

---

# HAR Model on RV

## Table of Contents

Data Processing and Running In-Sample Regression and Out-of-Sample Test .....	1
Normality Assumption Diagnostic .....	1
Regression Goodness of fit and Out-of-Sample Test .....	2

Author: Peiliang Guo

Regression on the log of RV with different lags

## Data Processing and Running In-Sample Regression and Out-of-Sample Test

```
sp500data = readtable('RV_DATA.xlsx');
sp500rv = table2array(sp500data(:,3));
x_yr = x2mdate(sp500data.date);

%calculating predictor values
win_sizes = [1 5 10];
ma_p = zeros(length(sp500rv),length(win_sizes));
for i = 1:length(win_sizes)
    ma_p(:,i) = filter([0 ones(1,win_sizes(i))], win_sizes(i), sp500rv);
end

%regressing log rv on log regressor
rv_X = [ones(length(sp500rv)-max(win_sizes),1)...
        log(ma_p(max(win_sizes)+1:end,:))];
rv_y = log(sp500rv(max(win_sizes)+1:end));
phi = regress(rv_y,rv_X);
res = rv_y - rv_X*phi;
sig_res = std(res);
%transforming back to RV
rv_fitted = exp(rv_X*phi)*exp(sig_res^2/2);

%out of sample forecast after 2000
test_idx = find(x_yr== datenum('01/03/2000','mm/dd/yyyy'));
test_rv_pred = zeros(length(sp500rv)-test_idx,1);
for i = test_idx:length(rv_y)
    rv_Xi = rv_X(1:i-1,:);
    rv_yi = rv_y(1:i-1);
    phi = regress(rv_yi,rv_Xi);
    sig_resi = std(rv_yi - rv_Xi*phi);
    test_rv_pred(i-test_idx+1) = exp(rv_X(i,:)*phi)*exp(sig_resi^2/2);
end
```

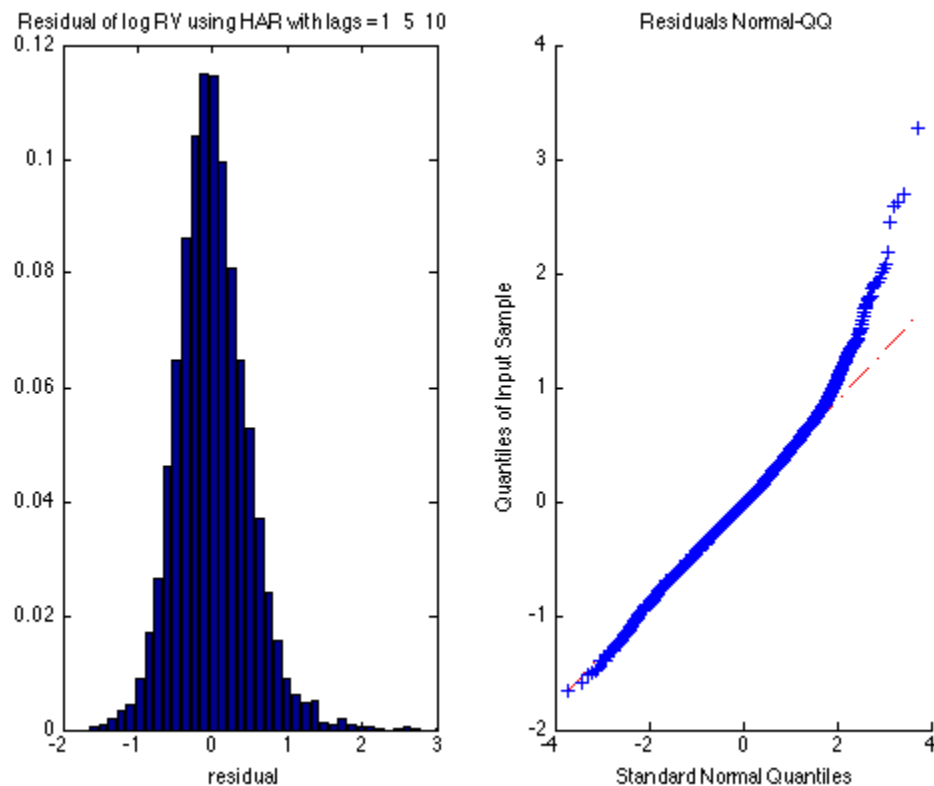
## Normality Assumption Diagnostic

```
figure(1)
```

```

subplot(1,2,1)
[c,x_hist] = hist(res,40,'Normalization','probability');
bar(x_hist,c/sum(c));
title(strcat('Residual of log RV using HAR with lags = ', num2str(win_sizes)))
xlabel('residual')
xlim([-2 3])
subplot(1,2,2)
qqplot(res)
title('Residuals Normal-QQ')

```



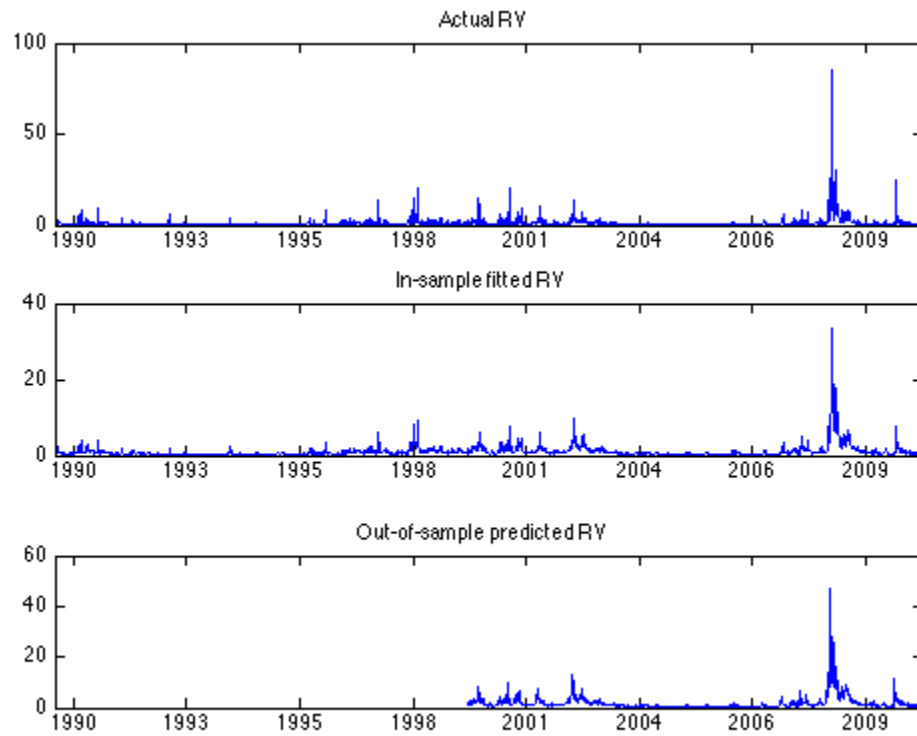
## Regression Goodness of fit and Out-of-Sample Test

```

figure(2)
subplot(3,1,1)
plot(x_yr(max(win_sizes)+1:end),sp500rv(max(win_sizes)+1:end))
xlim([x_yr(1) x_yr(end)])
datetick('x','kepticks','keeplimits')
title('Actual RV')
subplot(3,1,2)
plot(x_yr(max(win_sizes)+1:end),rv_fitted)
xlim([x_yr(1) x_yr(end)])
datetick('x','kepticks','keeplimits')
title('In-sample fitted RV')
subplot(3,1,3)

```

```
plot(x_yr(test_idx:end-1),test_rv_pred)
xlim([x_yr(1) x_yr(end)])
datetick('x','kepticks','keeplimits')
title('Out-of-sample predicted RV')
```



*Published with MATLAB® R2014a*