque-Bera (JB):       53325.423
Skew:                 4.971    Prob(JB):                0.00
Kurtosis:             48.582    Cond. No.                80.0
=====

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
MSE: 0.02074760329089548
RMSE: 0.144040283569894
MAE: 0.01258171241269164

```

Figure 19: Regressing Conditional and Realized Volatilities for S&P 500

```

Regression
      OLS Regression Results
=====
Dep. Variable:          y    R-squared:          0.327
Model:                  OLS    Adj. R-squared:    0.326
Method:                 Least Squares    F-statistic:    284.5
Date:                   Wed, 15 May 2019    Prob (F-statistic):    2.53e-52
Time:                   21:19:59    Log-Likelihood:    1178.6
No. Observations:       588    AIC:            -2353.
Df Residuals:           586    BIC:            -2344.
Df Model:               1
Covariance Type:        nonrobust
=====
               coef    std err          t      P>|t|    [0.025    0.975]
-----
const          0.0017    0.004       0.390    0.697    -0.007    0.010
x1             0.9399    0.056     16.867    0.000    0.830    1.049
=====
Omnibus:             259.460    Durbin-Watson:           1.482
Prob(Omnibus):        0.000    Jarque-Bera (JB):       1615.945
Skew:                 1.856    Prob(JB):                0.00
Kurtosis:             10.224    Cond. No.                41.6
=====

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
MSE: 0.03275639135296147
RMSE: 0.18098726848306615
MAE: 0.022813610643813383

```

Figure 20: Regressing Conditional and Realized Volatilities for GE

```

Regression
      OLS Regression Results
=====
Dep. Variable:      y      R-squared:      0.262
Model:              OLS      Adj. R-squared: 0.260
Method:             Least Squares      F-statistic: 207.7
Date:               Wed, 15 May 2019      Prob (F-statistic): 1.60e-40
Time:               21:20:09      Log-Likelihood: 1266.9
No. Observations:   588      AIC: -2530.
Df Residuals:       586      BIC: -2521.
Df Model:            1
Covariance Type:    nonrobust
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const         -0.0107      0.006      -1.930      0.054      -0.022      0.000
x1             1.1232      0.078      14.411      0.000      0.970      1.276
=====
Omnibus:          341.978      Durbin-Watson:      1.473
Prob(Omnibus):    0.000      Jarque-Bera (JB): 4617.936
Skew:             2.294      Prob(JB):      0.00
Kurtosis:         15.940      Cond. No.      67.6
=====

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
MSE: 0.02819623125759392
RMSE: 0.16791733459531186
MAE: 0.0201179591229439

```

Figure 21: Regressing Conditional and Realized Volatilities for IBM

- (c) RMSE for S&P 500 is 0.021.
 RMSE for GE is 0.033.
 RMSE for IBM is 0.028.

Looking at the RMSE, we would say that GJR-GARCH also performs well in predicting volatility for all 3 assets.

4. Table of RMSE and MAE of the AR(1), GARCH, and GJR-GARCH:

Error Type	SP500	GE	IBM
AR(1) RMSE	0.019	0.033	0.028
GARCH RMSE	0.021	0.033	0.028
GJR-GARCH RMSE	0.021	0.033	0.028
AR(1) MAE	0.011	0.022	0.019
GARCH MAE	0.013	0.023	0.020
GJR-GARCH MAE	0.013	0.023	0.020

Table 2: Comparing RMSE and MAE of the models

From the table above, we can see that the results for all 3 models are very similar. The AR(1) model tends to produce sometimes slightly smaller RMSE and MAE. AR(1) and GARCH seem like reasonable options as they are the most parsimonious (being AR(1) the simplest). GARCH gives the flexibility of also modeling mean returns as a time series while accounting for persistence in volatility. However, for this particular setting, we could choose any model as results would be very similar in practice.

Assignment 7

May 16, 2019

```
In [1]: import numpy as np
import statsmodels.api as sm
import pandas as pd
from datetime import datetime
import copy
from statsmodels.stats.sandwich_covariance import cov_hac as cov
import matplotlib.pyplot as plt
from scipy.stats import chi2
from arch import arch_model
from pandas.plotting import register_matplotlib_converters
register_matplotlib_converters()

In [2]: def OLS(x, y, addcon=True, cov_type=None, sig_level=.05, summary=0, cov_kwds = None):
    """Wrapper for statsmodels OLS regression"""
    if addcon:
        X = sm.add_constant(x)
    else:
        X = x
    if cov_type==None:
        ols_results = sm.OLS(y,X).fit(cov_type='nonrobust')
    else:
        ols_results = sm.OLS(y,X).fit(cov_type=cov_type, cov_kwds=cov_kwds)

    ### print out the OLS estimation results
    if summary==1:
        print(ols_results.summary())

    ols_cov_mat = cov(ols_results)
    ols_beta_hat = ols_results.params # beta_hat
    ols_resids = ols_results.resid # resids
    ols_se = ols_results.bse
    ols_pvalues = ols_results.pvalues

    return ols_beta_hat, ols_resids, ols_se, ols_cov_mat, ols_pvalues

def latex_table(df, caption="", label="", index=False):
```

```

return "\\begin{table}[H]\\n\\centering\\n"+df.to_latex(index=index)+"\\caption{"+caption+
\\n\\end{table}"

```

```

In [3]: def question_1_wrapper(data_name, print_OLS):
    '''
    data_name can be either 'SP', 'GE' or 'IBM'
    '''

    # Data
    data = pd.read_csv(data_name+'.csv')
    data['Date'] = pd.to_datetime(data['Date']).dt.date
    data = data.set_index('Date')
    data = data.loc[data.index.astype(str) <= '2018-12-31'] # Every year and month are c

    # Daily continuous returns (not scaled)
    Daily_Returns = []
    for i in range(len(data['Adj Close'])-1):
        Daily_Returns.append(np.log(data['Adj Close'][i+1]/data['Adj Close'][i]))
    data = data.iloc[1:]
    data['Daily Returns'] = Daily_Returns
    data = data.reset_index()

    # Annualized Volatility with Daily Data (Monthly)
    RV_annual_from_daily_monthly = []
    RV_List = [] # Daily returns in the month
    cur_month = data['Date'][0].month
    for i in range(len(data['Date'])):
        if data['Date'][i].month == cur_month and i < (len(data['Date'])-1):
            RV_List.append(data['Daily Returns'][i])
        else:
            if i == (len(data['Date']) - 1):
                RV_List.append(data['Daily Returns'][i])
            RV_annual_from_daily_monthly.append(np.sqrt(np.sum((np.array(RV_List))**2.0)))
            RV_List = []
            cur_month = data['Date'][i].month
            RV_List.append(data['Daily Returns'][i])
    RV_annual_from_daily_monthly_var = np.array(RV_annual_from_daily_monthly)**2
    RV_annual_from_daily_monthly = np.array(RV_annual_from_daily_monthly)
    RV_annual_from_daily_monthly_annualized = RV_annual_from_daily_monthly * np.sqrt(12)

    # Realized monthly volatility from daily data
    Monthly_Returns = []
    RV_monthly_from_daily = []
    RV_List = [] # Daily returns in the month
    cur_month = data['Date'][0].month
    Month_List = []
    for i in range(len(data['Date'])):
        if data['Date'][i].month == cur_month and i < (len(data['Date'])-1):

```

```

RV_List.append(data['Daily Returns'][i])
else:
    if i == (len(data['Date']) - 1):
        RV_List.append(data['Daily Returns'][i])
        return_for_the_month = np.sum(np.array(RV_List))
        Monthly_Returns.append(return_for_the_month)
        RV_monthly_from_daily.append(np.sqrt(return_for_the_month**2.0))
        RV_List = []
        Month_List.append(data['Date'][i])
        cur_month = data['Date'][i].month
        RV_List.append(data['Daily Returns'][i])
RV_monthly_from_daily_var = np.array(RV_monthly_from_daily)**2
RV_monthly_from_daily = np.array(RV_monthly_from_daily) * np.sqrt(12)

# Realized volatility with daily returns
RV_annual_from_daily_var = np.array((data['Daily Returns'])**2)
RV_annual_from_daily = np.sqrt(np.array((data['Daily Returns'])**2)) * np.sqrt(252)

# Output
Daily_Returns = np.array(Daily_Returns)
Monthly_Returns = np.array(Monthly_Returns)
RV_annual_from_daily = np.array(RV_annual_from_daily)
RV_annual_from_monthly = np.array(RV_monthly_from_daily)
RV_annual_from_monthly_daily = np.array(RV_annual_from_daily_monthly)
vol = RV_annual_from_monthly_daily

# Comparisson
plt.figure(figsize=(8,6))
plt.rcParams.update({'font.size': 12})
plt.plot(data['Date'], RV_annual_from_daily, label='Annual vol from daily returns')
plt.plot(Month_List, RV_annual_from_monthly, label='Annual vol from monthly returns')
plt.plot(Month_List, RV_annual_from_daily_monthly_annualized, label='Annual monthly')
plt.xlabel('Year')
plt.ylabel('Volatility')
plt.legend(loc='upper left')
plt.show()

# Fit AR(1)
x = vol[0:len(vol)-1]
y = vol[1:len(vol)]
ols_beta_hat, ols_resids, ols_se, ols_cov_mat, ols_pvalues = OLS(x, y, addcon = True)

# Forecast
y_hat = x*ols_beta_hat[1] + ols_beta_hat[0]
plt.figure(figsize=(8,6))
plt.rcParams.update({'font.size': 12})
plt.plot(y, label='Realized')
plt.plot(y_hat, label='Predicted')

```

```

plt.legend()
plt.xlabel('Months')
plt.ylabel('Volatility %')
plt.show()

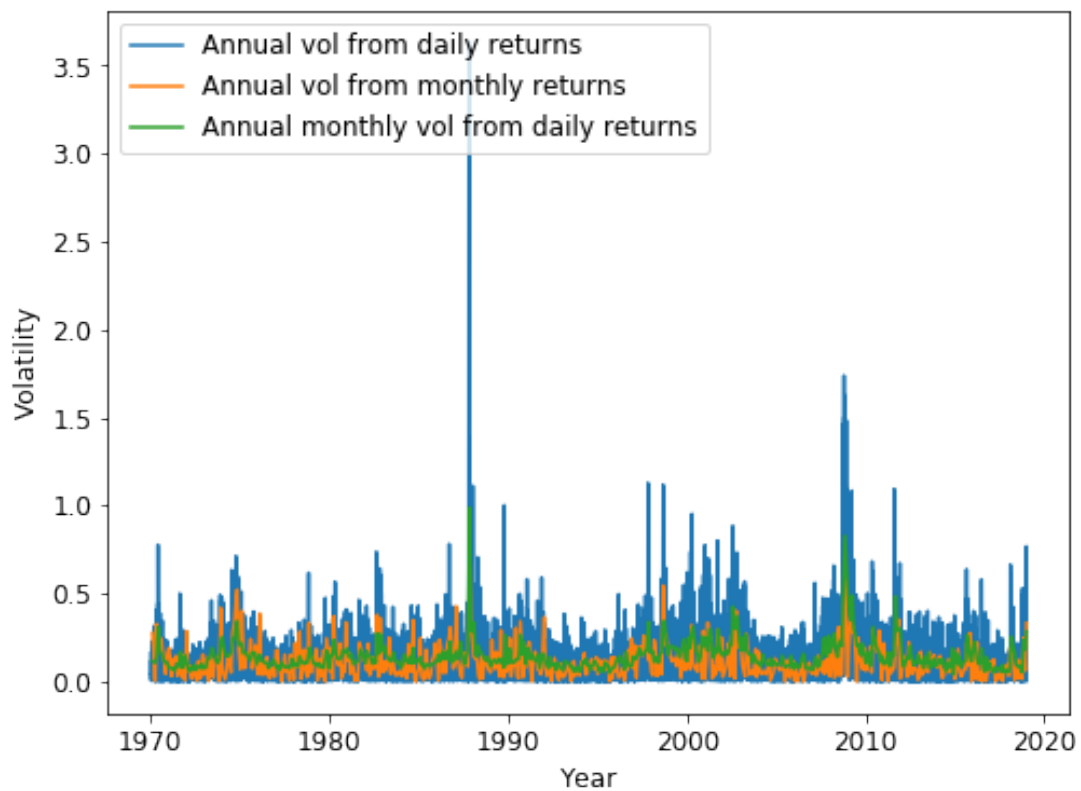
# Error
mse = np.mean((y - y_hat)**2)
rmse = np.sqrt(mse)
mae = np.mean(np.abs(y-y_hat))
print('MSE:', mse)
print('RMSE:', rmse)
print('MAE:', mae)
print('Annualized Volatility with Daily Returns:', np.sqrt(np.mean(RV_annual_from_da
print('Annualized Volatility with Monthly Returns:', np.sqrt(np.mean(RV_monthly_from
print('Annualized Volatility with Monthly Vol and Daily Data:', np.sqrt(np.mean(RV_a

return Monthly>Returns, RV_annual_from_monthly_daily

```

0.1 Problem 1

In [4]: `SP_returns, SP_vol = question_1_wrapper('SP', 1)`



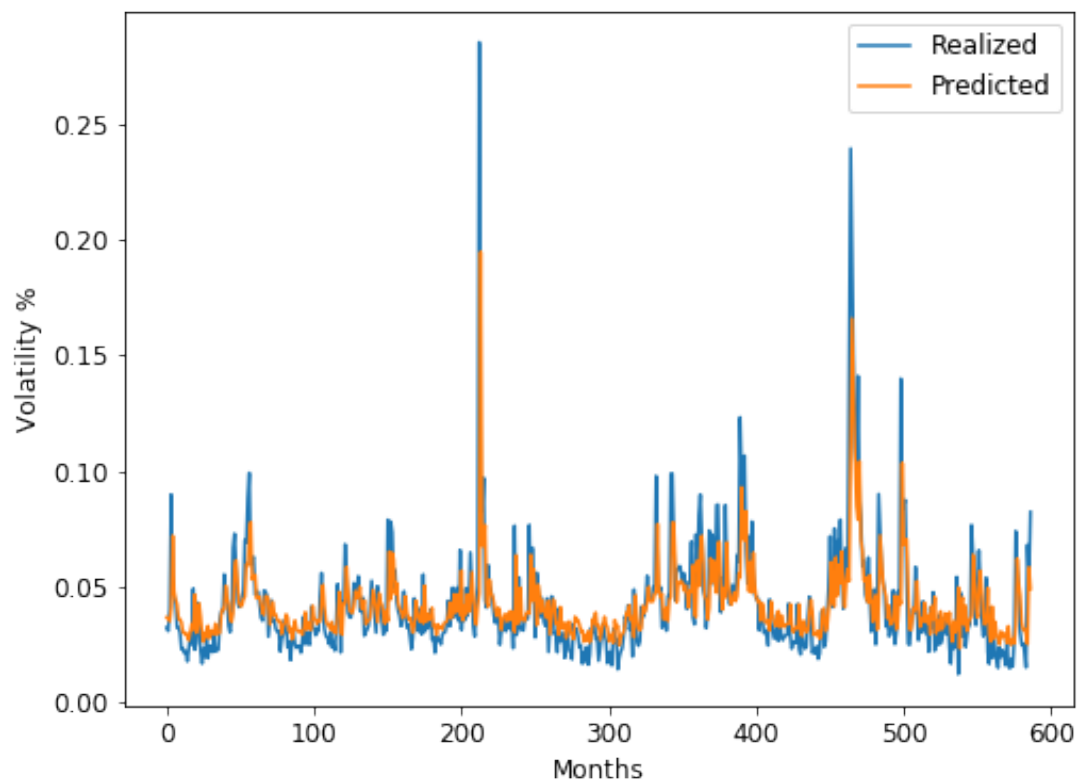
OLS Regression Results

=====						
Dep. Variable:	y	R-squared:	0.392			
Model:	OLS	Adj. R-squared:	0.391			
Method:	Least Squares	F-statistic:	377.0			
Date:	Wed, 15 May 2019	Prob (F-statistic):	3.44e-65			
Time:	22:37:28	Log-Likelihood:	1493.4			
No. Observations:	587	AIC:	-2983.			
Df Residuals:	585	BIC:	-2974.			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	0.0157	0.002	10.029	0.000	0.013	0.019
x1	0.6274	0.032	19.416	0.000	0.564	0.691
=====						
Omnibus:	609.702	Durbin-Watson:	2.226			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	59397.335			
Skew:	4.444	Prob(JB):	0.00			
Kurtosis:	51.472	Cond. No.	41.2			
=====						

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



MSE: 0.00036122005150237944

RMSE: 0.019005789946813036

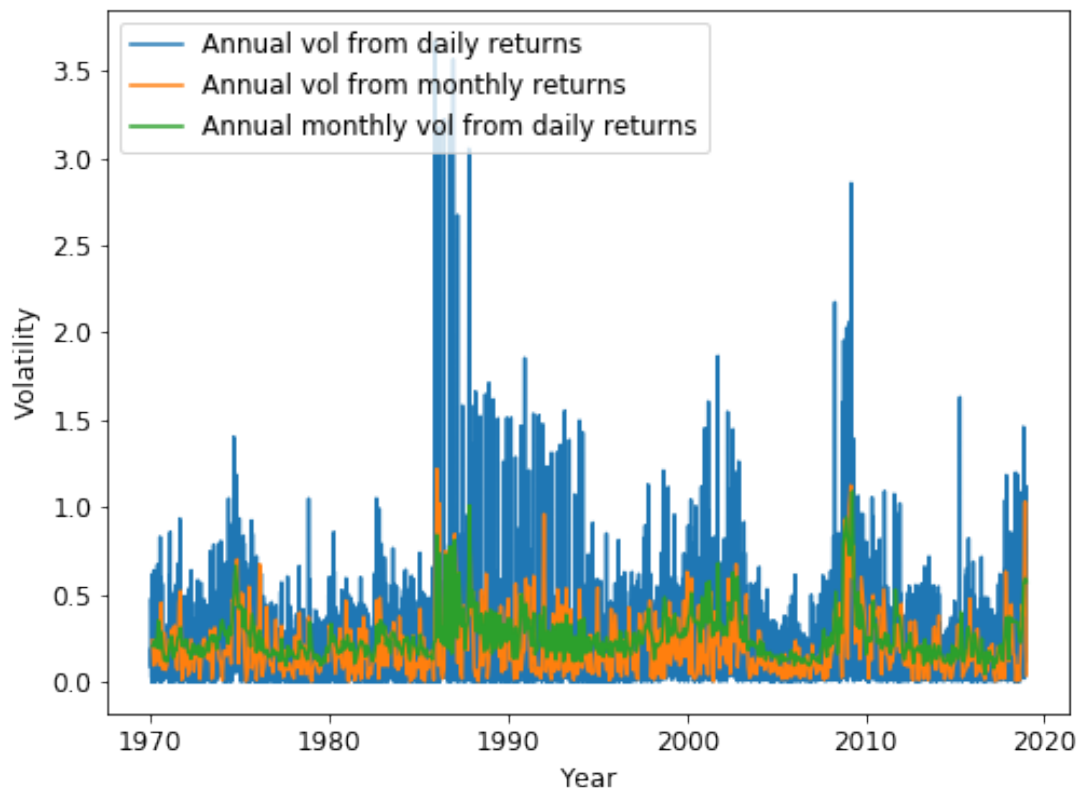
MAE: 0.010968749039436572

Annualized Volatility with Daily Returns: 0.16818403557750747

Annualized Volatility with Monthly Returns: 0.15267557334233467

Annualized Volatility with Monthly Vol and Daily Data: 0.16827254453869964

```
In [5]: GE_returns, GE_vol = question_1_wrapper('GE', 1)
```



OLS Regression Results

```

=====
Dep. Variable:          y      R-squared:          0.308
Model:                  OLS    Adj. R-squared:       0.307
Method:                 Least Squares    F-statistic:       260.4
Date:                  Wed, 15 May 2019    Prob (F-statistic): 1.01e-48
Time:                  22:37:31    Log-Likelihood:    1168.0
No. Observations:      587    AIC:              -2332.
Df Residuals:          585    BIC:              -2323.
Df Model:              1
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	0.0321	0.003	11.295	0.000	0.027	0.038
x1	0.5575	0.035	16.136	0.000	0.490	0.625

```

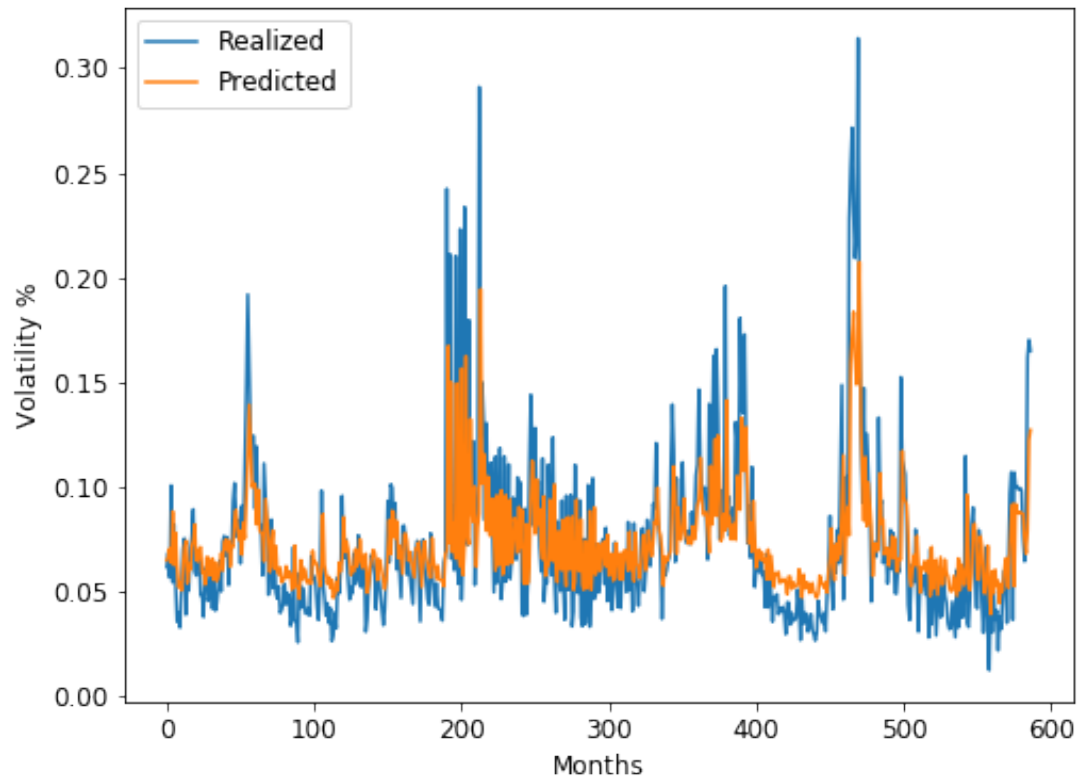
=====
Omnibus:                268.443    Durbin-Watson:       2.351
Prob(Omnibus):          0.000    Jarque-Bera (JB):    1855.617
Skew:                   1.897    Prob(JB):            0.00
Kurtosis:               10.841    Cond. No.            25.4
=====

```

=====

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



MSE: 0.0010943459149488828

RMSE: 0.033080899548665284

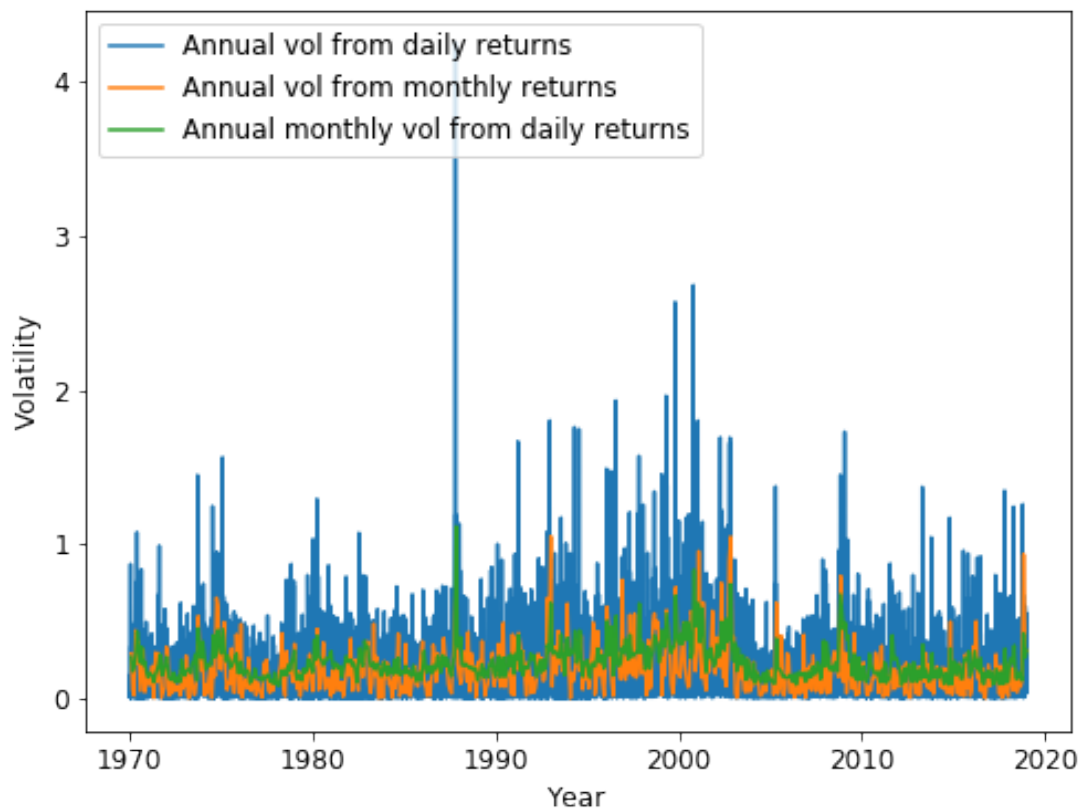
MAE: 0.02199152059292526

Annualized Volatility with Daily Returns: 0.2859189536300444

Annualized Volatility with Monthly Returns: 0.2731223941114715

Annualized Volatility with Monthly Vol and Daily Data: 0.2860694220706669

In [6]: IBM_returns, IBM_vol = question_1_wrapper('IBM', 1)



OLS Regression Results

```

=====
Dep. Variable:          y      R-squared:          0.278
Model:                  OLS    Adj. R-squared:       0.277
Method:                  Least Squares    F-statistic:       225.0
Date:                    Wed, 15 May 2019    Prob (F-statistic): 2.84e-43
Time:                    22:37:33    Log-Likelihood:    1270.8
No. Observations:       587    AIC:              -2538.
Df Residuals:           585    BIC:              -2529.
Df Model:                1
Covariance Type:        nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	0.0318	0.003	12.104	0.000	0.027	0.037
x1	0.5272	0.035	15.001	0.000	0.458	0.596

```

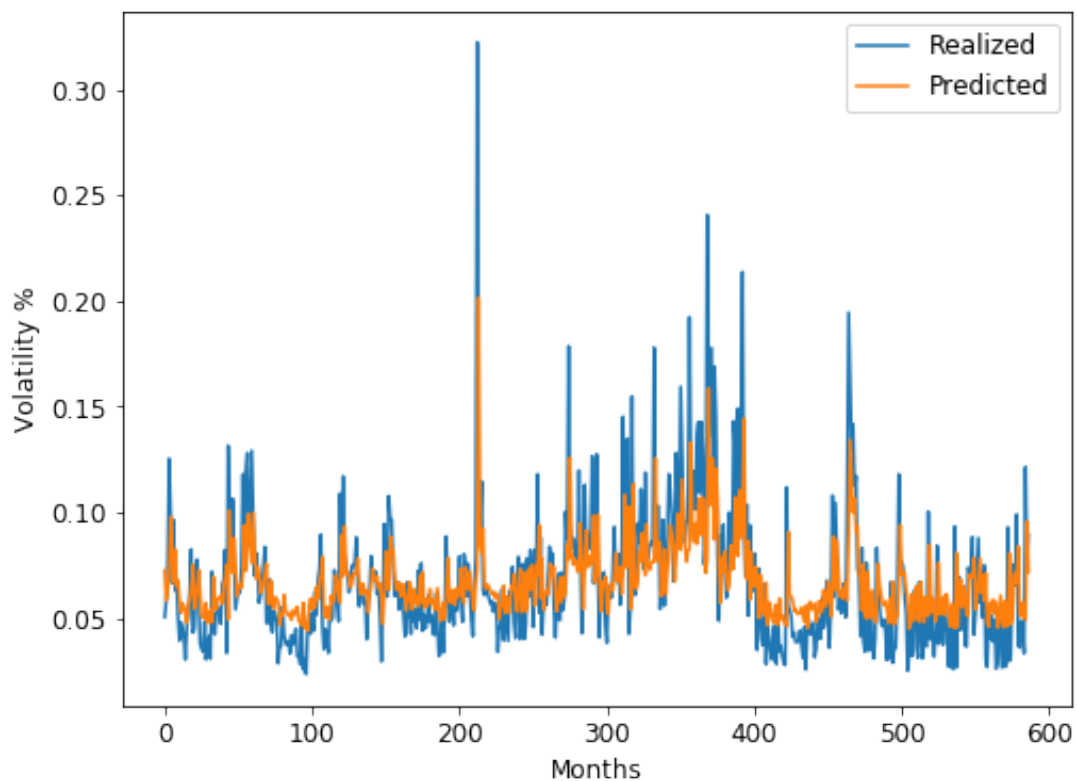
=====
Omnibus:                354.723    Durbin-Watson:       2.252
Prob(Omnibus):          0.000    Jarque-Bera (JB):    5329.383
Skew:                   2.377    Prob(JB):            0.00
=====

```

Kurtosis:	16.975	Cond. No.	30.8
=====			

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



MSE: 0.0007711503591662113

RMSE: 0.027769594148388473

MAE: 0.018829690059106226

Annualized Volatility with Daily Returns: 0.25863424149889247

Annualized Volatility with Monthly Returns: 0.2468062143866231

Annualized Volatility with Monthly Vol and Daily Data: 0.25877035101704726

0.2 Problem 2

```
In [7]: def question_2_3_wrapper(returns, realized_vol, p=1, o=0, q=1):
        """
        p (int, optional) Lag order of the symmetric innovation
        o (int, optional) Lag order of the asymmetric innovation
```

q (int, optional) Lag order of lagged volatility or equivalent

For GARCH: [p, o, q] = [1, 0, 1]

For GJR-GARCH: [p, o, q] = [1, 1, 1]

'''

Plot returns

```
plt.figure(figsize=(8,6))
plt.rcParams.update({'font.size': 12})
plt.plot(returns)
plt.xlabel('Months')
plt.ylabel('Returns')
plt.show()
```

GARCH (part a)

```
GARCH_model = arch_model(returns, p=p, o=o, q=q)
Res = GARCH_model.fit(update_freq = 5)
conditional_vol = Res.conditional_volatility
print('\n\n\n\n\n GARCH')
print(Res)
```

Regression (part b)

```
print('\n\n\n\n\n Regression')
vol_beta, vol_resids, _, _, _ = OLS(conditional_vol, realized_vol, summary = 1)
```

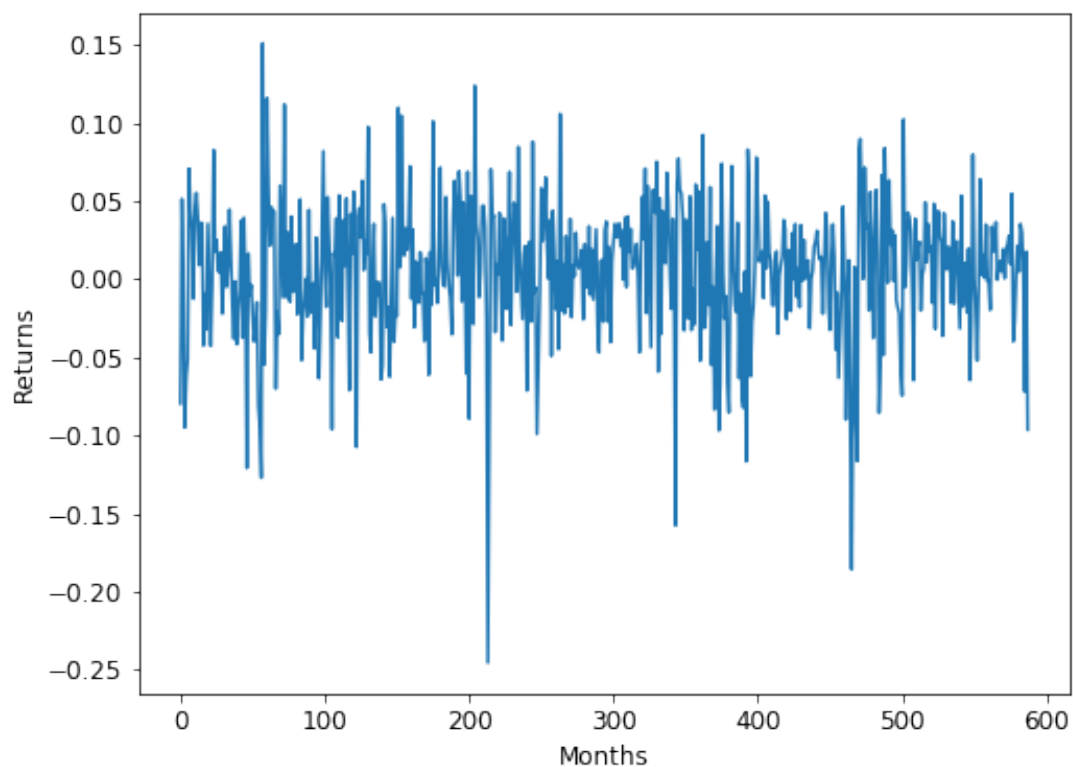
Plot

```
plt.figure(figsize=(8,6))
plt.rcParams.update({'font.size': 12})
plt.plot(realized_vol, label='Realized volatility')
plt.plot(conditional_vol, label='Conditional volatility')
plt.legend()
plt.xlabel('Months')
plt.ylabel('Volatility')
```

Error

```
mse = np.mean((realized_vol - conditional_vol)**2)
rmse = np.sqrt(mse)
mae = np.mean(np.abs(realized_vol-conditional_vol))
print('MSE:', mse)
print('RMSE:', rmse)
print('MAE:', mae)
```

In [8]: question_2_3_wrapper(SP_returns, SP_vol)



```

Positive directional derivative for linesearch      (Exit mode 8)
Current function value: -1028.5391663334103
Iterations: 6
Function evaluations: 18
Gradient evaluations: 2

```

GARCH

Constant Mean - GARCH Model Results

```

=====
Dep. Variable:          y      R-squared:          -0.000
Mean Model:             Constant Mean    Adj. R-squared:      -0.000
Vol Model:              GARCH    Log-Likelihood:    1028.54
Distribution:           Normal    AIC:              -2049.08
Method:                Maximum Likelihood BIC:            -2031.57
                               No. Observations:      588
Date:                  Wed, May 15 2019    Df Residuals:      584
Time:                  22:37:34    Df Model:          4
                               Mean Model

```

```

=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
mu          6.4621e-03  2.013e-03      3.210  1.329e-03  [2.516e-03,1.041e-02]
              Volatility Model
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega       1.9111e-04  1.008e-03      0.190      0.850  [-1.784e-03,2.166e-03]
alpha[1]     0.1000  2.725e-02      3.669  2.433e-04  [4.658e-02, 0.153]
beta[1]      0.8000      0.553      1.446      0.148  [ -0.285, 1.885]
=====

```

Covariance estimator: robust

WARNING: The optimizer did not indicate successful convergence. The message was Positive directional derivative for linesearch. See convergence_flag.

Regression

OLS Regression Results

```

=====
Dep. Variable:          y      R-squared:                0.250
Model:                  OLS      Adj. R-squared:          0.249
Method:                 Least Squares      F-statistic:        195.5
Date:                  Wed, 15 May 2019      Prob (F-statistic):    1.53e-38
Time:                  22:37:34      Log-Likelihood:       1434.8
No. Observations:      588      AIC:                 -2866.
Df Residuals:          586      BIC:                 -2857.
Df Model:              1
Covariance Type:       nonrobust
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const         -0.0189      0.004     -4.256      0.000      -0.028      -0.010
x1             1.4221      0.102     13.982      0.000       1.222       1.622
=====
Omnibus:              629.808      Durbin-Watson:          1.170
Prob(Omnibus):        0.000      Jarque-Bera (JB):       46605.888
Skew:                 4.824      Prob(JB):               0.00
Kurtosis:             45.535      Cond. No.               117.
=====

```

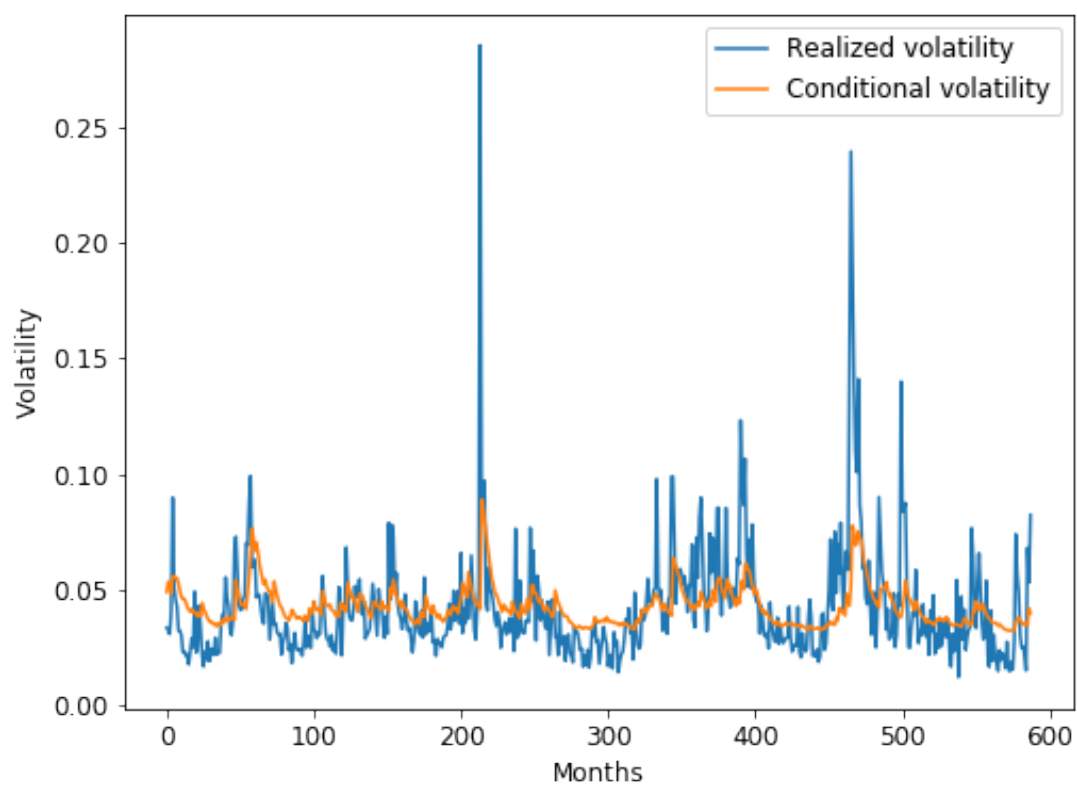
Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

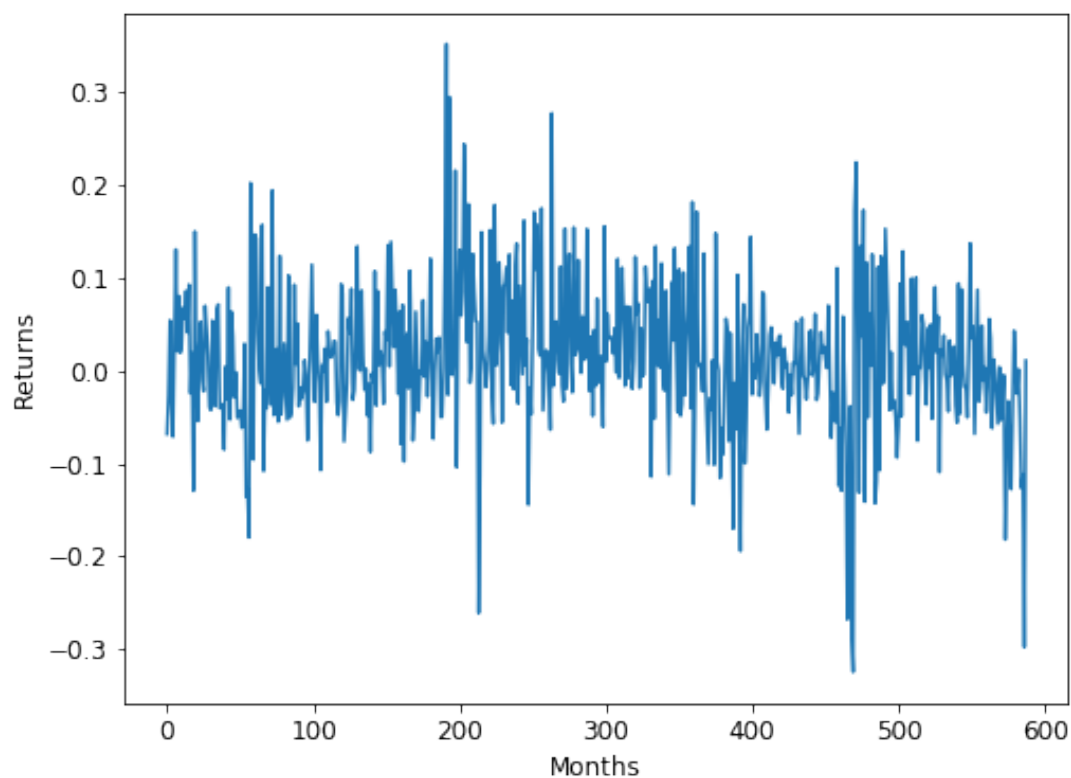
MSE: 0.0004585009094400436
RMSE: 0.021412634341435983
MAE: 0.013028995258359091

C:\Users\conan\Anaconda3\lib\site-packages\arch\univariate\base.py:577: ConvergenceWarning:
The optimizer returned code 8. The message is:
Positive directional derivative for linesearch
See scipy.optimize.fmin_slsqp for code meaning.

ConvergenceWarning)



In [9]: question_2_3_wrapper(GE_returns, GE_vol)



```

Iteration:      5,   Func. Count:    44,   Neg. LLF: -709.5464610335621
Iteration:     10,   Func. Count:    76,   Neg. LLF: -709.5994262569297
Optimization terminated successfully.   (Exit mode 0)
    Current function value: -709.5994262567933
    Iterations: 10
    Function evaluations: 76
    Gradient evaluations: 10

```

GARCH

Constant Mean - GARCH Model Results

```

=====
Dep. Variable:                y      R-squared:                -0.000
Mean Model:      Constant Mean  Adj. R-squared:           -0.000
Vol Model:      GARCH          Log-Likelihood:             709.599
Distribution:    Normal        AIC:                  -1411.20
Method:         Maximum Likelihood  BIC:                  -1393.69
                                           No. Observations:      588

```

Date: Wed, May 15 2019 Df Residuals: 584
Time: 22:37:35 Df Model: 4

Mean Model

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
mu           0.0145   2.779e-03      5.233   1.672e-07   [9.096e-03,1.999e-02]
```

Volatility Model

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega        2.4052e-04  1.141e-04      2.108   3.499e-02   [1.694e-05,4.641e-04]
alpha[1]      0.1412   3.855e-02      3.661   2.509e-04   [6.559e-02, 0.217]
beta[1]       0.8291   3.571e-02     23.216  3.103e-119   [ 0.759, 0.899]
=====
```

Covariance estimator: robust

Regression

OLS Regression Results

```
=====
Dep. Variable:          y      R-squared:          0.293
Model:                  OLS      Adj. R-squared:        0.292
Method:                 Least Squares      F-statistic:        243.0
Date:                   Wed, 15 May 2019      Prob (F-statistic):    4.36e-46
Time:                   22:37:35      Log-Likelihood:       1164.2
No. Observations:       588      AIC:                -2324.
Df Residuals:           586      BIC:                -2316.
Df Model:                1
Covariance Type:        nonrobust
=====
```

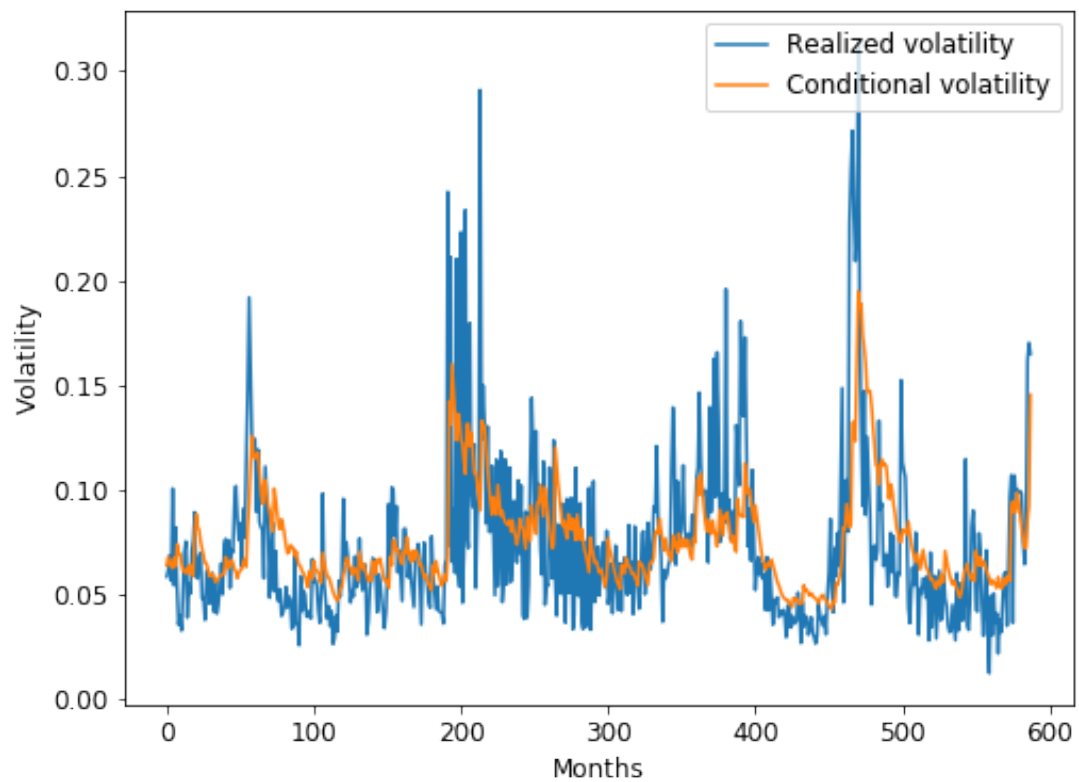
```
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const         0.0040      0.005      0.877      0.381      -0.005      0.013
x1            0.9064      0.058     15.589      0.000      0.792      1.021
=====
```

```
=====
Omnibus:          250.201      Durbin-Watson:          1.436
Prob(Omnibus):    0.000      Jarque-Bera (JB):       1427.327
Skew:             1.810      Prob(JB):               1.15e-310
Kurtosis:         9.719      Cond. No.               42.4
=====
```

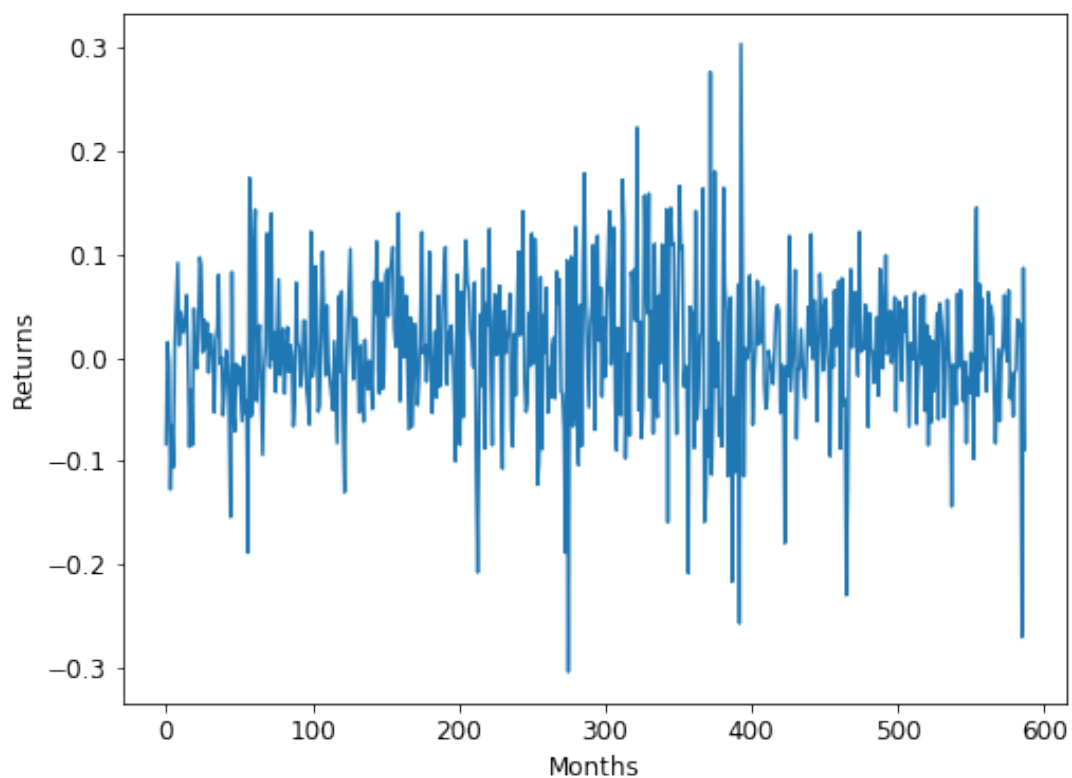
Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

MSE: 0.0011303178725781107
RMSE: 0.03362020036493106
MAE: 0.023377654172692438



```
In [10]: question_2_3_wrapper(IBM_returns, IBM_vol)
```



Iteration: 5, Func. Count: 40, Neg. LLF: -744.0163485046576
 Optimization terminated successfully. (Exit mode 0)
 Current function value: -744.0373348127415
 Iterations: 9
 Function evaluations: 69
 Gradient evaluations: 9

GARCH

Constant Mean - GARCH Model Results

```

=====
Dep. Variable:          y      R-squared:          -0.000
Mean Model:           Constant Mean  Adj. R-squared:        -0.000
Vol Model:              GARCH      Log-Likelihood:       744.037
Distribution:           Normal     AIC:                  -1480.07
Method:                Maximum Likelihood  BIC:                  -1462.57
                                           No. Observations:     588
Date:                  Wed, May 15 2019   Df Residuals:         584
  
```

Time: 22:37:35 Df Model: 4
Mean Model

	coef	std err	t	P> t	95.0% Conf. Int.
mu	7.1227e-03	2.619e-03	2.720	6.529e-03	[1.990e-03,1.226e-02]
Volatility Model					
	coef	std err	t	P> t	95.0% Conf. Int.
omega	4.0907e-04	2.831e-04	1.445	0.148	[-1.457e-04,9.638e-04]
alpha[1]	0.0920	4.769e-02	1.928	5.381e-02	[-1.506e-03, 0.185]
beta[1]	0.8272	9.068e-02	9.122	7.353e-20	[0.649, 1.005]

Covariance estimator: robust

Regression

OLS Regression Results

=====						
Dep. Variable:	y	R-squared:		0.258		
Model:	OLS	Adj. R-squared:		0.257		
Method:	Least Squares	F-statistic:		204.2		
Date:	Wed, 15 May 2019	Prob (F-statistic):		5.74e-40		
Time:	22:37:35	Log-Likelihood:		1265.6		
No. Observations:	588	AIC:		-2527.		
Df Residuals:	586	BIC:		-2519.		
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

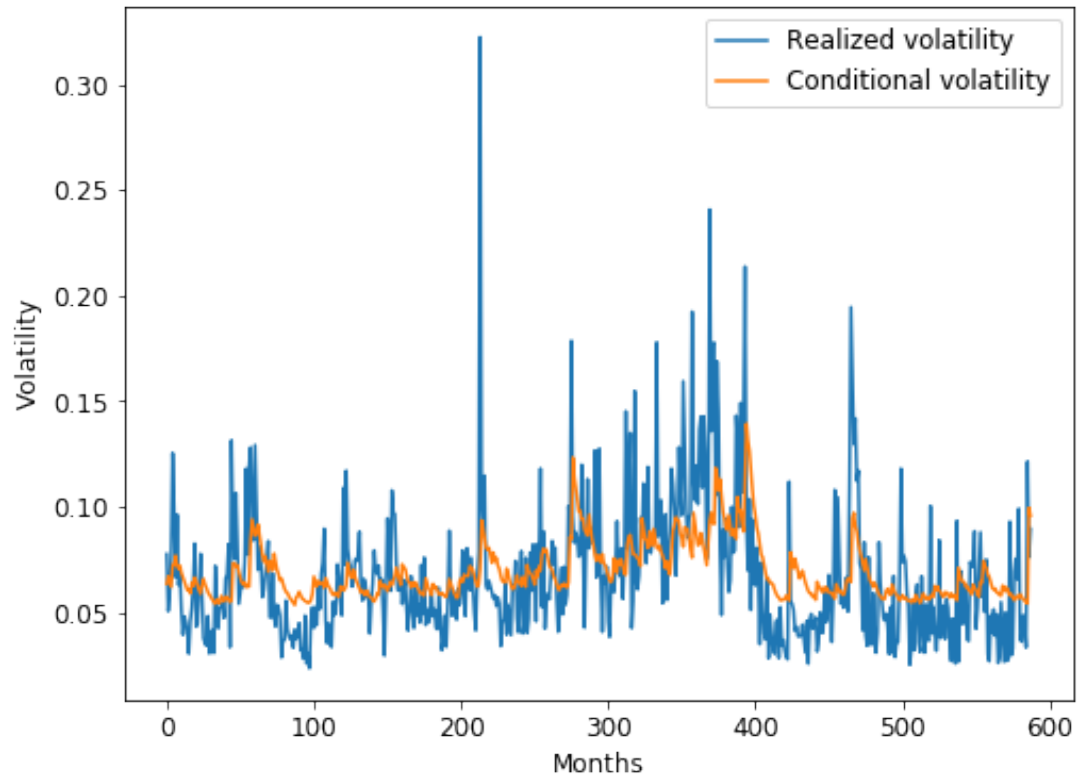
const	-0.0174	0.006	-2.889	0.004	-0.029	-0.006
x1	1.2202	0.085	14.292	0.000	1.053	1.388
=====						
Omnibus:	349.718	Durbin-Watson:		1.460		
Prob(Omnibus):	0.000	Jarque-Bera (JB):		4941.665		
Skew:	2.348	Prob(JB):		0.00		
Kurtosis:	16.403	Cond. No.		73.9		

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
MSE: 0.0008042179880348184

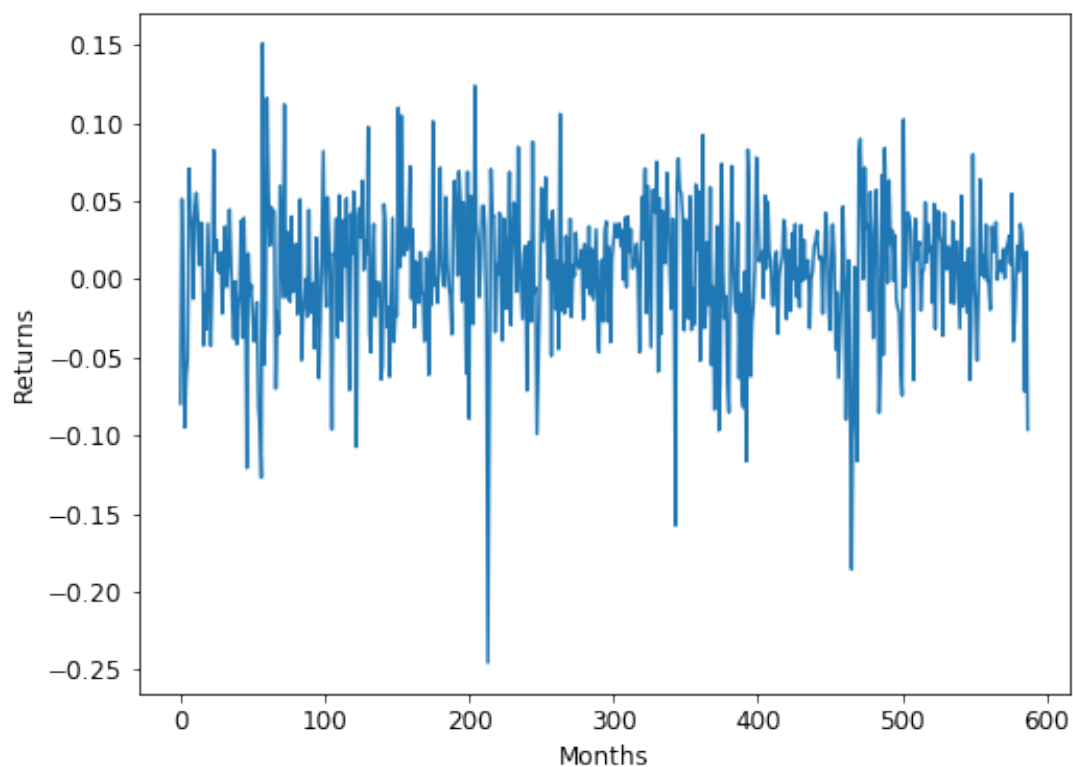
RMSE: 0.028358737419617582

MAE: 0.02011790230504535



0.3 Problem 3

```
In [11]: question_2_3_wrapper(SP_returns, SP_vol, p=1, o=1, q=1)
```



```

Iteration:      5,   Func. Count:    48,   Neg. LLF: -1033.2384143407485
Iteration:     10,   Func. Count:    88,   Neg. LLF: -1033.4340740534335
Optimization terminated successfully.   (Exit mode 0)
    Current function value: -1033.4444615316143
    Iterations: 13
    Function evaluations: 109
    Gradient evaluations: 13

```

GARCH

Constant Mean - GJR-GARCH Model Results

```

=====
Dep. Variable:                y      R-squared:                -0.000
Mean Model:      Constant Mean    Adj. R-squared:           -0.000
Vol Model:       GJR-GARCH        Log-Likelihood:         1033.44
Distribution:    Normal           AIC:                      -2056.89
Method:         Maximum Likelihood BIC:                      -2035.01
                                           No. Observations:      588
Date:           Wed, May 15 2019    Df Residuals:          583

```


Time: 22:37:36 Df Model: 5
Mean Model

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
mu          5.5600e-03  1.619e-03      3.434  5.941e-04 [2.387e-03,8.733e-03]
              Volatility Model
=====
```

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega       1.2773e-04  1.286e-04      0.994    0.320 [-1.242e-04,3.797e-04]
alpha[1]     0.0557  8.823e-02      0.631    0.528 [ -0.117, 0.229]
gamma[1]     0.1156    0.134      0.860    0.390 [ -0.148, 0.379]
beta[1]      0.8216  6.827e-02     12.036  2.303e-33 [ 0.688, 0.955]
=====
```

Covariance estimator: robust

Regression

OLS Regression Results

```
=====
Dep. Variable:          y      R-squared:          0.276
Model:                  OLS      Adj. R-squared:      0.275
Method:                 Least Squares      F-statistic:      223.8
Date:                  Wed, 15 May 2019      Prob (F-statistic):      4.36e-43
Time:                  22:37:36      Log-Likelihood:      1445.2
No. Observations:      588      AIC:      -2886.
Df Residuals:          586      BIC:      -2878.
Df Model:              1
Covariance Type:       nonrobust
=====
```

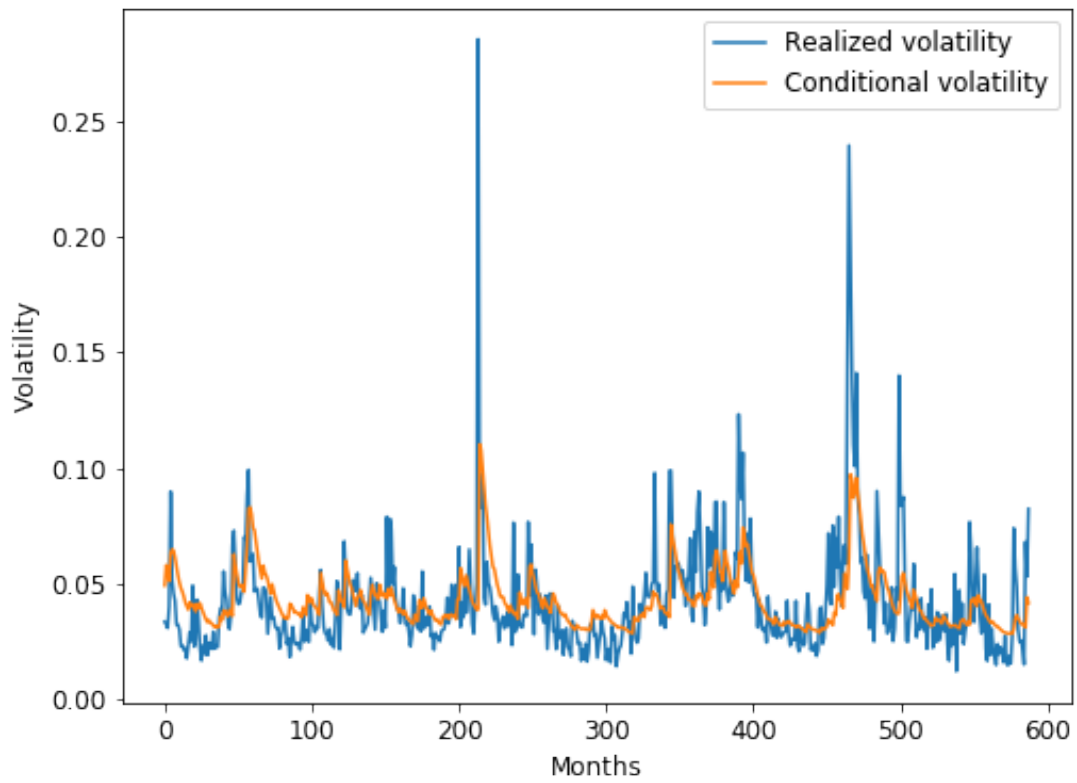
```
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const       -0.0020      0.003     -0.665    0.507     -0.008      0.004
x1          1.0222      0.068     14.959    0.000      0.888      1.156
=====
```

```
Omnibus:          644.224      Durbin-Watson:          1.204
Prob(Omnibus):    0.000      Jarque-Bera (JB):      53325.985
Skew:            4.971      Prob(JB):          0.00
Kurtosis:        48.582      Cond. No.          80.0
=====
```

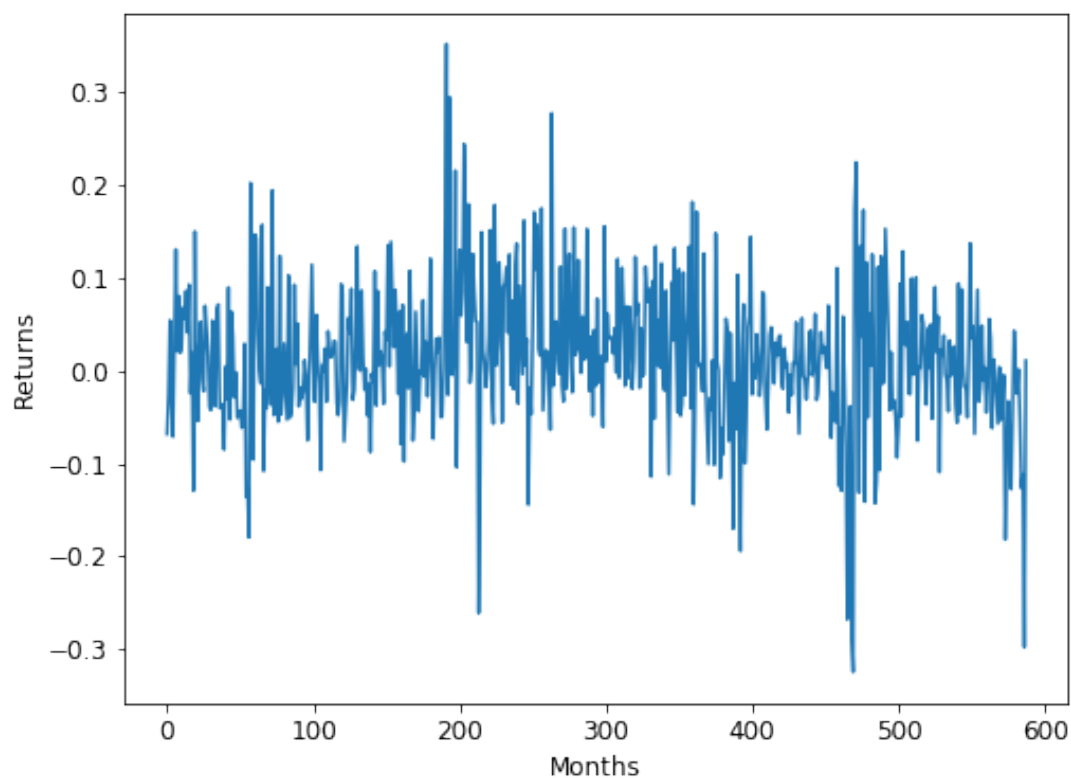
Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

MSE: 0.0004304619546796192
RMSE: 0.020747577079736786
MAE: 0.01258169591706444



In [12]: question_2_3_wrapper(GE_returns, GE_vol, p=1, o=1, q=1)



```

Iteration:      5,   Func. Count:    55,   Neg. LLF: -711.0202362292537
Iteration:     10,   Func. Count:    97,   Neg. LLF: -711.494028499884
Optimization terminated successfully.   (Exit mode 0)
Current function value: -711.508038708007
Iterations: 14
Function evaluations: 119
Gradient evaluations: 13

```

GARCH

Constant Mean - GJR-GARCH Model Results

```

=====
Dep. Variable:                y      R-squared:                -0.000
Mean Model:      Constant Mean  Adj. R-squared:           -0.000
Vol Model:      GJR-GARCH      Log-Likelihood:          711.508
Distribution:    Normal        AIC:                      -1413.02
Method:         Maximum Likelihood  BIC:                  -1391.13
                                           No. Observations:      588

```

Date: Wed, May 15 2019 Df Residuals: 583
Time: 22:37:37 Df Model: 5

Mean Model

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
mu           0.0135   2.902e-03      4.666   3.075e-06 [7.853e-03,1.923e-02]
```

Volatility Model

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega        2.3804e-04  1.197e-04      1.989   4.671e-02 [3.466e-06,4.726e-04]
alpha[1]      0.1037   3.450e-02      3.005   2.656e-03 [3.606e-02, 0.171]
gamma[1]      0.0695   4.516e-02      1.540     0.124 [-1.898e-02, 0.158]
beta[1]       0.8340   3.398e-02     24.543   5.181e-133 [ 0.767, 0.901]
=====
```

Covariance estimator: robust

Regression

OLS Regression Results

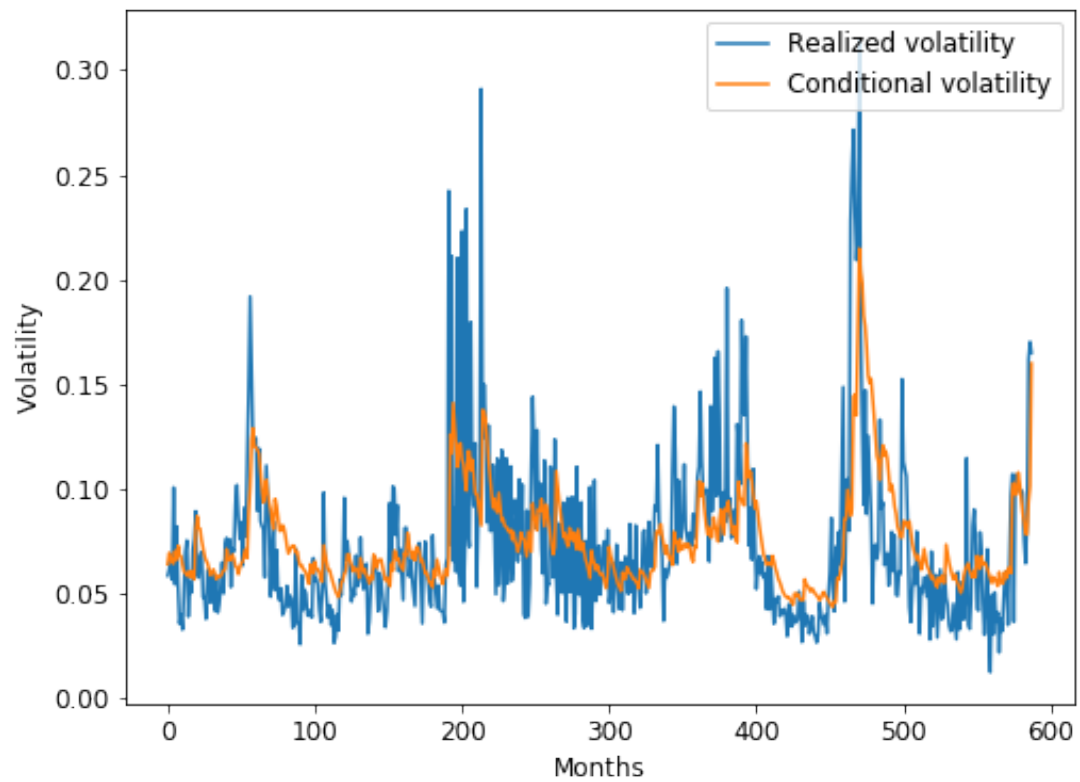
```
=====
Dep. Variable:          y      R-squared:          0.327
Model:                  OLS      Adj. R-squared:        0.326
Method:                 Least Squares      F-statistic:        284.5
Date:                   Wed, 15 May 2019      Prob (F-statistic):    2.53e-52
Time:                   22:37:37      Log-Likelihood:       1178.6
No. Observations:       588      AIC:                -2353.
Df Residuals:           586      BIC:                -2344.
Df Model:                1
Covariance Type:        nonrobust
=====
```

```
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const         0.0017      0.004      0.390      0.697      -0.007      0.010
x1            0.9399      0.056     16.867      0.000      0.830      1.049
=====
```

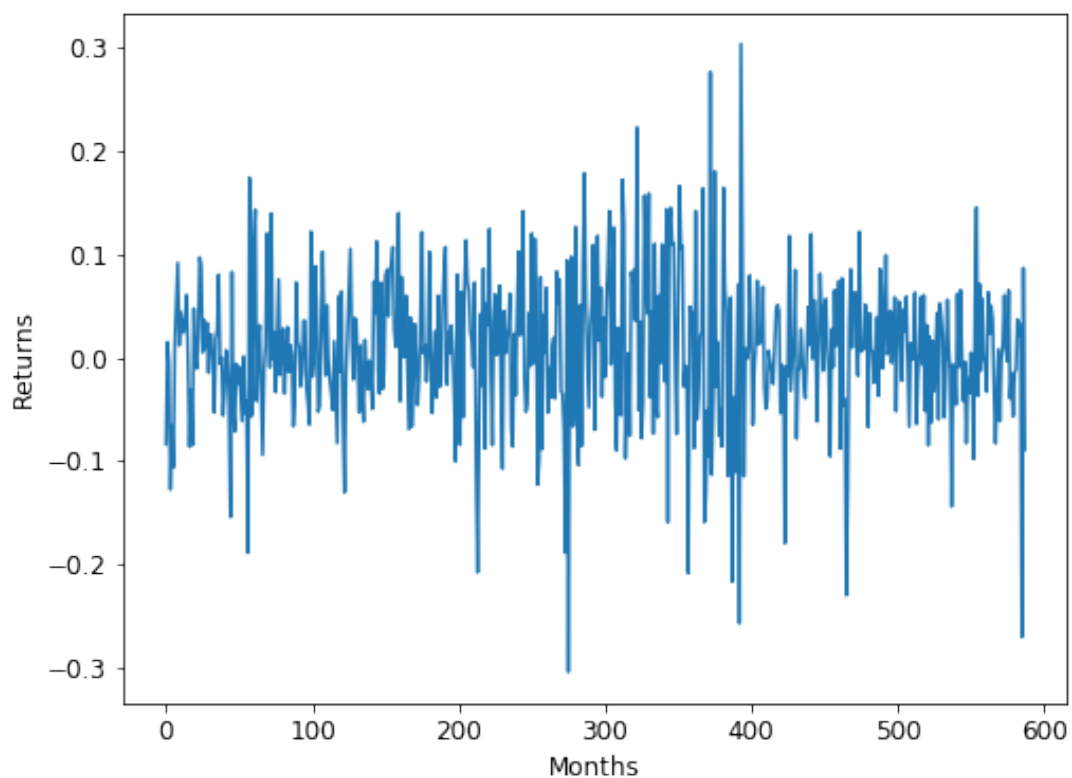
```
=====
Omnibus:          259.460      Durbin-Watson:          1.482
Prob(Omnibus):    0.000      Jarque-Bera (JB):       1615.945
Skew:             1.856      Prob(JB):               0.00
Kurtosis:         10.224      Cond. No.               41.6
=====
```

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
MSE: 0.001072981174468369
RMSE: 0.03275639135296147
MAE: 0.022813610643813383



In [13]: question_2_3_wrapper(IBM_returns, IBM_vol, p=1, o=1, q=1)



Iteration: 5, Func. Count: 50, Neg. LLF: -745.883298363608
 Optimization terminated successfully. (Exit mode 0)
 Current function value: -745.8882723590806
 Iterations: 8
 Function evaluations: 74
 Gradient evaluations: 8

GARCH

Constant Mean - GJR-GARCH Model Results

```

=====
Dep. Variable: y R-squared: -0.000
Mean Model: Constant Mean Adj. R-squared: -0.000
Vol Model: GJR-GARCH Log-Likelihood: 745.888
Distribution: Normal AIC: -1481.78
Method: Maximum Likelihood BIC: -1459.89
No. Observations: 588
Date: Wed, May 15 2019 Df Residuals: 583
  
```

Time: 22:37:38 Df Model: 5
Mean Model

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
mu          6.5981e-03  2.595e-03      2.543  1.099e-02  [1.513e-03,1.168e-02]
              Volatility Model
=====
```

```
=====
              coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega       5.0976e-04  3.418e-04      1.491      0.136  [-1.602e-04,1.180e-03]
alpha[1]      0.0552  3.123e-02      1.766  7.739e-02  [-6.056e-03, 0.116]
gamma[1]      0.0979  9.656e-02      1.014      0.310  [-9.132e-02, 0.287]
beta[1]       0.7961  9.053e-02      8.794  1.448e-18  [ 0.619, 0.973]
=====
```

Covariance estimator: robust

Regression

OLS Regression Results

```
=====
Dep. Variable:          y      R-squared:          0.262
Model:                  OLS      Adj. R-squared:      0.260
Method:                 Least Squares      F-statistic:      207.7
Date:                  Wed, 15 May 2019      Prob (F-statistic):      1.60e-40
Time:                  22:37:38      Log-Likelihood:      1266.9
No. Observations:      588      AIC:      -2530.
Df Residuals:          586      BIC:      -2521.
Df Model:              1
Covariance Type:      nonrobust
=====
```

```
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
const       -0.0107      0.006      -1.930      0.054      -0.022      0.000
x1          1.1232      0.078      14.411      0.000      0.970      1.276
=====
```

```
=====
Omnibus:          341.978      Durbin-Watson:          1.473
Prob(Omnibus):    0.000      Jarque-Bera (JB):      4617.936
Skew:            2.294      Prob(JB):          0.00
Kurtosis:        15.940      Cond. No.          67.6
=====
```

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

MSE: 0.0007950274571317165
RMSE: 0.02819623125759392
MAE: 0.0201179591229439

