Empirical Methods in Financial Econometrics:Midterm Exam

Xijie Zhou

October 2, 2019

Question 0

The stocks I will use for solving this exam:

Stock	Ticker
1	DIS
2	PG
3	SPY

Question 1

A

The “t” for the date ”September 16, 2008” is 430.

It is impossible to estimate IV for today if it is still in the morning. Because we don’t have enough returns to estimate IV.

B

The estimate for yesterday’s integrated variance of DIS is 4.5898×10^{-04}

The estimate for yesterday’s integrated variance of PG is 2.2365×10^{-04}

Interpret: The estimate for integrated variance I used is the truncated variance, which is the sum of the square of the continuous returns. Because we can not get the integrated variance directly, the truncated variance is used to get an asymptotic one.

The estimate for next day’s integrated variance of DIS is 9.2363×10^{-04}

The estimate for next day’s integrated variance of PG is 7.5846×10^{-04}

Justify: We can see the estimate of next day’s integrated variance is obviously bigger than the estimate of yesterdays’s integrated variance, so it is not wise to use the yesterday’s estimate to forecast the next one. The reason could be the data of one day is much less than what we need to forecast.

C

		Estimate of $\mathrm{P}(r_t^c \leq -q)$		
Stock	q	Estimate	Lower bound	Upper bound
1	2%	0.1753	0.1268	0.2093
1	4%	0.0309	0.0112	0.0528
2	2%	0.0906	0.0397	0.1308
2	4%	0.0037	2.2492×10^{-4}	0.0124

Comment: The width of the confidence interval in DIS is 0.0825 and 0.0416 respectively. The width of the confidence interval in DIS is 0.0911 and 0.0121 respectively. Most of them are around or in the range 0.01-0.06, so I think the estimation accuracy is good.

D

		VaR Estimate and 95% Confidence Interval		
Stock	p	VaR	Lower bound	Upper bound
1	1%	-9.9679×10^6	-1.1502×10^7	-8.1495×10^6
1	5%	-7.0478×10^6	-8.1327×10^6	-5.7622×10^6
2	1%	-6.9580×10^6	-8.2884×10^6	-5.3038×10^6
2	5%	-4.9197×10^6	-5.8604×10^6	-3.7501×10^6

Comment: The width of DIS is 3.3528 million(1%) and 2.3706 million(5%) respectively; The width of PG is 2.9847 million(1%) and 2.1103 million(5%) respectively. At least we can find the range of the loses of each stocks, or the maximum and minimum of the loses of each stocks, so I think my boss could find them helpful by using this information to adjust stock holding.

Question 2

A

		Summary Statistics of Long-Short Continuous Returns		
Average	Min	5% Percentile	95% Percentile	Max
-1.5566×10^{-6}	-0.0357	-0.0022	0.0021	0.0416

B

		Summary Statistics of the Realized Beta		
Average	Min	5% Percentile	95% Percentile	Max
0.3899	-1.3388	-0.2081	0.9650	2.2904

C

The figure is below:

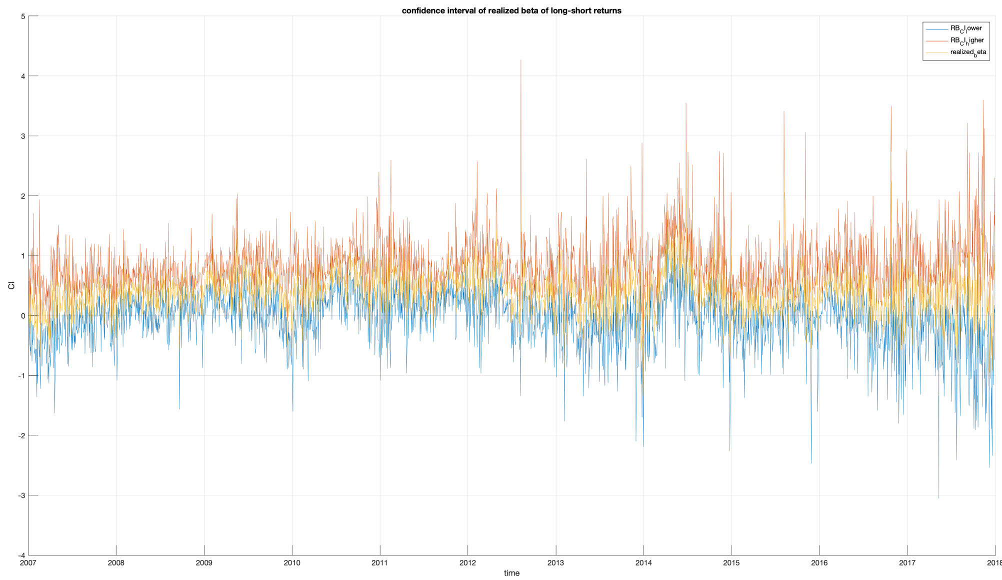

Figure 1: confidence interval of realized beta of long-short returns

Comment: I think the long-short position appears market neutral at least visually, because the realized beta of the long-short portfolio is around 0.

D

The figure is below:

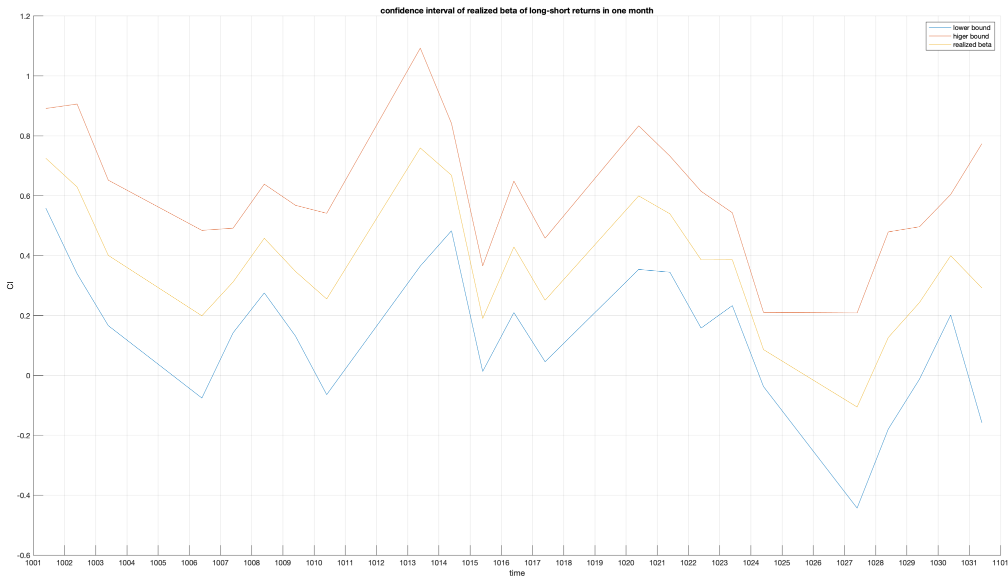

Figure 2: confidence interval of realized beta of long-short returns in one month

Interpret: The trend of the realized beta of long-short returns in one month is fluctuated from 0 to 1. The confidence intervals of the realized beta of long-short returns are good because the plot of the realized beta of long-short returns is under the upper confidence interval and above the lower of confidence interval.

E

The total number of rejections are 1397.

The percentage of rejections is 50.45%. It is far bigger than 10%, so I think the hedge fund holding long stock 1 and short stock 2 would not be considered market neutral.

Question 3

A

		Summary Statistics for Continuous Returns			
Stock	Average	Min	5% Percentile	95% Percentile	Max
1	4.3967×10^{-6}	-0.0323	-0.0022	0.0022	0.0474
2	5.4237×10^{-6}	-0.0254	-0.0015	0.0016	0.0420
SPY	1.6385×10^{-6}	-0.0233	-0.0015	0.0015	0.0376

B

I detected 116 jumps in the market.

C

		Summary Statistics of the Realized Beta			
Stock	Average	Min	5% Percentile	95% Percentile	Max
1	0.9312	-1.0811	0.4731	1.3575	-1.0811
2	0.5412	-0.4167	0.1475	0.9690	-0.4167

D

The figure is below:

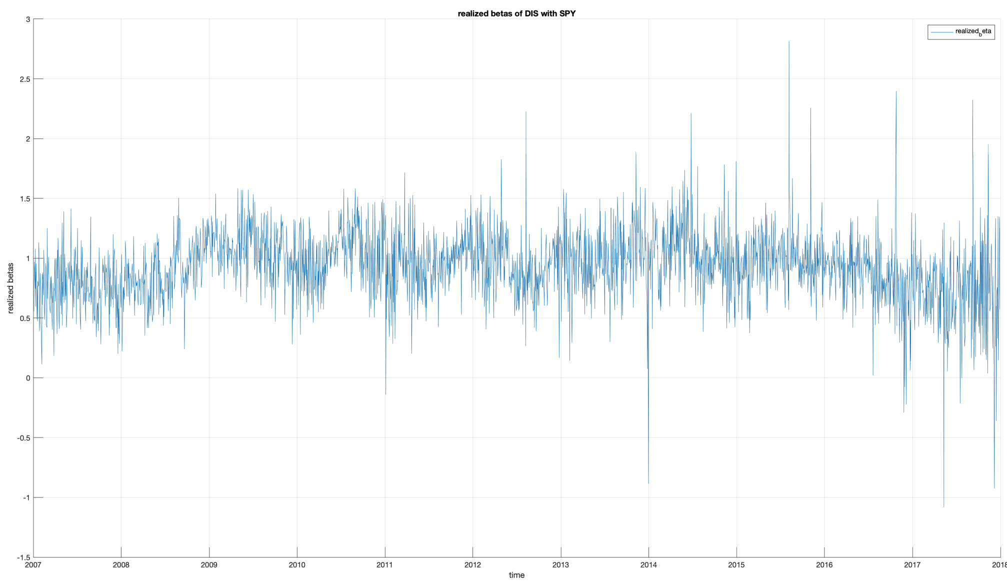

Figure 3: realized betas of DIS with SPY

The figure is below:

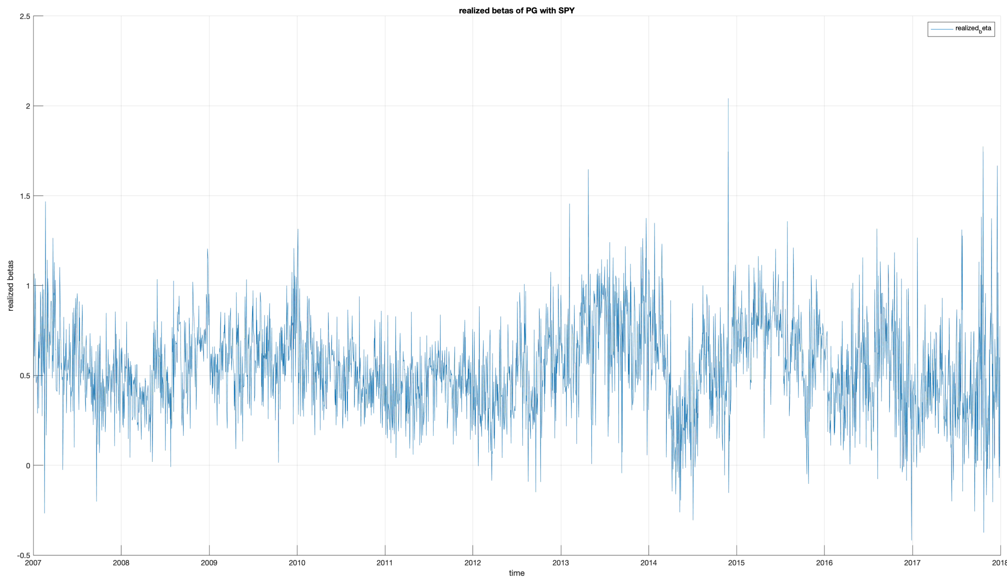

Figure 4: realized betas of PG with SPY

From the plots I can see, the DIS is generally as risky as the market because the β is around 1 ; the PG is generally less risky than the market because theβ is below than 1.

E

		Summary Statistics of the Residual Correlation		
Average	Min	5% Percentile	95% Percentile	Max
0.0196	-0.4993	-0.2476	0.2909	0.5563

F

The figure is below:

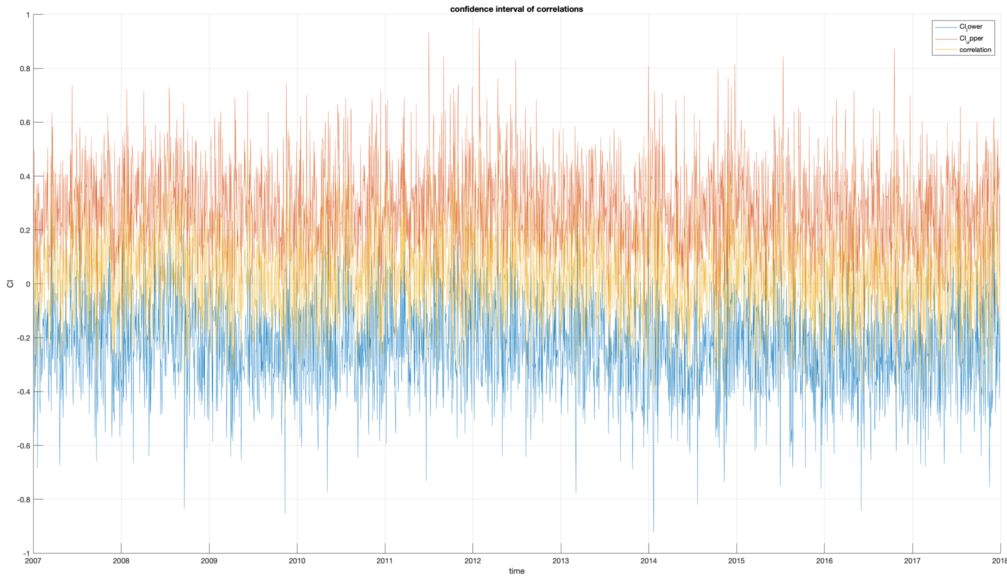

Figure 5: confidence interval of realized correlations

Interpret: The correlation is fluctuated dramatically around 0 from 2007 to 2017. The confidence intervals of the correlation follow the correlation to fluctuate at the same time. The upper and lower limits of the confidence intervals are 1 and -1 respectively.

G

The figure is below:

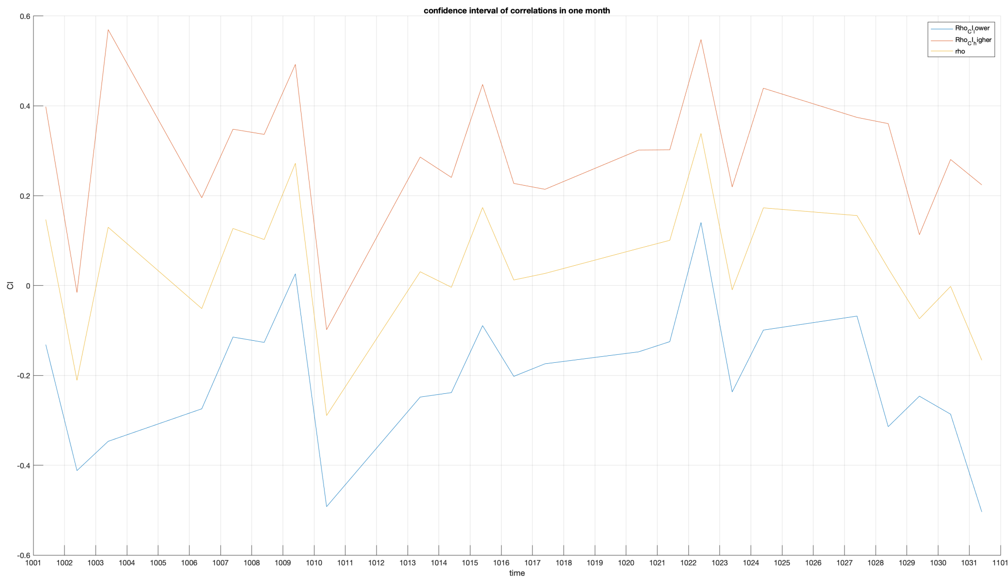

Figure 6: confidence interval of realized correlations in one month

Interpret: We can see the trend more clearly in one month. The correlation is fluctuate around 0 from -0.3 to 0.4. The upper and lower limits of the confidence intervals are 0.6 and -0.6 respectively.

H

The number of p0 is 10.9831.

The number of p+ is 0.7897.

The number of p- is 0.4722.

Interpret: The p0 is 10.9831, which is most of the 11 years, so we do not have enough evidences to reject H0, which is the realized correlation is 0.

Code

The code is below:

Listing 1: q_1.m

```
% A
[dates, prices] = load_stock('DIS.csv');
[dates_return,deltax] = log_returns(dates, prices);
[n,T] = size(deltax);

```

```
zoom_end_916 = find(dates_return==datenum(2008,9,16,16,00,0))/n;
dates_zoom_916 = dates_return(:,zoom_end_916);

% B
zoom_end_915 = find(dates_return==datenum(2008,9,15,16,00,0))/n;
dates_zoom_915 = dates_return(:,zoom_end_915);

deltan = 1/n;
deltax_915 = deltax(:,zoom_end_915);
tau_915 = tau_f(deltax_915);
BV_915 = bipower_var(deltax_915);
alpha = 5;
cutoff = alpha*deltan^0.49*sqrt(tau_915*BV_915);
rc_915 = deltax_915;
rc_915(abs(deltax_915)>cutoff)=0;

TV_915 = sum((rc_915).^2);

% use the same way on 916

deltax_916 = deltax(:,zoom_end_916);
tau_916 = tau_f(deltax_916);
BV_916 = bipower_var(deltax_916);
alpha = 5;
cutoff = alpha*deltan^0.49*sqrt(tau_916*BV_916);
rc_916 = deltax_916;
rc_916(abs(deltax_916)>cutoff)=0;

TV_916 = sum((rc_916).^2);

% C
QIV_915 = 1/(3*deltan)*sum((rc_915).^4);
CI_IV_915 = [TV_915-1.96*sqrt(2*deltan*QIV_915),TV_915+1.96*sqrt(2*deltan*
    QIV_915)];

pd = makedist('Normal','mu',0,'sigma',1);
```

```
p_2 = cdf(pd,-0.02/sqrt(TV_915));
p_4 = cdf(pd,-0.04/sqrt(TV_915));

p2_lower = cdf(pd,-0.02/sqrt(CI_IV_915(1,1)));
p2_upper = cdf(pd,-0.02/sqrt(CI_IV_915(1,2)));
P2_CI = [p2_lower,p2_upper];
P2_width = p2_upper - p2_lower;

p4_lower = cdf(pd,-0.04/sqrt(CI_IV_915(1,1)));
p4_upper = cdf(pd,-0.04/sqrt(CI_IV_915(1,2)));
P4_CI = [p4_lower,p4_upper];
P4_width = p4_upper - p4_lower;

% D
Q_1 = icdf(pd,0.01);
Q_5 = icdf(pd,0.05);

VaR_1 = Q_1*sqrt(TV_915)*200000000;
VaR_1_upper = icdf(pd,0.01)*sqrt(CI_IV_915(1,1))*200000000;
VaR_1_lower = icdf(pd,0.01)*sqrt(CI_IV_915(1,2))*200000000;
VaR_1_CI = [VaR_1_lower,VaR_1_upper];

VaR_5 = Q_5*sqrt(TV_915)*200000000;
VaR_5_upper = icdf(pd,0.05)*sqrt(CI_IV_915(1,1))*200000000;
VaR_5_lower = icdf(pd,0.05)*sqrt(CI_IV_915(1,2))*200000000;
VaR_5_CI = [VaR_5_lower,VaR_5_upper];
```

Listing 2: q_2.m

```
% A
[dates_1, prices_1] = load_stock('DIS.csv');
[dates_2, prices_2] = load_stock('PG.csv');

[dates_return_1,deltax_1] = log_returns(dates_1, prices_1);
[dates_return_2,deltax_2] = log_returns(dates_2, prices_2);
[n,T] = size(deltax_1);
deltan = 1/n;
```

```

r_long_short = deltax_1-deltax_2;

tau_r_long_short = tau_f(r_long_short);
BV_r_long_short = bipower_var(r_long_short);
alpha = 5;
cutoff = alpha*deltan^0.49*sqrt(tau_r_long_short*BV_r_long_short);
rc_long_short = r_long_short;
rc_long_short(abs(r_long_short)>cutoff)=0;

average_rc_long_short = mean(rc_long_short(:));
min_rc_long_short = min(rc_long_short(:));
p5_rc_long_short = quantile(rc_long_short(:),0.05);
p95_rc_long_short = quantile(rc_long_short(:),0.95);
max_rc_long_short = max(rc_long_short(:));

% B
[dates_m, prices_m] = load_stock('SPY.csv');
[dates_return_m,deltax_m] = log_returns(dates_m, prices_m);
tau_m = tau_f(deltax_m);
BV_m = bipower_var(deltax_m);
alpha = 5;
cutoff = alpha*deltan^0.49*sqrt(tau_m*BV_m);
rc_m = deltax_m;
rc_m(abs(deltax_m)>cutoff)=0;

realized_beta = sum((rc_m.*rc_long_short))./sum(rc_m.^2);

average_realized_beta = mean(realized_beta(:));
min_realized_beta = min(realized_beta(:));
p5_realized_beta = quantile(realized_beta(:),0.05);
p95_realized_beta = quantile(realized_beta(:),0.95);
max_realized_beta= max(realized_beta(:));

% C
kn = 11;
```

```
M = 7;

J = 1000;
booRB = boost_realized_beta(J,T,n,kn,M,rc_m,rc_long_short);

RB_CI = ci_f(0.05,mean(booRB),std(booRB));

f = figure(1);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_m(1,:),RB_CI(1,:))
hold on
plot(dates_m(1,:),RB_CI(2,:))
plot(dates_m(1,:),realized_beta)
hold off
datetick('x','yyyy');
legend('RB_CI_lower','RB_CI_higher','realized_beta');
title('confidence interval of realized beta of long-short returns');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/2C1');
close(f);

% D
zoom_start = find(dates_return_1==datenum(2008,10,1,16,0,0))/n;
zoom_end = find(dates_return_1==datenum(2008,10,31,16,0,0))/n;
realized_beta_zoom = realized_beta(zoom_start:zoom_end);
RB_CI_zoom = RB_CI(:,zoom_start:zoom_end);
dates_zoom = dates_return_1(1,zoom_start:zoom_end);


f = figure(2);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_zoom(1,:),RB_CI_zoom(1,:))
hold on
plot(dates_zoom(1,:),RB_CI_zoom(2,:))
```

```
plot(dates_zoom(1,:),realized_beta_zoom)
hold off
legend('lower bound','higer bound','realized beta')
% xticks(1:23)
% xticklabels
    ({'1','2','3','7','8','9','10','11','14','15','16','17','18','21','22','23','24','2

datetick('x','mmdd');
title('confidence interval of realized beta of long-short returns in one
    month');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/2D1');
close(f);

% E
contain_0 = zeros(1,T);
for i = 1:T
    contain_0(i) = RB_CI(1,i)<0 && RB_CI(2,i)>0;
end

sum_contain_0 = sum(contain_0);
not_contaion_0 = T - sum_contain_0;
p_not_contaion_0 = not_contaion_0/T;
```

Listing 3: q3.m

```
% A
[dates_1, prices_1] = load_stock('DIS.csv');
[dates_2, prices_2] = load_stock('PG.csv');
[dates_m, prices_m] = load_stock('SPY.csv');

[dates_return_1,deltax_1] = log_returns(dates_1, prices_1);
[dates_return_2,deltax_2] = log_returns(dates_2, prices_2);
[dates_return_m,deltax_m] = log_returns(dates_m, prices_m);

```

```
[n,T] = size(deltax_1);
deltan = 1/n;

tau_1 = tau_f(deltax_1);
tau_2 = tau_f(deltax_2);
tau_m = tau_f(deltax_m);

BV_1 = bipower_var(deltax_1);
BV_2 = bipower_var(deltax_2);
BV_m = bipower_var(deltax_m);

alpha = 5;
cutoff_1 = alpha*deltan^0.49*sqrt(tau_1*BV_1);
cutoff_2 = alpha*deltan^0.49*sqrt(tau_2*BV_2);
cutoff_m = alpha*deltan^0.49*sqrt(tau_m*BV_m);

rc_1 = deltax_1;
rc_1(abs(deltax_1)>cutoff_1)=0;

rc_2 = deltax_2;
rc_2(abs(deltax_2)>cutoff_2)=0;

rc_m = deltax_m;
rc_m(abs(deltax_m)>cutoff_2)=0;

average_1 = mean(rc_1(:));
min_1 = min(rc_1(:));
p5_1 = quantile(rc_1(:),0.05);
p95_1 = quantile(rc_1(:),0.95);
max_1 = max(rc_1(:));

average_2 = mean(rc_2(:));
min_2 = min(rc_2(:));
p5_2 = quantile(rc_2(:),0.05);
p95_2 = quantile(rc_2(:),0.95);
max_2 = max(rc_2(:));
```

```

average_m = mean(rc_m(:));
min_m = min(rc_m(:));
p5_m = quantile(rc_m(:),0.05);
p95_m = quantile(rc_m(:),0.95);
max_m = max(rc_m(:));

% B
rd_m = deltax_m;
rd_m(abs(deltax_m)<=cutoff_m)=0;

jumps = sum(rd_m~=0);
y = zeros(12,1);
[yyyy,mm,dd]= ymd_f('SPY.csv');
for i = 1:11
    y(i+1) = sum(ismember(yyyy,2006+i));
end
jumps_y = zeros(11,1);
for i = 1:11
    jumps_y(i) = sum(jumps(1,(sum(y(1:i))+1):sum(y(1:i+1))));
end
sum_jumps_y = sum(jumps_y);

% C
realized_beta_1 = sum((rc_m.*rc_1))./sum(rc_m.^2);
realized_beta_2 = sum((rc_m.*rc_2))./sum(rc_m.^2);

average_rb_1 = mean(realized_beta_1(:));
min_rb_1 = min(realized_beta_1(:));
p5_rb_1 = quantile(realized_beta_1(:),0.05);
p95_rb_1 = quantile(realized_beta_1(:),0.95);
max_rb_1 = max(realized_beta_1(:));

average_rb_2 = mean(realized_beta_2(:));
min_rb_2 = min(realized_beta_2(:));
p5_rb_2 = quantile(realized_beta_2(:),0.05);
```

```
p95_rb_2 = quantile(realized_beta_2(:),0.95);
max_rb_2 = max(realized_beta_2(:));

% D
f = figure(3);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_1,realized_beta_1)
datetick('x','yyyy');
legend('realized_beta');
title('realized betas of DIS with SPY');
box off; grid on;
xlabel('time');
ylabel('realized betas');
print(f,'-dpng','-r200','figures/3D1');
close(f);

f = figure(4);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_2,realized_beta_2)
datetick('x','yyyy');
legend('realized_beta');
title('realized betas of PG with SPY');
box off; grid on;
xlabel('time');
ylabel('realized betas');
print(f,'-dpng','-r200','figures/3D2');
close(f);

% E
e_1 = rc_1 - realized_beta_1.*rc_m;

e_2 = rc_2 - realized_beta_2.*rc_m;

rho = zeros(1,T);
for i = 1:T
rho(i) = corr(e_1(:,i),e_2(:,i));
```

```
end

average_rho = mean(rho(:));
min_rho = min(rho(:));
p5_rho = quantile(rho(:),0.05);
p95_rho = quantile(rho(:),0.95);
max_rho = max(rho(:));

% F
kn = 11;
M = 7;

J = 1000;

new_rho = boost_correlation(J,T,n,kn,M,rc_1,rc_2,rc_m);

CI_new_rho = ci_f(0.05,mean(new_rho),std(new_rho));

f = figure(5);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_m(1,:),CI_new_rho(1,:))
hold on
plot(dates_m(1,:),CI_new_rho(2,:))
plot(dates_m(1,:),rho)
hold off
datetick('x','yyyy');
legend('CI_lower','CI_upper','correlation');
title('confidence interval of correlations');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/3F');
close(f);

% G
zoom_start = find(dates_return_1==datenum(2008,10,1,16,0,0))/n;
```

```
zoom_end = find(dates_return_1==datenum(2008,10,31,16,0,0))/n;
rho_zoom = rho(zoom_start:zoom_end);
CI_new_rho_zoom = CI_new_rho(:,zoom_start:zoom_end);
dates_m_zoom = dates_m(1,zoom_start:zoom_end);

f = figure(6);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_m_zoom(1,:),CI_new_rho_zoom(1,:))
hold on
plot(dates_m_zoom(1,:),CI_new_rho_zoom(2,:))
plot(dates_m_zoom(1,:),rho_zoom)
hold off
datetick('x','mmdd');
legend('Rho_CI_lower','Rho_CI_higher','rho');
title('confidence interval of correlations in one month');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/3G');
close(f);

% H
p_1 = (length(find(CI_new_rho(1,:)>0)))/252;
p_2 = (length(find(CI_new_rho(2,:)<0)))/252;
p_0 = (T-p_1-p_2)/252;
```