Empirical Methods in Financial Econometrics:Project 7

Xijie Zhou

November 1, 2019

Exercise 1

A

The function is below:

Listing 1: realized_var_s.m

```
function rv_s = realized_var_s(kn,prices)
prices_s = prices(1:kn:end,:);
deltax_s = diff(prices_s);
rv_s = realized_var(deltax_s);
end
```

B

The figure is below:

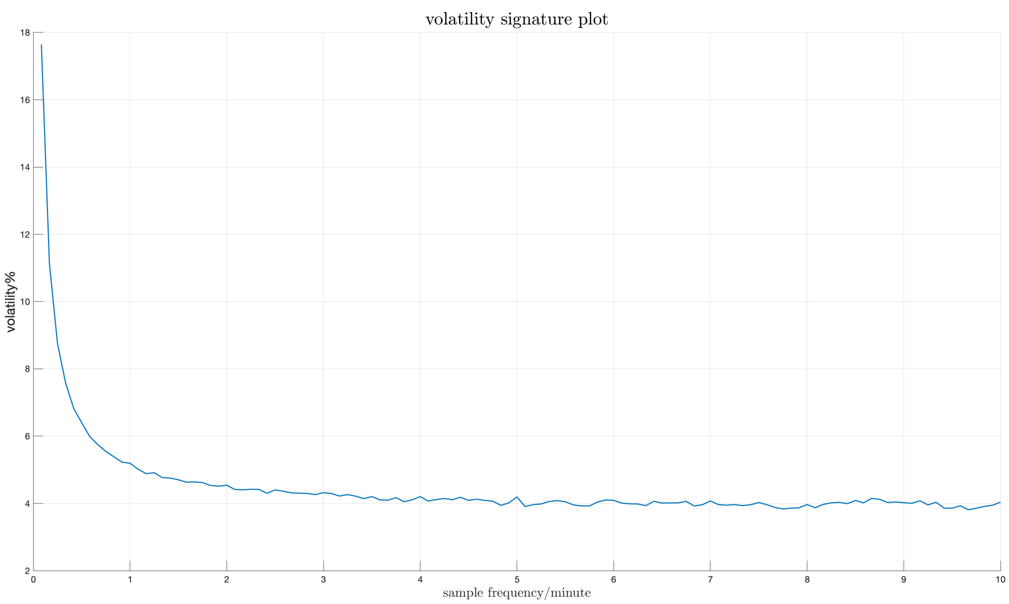

Figure 1: Volatility signature plot

Interpret: As we can see, as the sampling size increased, the RV decreased from 18% to 4%, then kept relatively stable.

C

The price data we have could be measured with error. When using the high-frequency data to compute RV, we obtain the noise, because RV is almost equal to $IV + 2n\sigma_\chi^2$. IV is so small, so the noise dominant the whole result. Then, $\hat{\sigma}_\chi^2$ is the variance of the difference between theoretical prices and observed prices.

D

Noise is $2\frac{n}{k_n}\sigma_\chi^2$, and RV is IV plus Noise, so the contribution is as the question shows.

E

The figure is below:

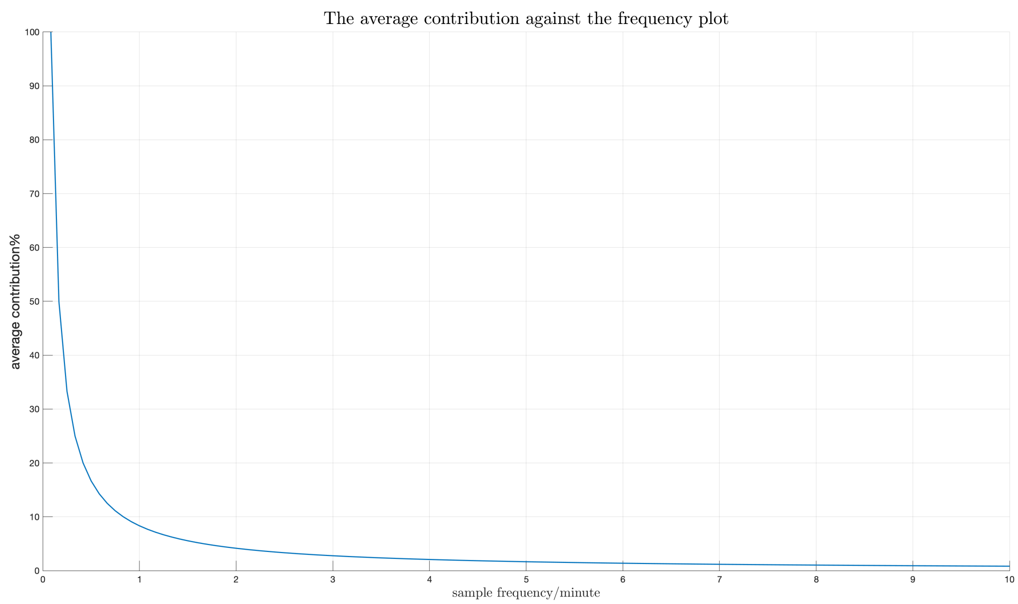

Figure 2: The average contribution against the frequency plot

Interpret: When the sampling size is enough small, the contribution is close to 100%. The noise is unimportant at low frequency in my data set, because it is obvious to see that the noise contributes less than 10% of the total variation at 5-min and 8-min in my figure.

F

The figure is below:

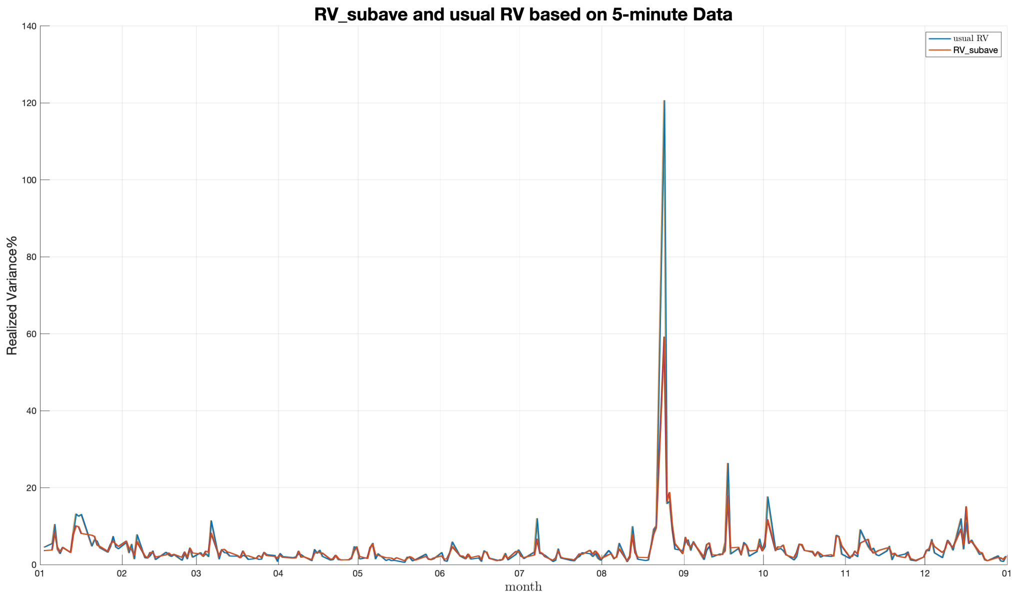

Figure 3: RV_subave and usual RV based on 5-minute Data

Interpret: There are some small differences between the estimates. There should be some little differences, because we take the average of different RV by different observations with the same sampling frequency.

G

The figure is below:

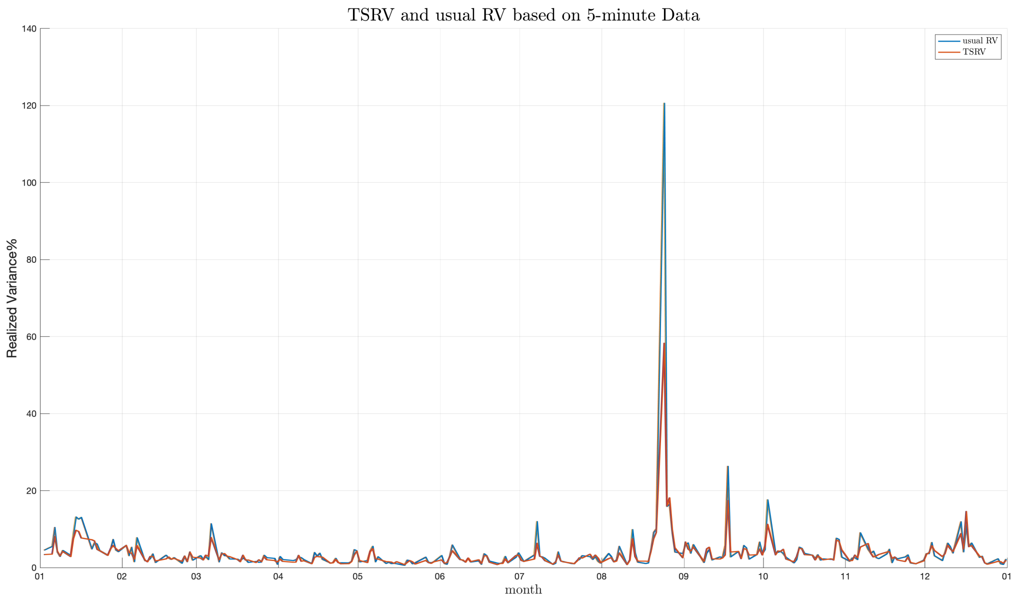

Figure 4: TSRV and usual RV based on 5-minute Data

Interpret: There is only little difference between two lines, because the subaverage RV takes the average of different samples with same frequency.

H

The figure is below:

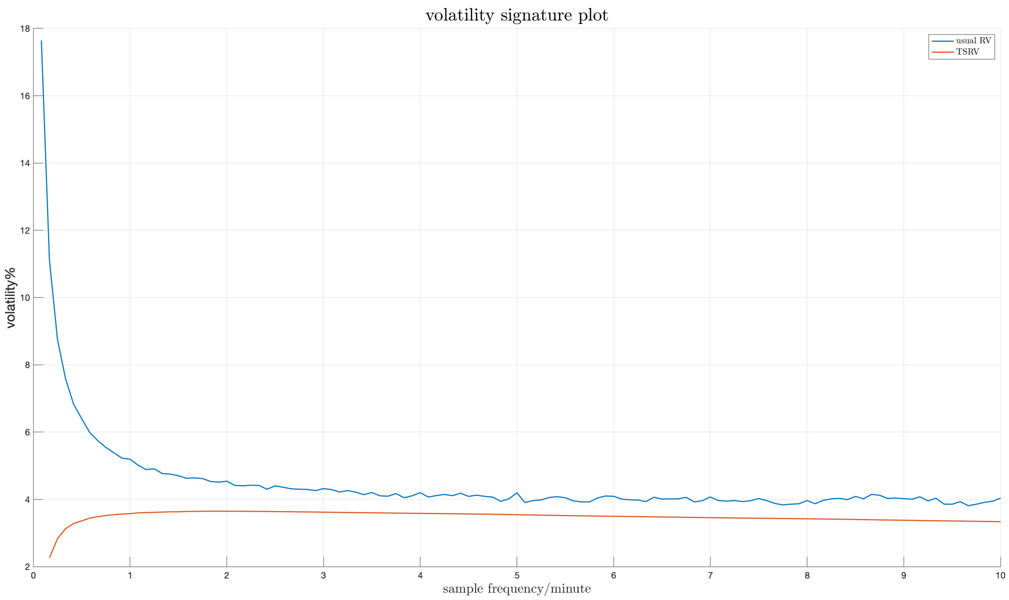

Figure 5: volatility signature plot

Interpret: Yes, it does. Because it deleted microstructure noise when computing the value. The curve relatively flat when the sampling frequency increases. Compared to the usual RV estimator, TSRV is smaller when sampling at high frequency, and always smaller than the usual RV with the smaller difference when sampling frequency increases.

Code

The code is below:

Listing 2: ex_1.m

```
% question A
[dates, prices] = load_stock2('BAC-2015.csv');
kn = 60;
rv_s = realized_var_s(kn,prices);

% question B
average_rv_s = zeros(1,120);
for i =1:1:120
    rv_s=realized_var_s(i,prices);
    average_rv_s(i) = mean(rv_s);
end
```

```
frequency=1:120;

f = figure(1);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(frequency/12,100*252*average_rv_s)
title('volatility signature plot');
box off; grid on;
xlabel('sample frequency/minute');
ylabel('volatility%');
print(f,'-dpng','-r200','figures/1B');
close(f);

% question C
deltax_y = diff(prices);
rv_y = realized_var(deltax_y);
[n,T] = size(deltax_y);
sigma_chi_2 = rv_y/(2*n);

% question D
% question E
contribution_kn = zeros(1,120);
average_contribution_kn = zeros(1,120);
for i = 1:1:120
contribution_kn = (2*(n/i)*sigma_chi_2)./rv_y;
average_contribution_kn(i) = 100*mean(contribution_kn);
end

f = figure(2);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(frequency/12,average_contribution_kn)
title('The average contribution against the frequency plot');
box off; grid on;
xlabel('sample frequency/minute');
ylabel('average contribution%');
print(f,'-dpng','-r200','figures/1E');
close(f);
```

```

% question F
rv_subave = realized_var_sub(kn,prices);

f = figure(3);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates,100*252*rv_s)
hold on
plot(dates,100*252*rv_subave)
hold off
datetick('x','mm')
xlabel('month')
ylabel('Realized Variance%')
title('RV_subave and usual RV based on 5-minute Data');
legend('usual RV','RV_subave')
box off; grid on;
print(f,'-dpng','-r200','figures/1F');
close(f);

% question G
tsrv = rv_subave - 2*n/60*sigma_chi_2;

f = figure(4);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates,100*252*rv_s)
hold on
plot(dates,100*252*tsrv)
hold off
datetick('x','mm')
xlabel('month')
ylabel('Realized Variance%')
title('TSRV and usual RV based on 5-minute Data');
legend('usual RV','TSRV')
box off; grid on;
print(f,'-dpng','-r200','figures/1G');
close(f);
```

```

% question H
tsrv_s = zeros(120,T);
for kn = 1:120
    rv_subave = realized_var_sub(kn,prices);
    tsrv_s(kn,:) = rv_subave - 2*n/kn*sigma_chi_2;
end

average_tsrv = mean(tsrv_s,2);

f = figure(5);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(frequency/12,100*252*average_rv_s)
hold on
plot(frequency(2:end)/12,100*252*average_tsrv(2:end))
hold off
title('volatility signature plot');
box off; grid on;
xlabel('sample frequency/minute');
ylabel('volatility%');
legend('usual RV','TSRV')
print(f,'-dpng','-r200','figures/1H');
close(f);
```