MMAT Homework 2

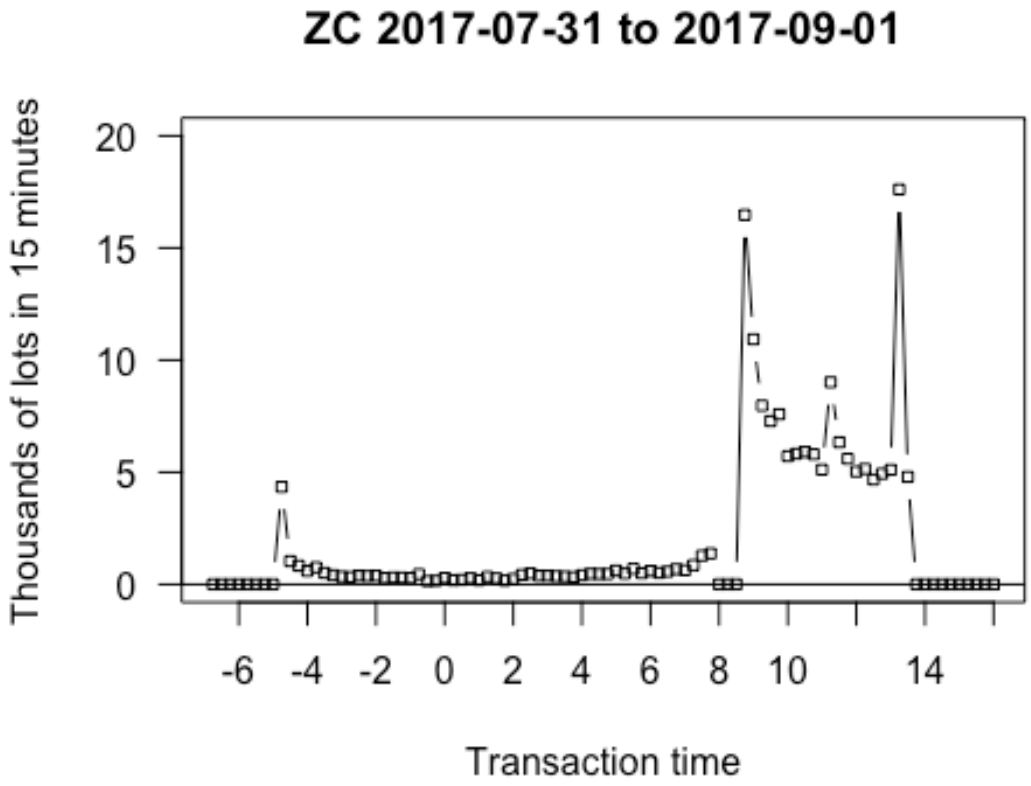
Jingyi Guo(jingyig1) , Jiawen Zhang(jiawenz2) Pittsburgh Campus

9/19/2017

1 Intraday Volume Profile

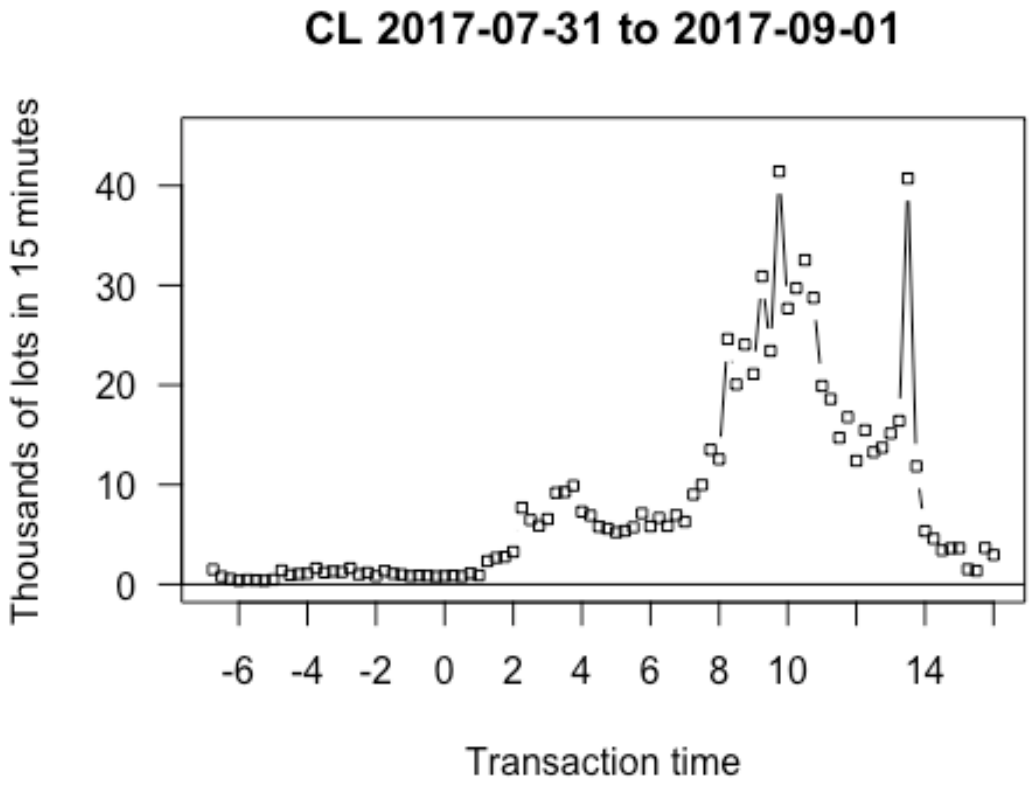
(1) AG Class: ZC (Corn)

```
library(rkdb)
h<-open_connection("172.19.10.167",7000)
data<-execute(h,"select v: sum siz by tbin:1+((`long$((rcvtime-rcvtime mod 10
00000000)+07:00:00))div(1000000000))div(15*60) from trade where sym2inst[sym]
=`ZC,date>=2017.07.31")
volume=matrix(0, nrow = 92, ncol = 1)
for (i in data$tbin){volume[i,1]=data$v[match(i,data$tbin)]}
volume_ave = volume/25/1000 #average number of lots per day
par(las=1)
plot(seq(-6.75,16,by=0.25),volume_ave,type='b',pch=0,cex=0.6,xaxt="n", ylim=c
(0, 20),ylab = "Thousands of lots in 15 minutes",xlab="Transaction time",main
= "ZC 2017-07-31 to 2017-09-01")
axis(1, at = seq(-6, 16, by = 2),las=1)
abline(a=0, b=0) #zero line
```



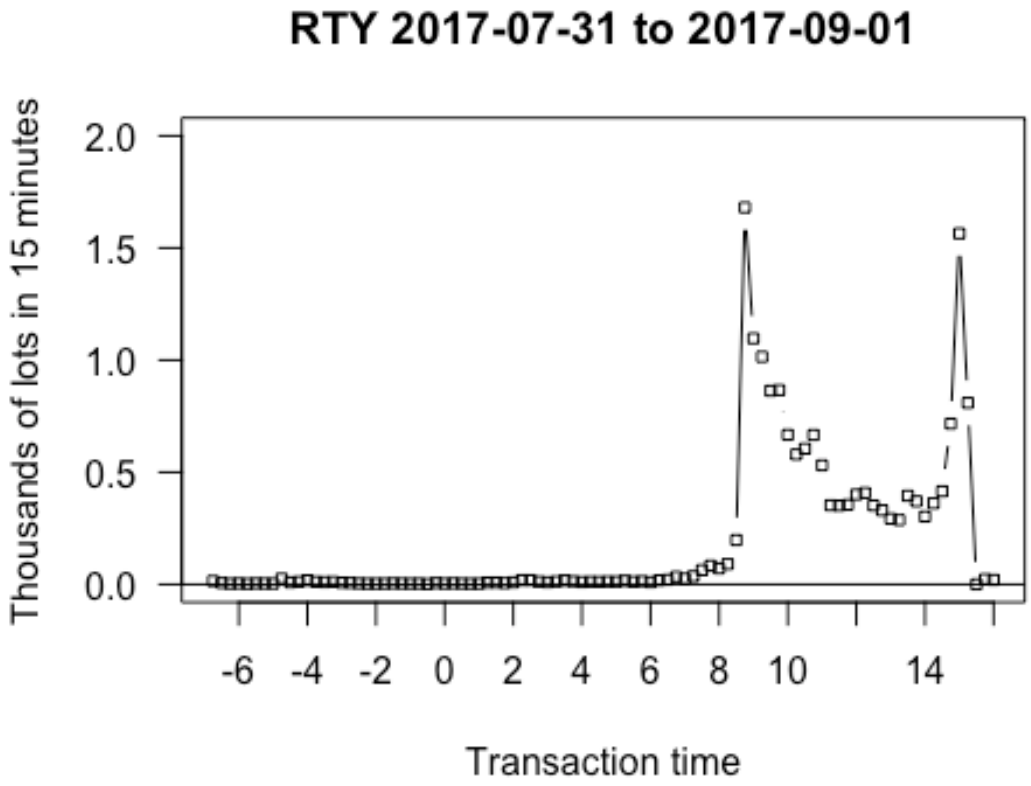
(2) EN Class: CL (Crude Oil)

```
data<-execute(h,"select v: sum siz by tbin:1+((`long$((rcvtime-rcvtime mod 10
00000000)+07:00:00))div(1000000000))div(15*60) from trade where sym2inst[sym]
=`CL,date>=2017.07.31")
volume=matrix(0, nrow = 92, ncol = 1)
for (i in data$tbin){volume[i,1]=data$v[match(i,data$tbin)]}
volume_ave = volume/25/1000 #average number of lots per day
par(las=1)
plot(seq(-6.75,16,by=0.25),volume_ave,type='b',pch=0,cex=0.6,xaxt="n", ylim=c
(0, 45),ylab = "Thousands of lots in 15 minutes",xlab="Transaction time",main
= "CL 2017-07-31 to 2017-09-01")
axis(1, at = seq(-6, 16, by = 2),las=1)
abline(a=0, b=0) #zero line
```



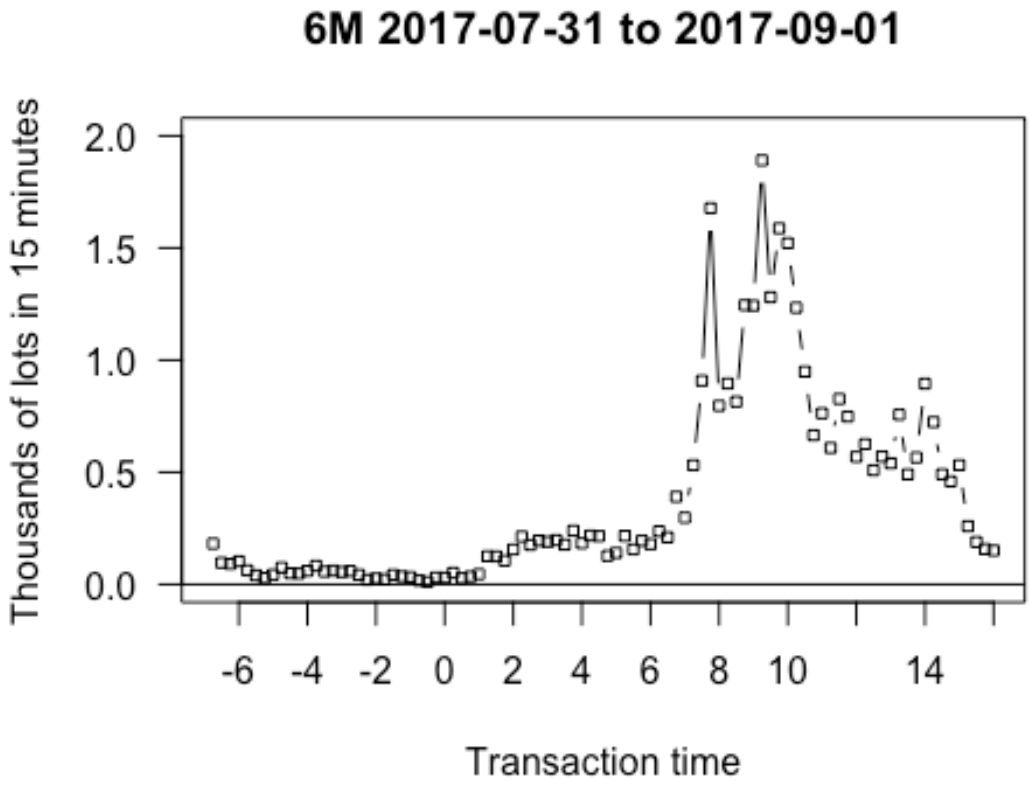
(3) EQ Class: RTY (R2000 Ind Mini)

```
data<-execute(h,"select v: sum siz by tbin:1+((`long$((rcvtime-rcvtime mod 10
00000000)+07:00:00))div(1000000000))div(15*60) from trade where sym2inst[sym]
=`RTY,date>=2017.07.31")
volume=matrix(0, nrow = 92, ncol = 1)
for (i in data$tbin){volume[i,1]=data$v[match(i,data$tbin)]}
volume_ave = volume/25/1000 #average number of lots per day
par(las=1)
plot(seq(-6.75,16,by=0.25),volume_ave,type='b',pch=0,cex=0.6,xaxt="n", ylim=c
(0, 2),ylab = "Thousands of lots in 15 minutes",xlab="Transaction time",main=
"RTY 2017-07-31 to 2017-09-01")
axis(1, at = seq(-6, 16, by = 2),las=1)
abline(a=0, b=0) #zero line
```



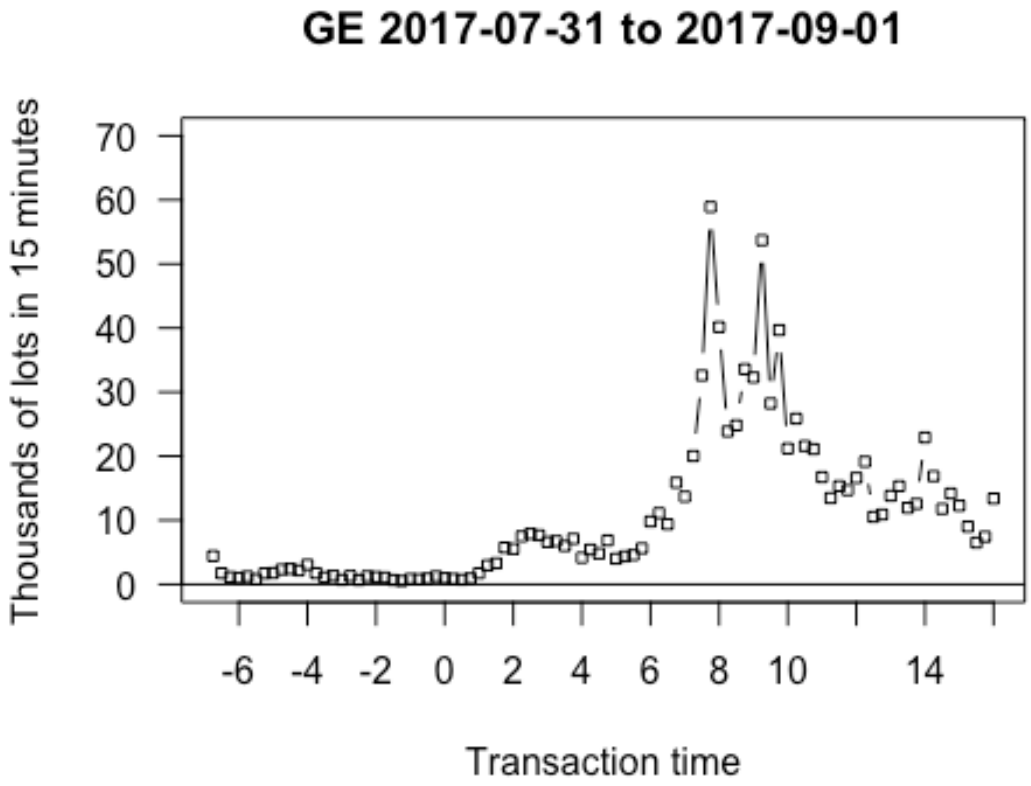
(4) FX Class: 6M (Mexican Peso)

```
data<-execute(h,"select v: sum siz by tbin:1+((`long$((rcvtime-rcvtime mod 10
00000000)+07:00:00))div(1000000000))div(15*60) from trade where sym2inst[sym]
=`6M,date>=2017.07.31")
volume=matrix(0, nrow = 92, ncol = 1)
for (i in data$tbin){volume[i,1]=data$v[match(i,data$tbin)]}
volume_ave = volume/25/1000 #average number of lots per day
par(las=1)
plot(seq(-6.75,16,by=0.25),volume_ave,type='b',pch=0,cex=0.6,xaxt="n", ylim=c
(0, 2),ylab = "Thousands of lots in 15 minutes",xlab="Transaction time",main=
"6M 2017-07-31 to 2017-09-01")
axis(1, at = seq(-6, 16, by = 2),las=1)
abline(a=0, b=0) #zero line
```



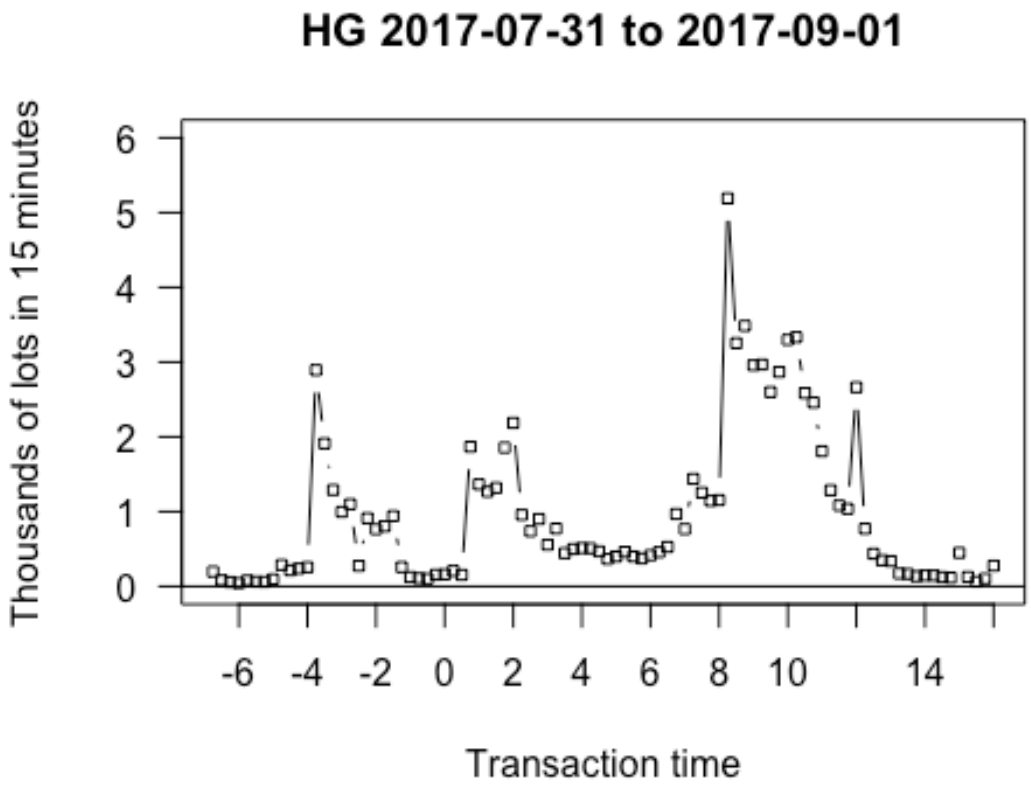
(5) IR Class: GE (Eurodollar)

```
h<-open_connection("172.19.10.167",6000)
data<-execute(h,"select v: sum siz by tbin:1+((`long$((rcvtime-rcvtime mod 10
00000000)+07:00:00))div(1000000000))div(15*60) from trade where sym2inst[sym]
=`GE,date>=2017.07.31")
volume=matrix(0, nrow = 92, ncol = 1)
for (i in data$tbin){volume[i,1]=data$v[match(i,data$tbin)]}
volume_ave = volume/25/1000 #average number of lots per day
par(las=1)
plot(seq(-6.75,16,by=0.25),volume_ave,type='b',pch=0,cex=0.6,xaxt="n", ylim=c
(0, 70),ylab = "Thousands of lots in 15 minutes",xlab="Transaction time",main
= "GE 2017-07-31 to 2017-09-01")
axis(1, at = seq(-6, 16, by = 2),las=1)
abline(a=0, b=0) #zero line
```



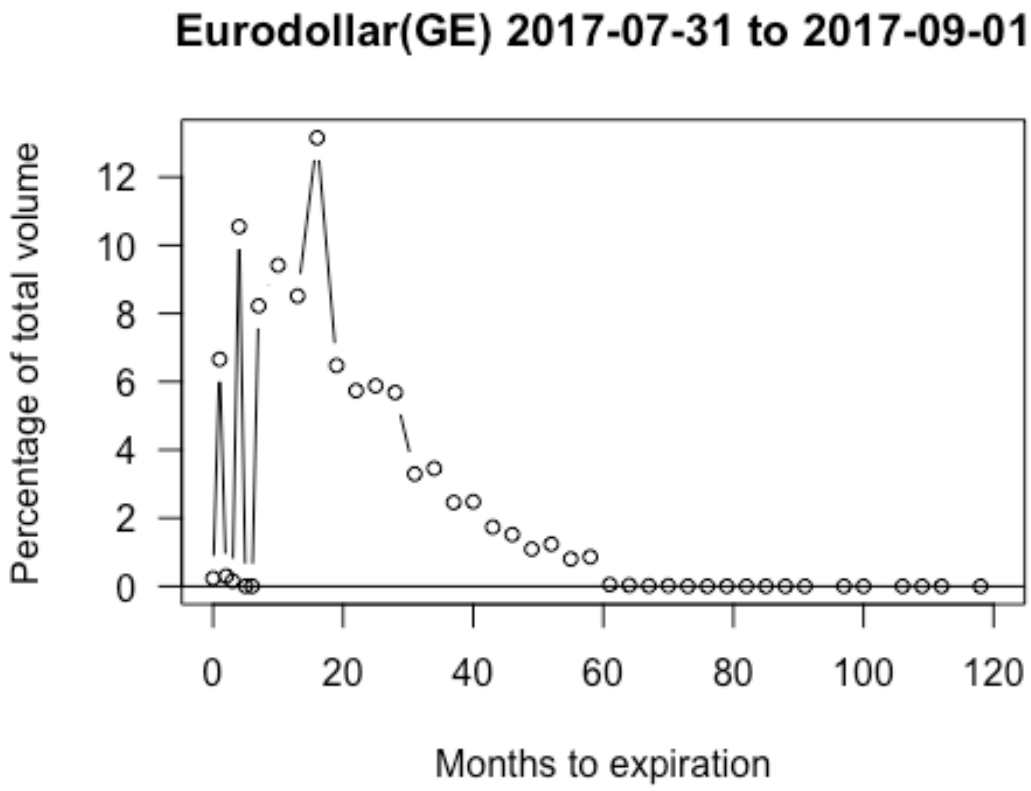
(6) MT Class: HG (Cooper)

```
h<-open_connection("172.19.10.167",7000)
data<-execute(h,"select v: sum siz by tbin:1+((`long$((rcvtime-rcvtime mod 10
00000000)+07:00:00))div(1000000000))div(15*60) from trade where sym2inst[sym]
=`HG,date>=2017.07.31")
volume=matrix(0, nrow = 92, ncol = 1)
for (i in data$tbin){volume[i,1]=data$v[match(i,data$tbin)]}
volume_ave = volume/25/1000 #average number of lots per day
par(las=1)
plot(seq(-6.75,16,by=0.25),volume_ave,type='b',pch=0,cex=0.6,xaxt="n", ylim=c
(0, 6),ylab = "Thousands of lots in 15 minutes",xlab="Transaction time",main=
"HG 2017-07-31 to 2017-09-01")
axis(1, at = seq(-6, 16, by = 2),las=1)
abline(a=0, b=0) #zero line
```



2 Volume Profile Across Maturities

```
h<-open_connection("172.19.10.167",6000)
volume_GE<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`GE,date>=2017.07.31")
total_vol_GE<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`GE,
date>=2017.07.31")
volume_GE$v = volume_GE$v/ total_vol_GE$v * 100 #percent
par(las=1)
plot(volume_GE,type='b',pch=1,cex=0.8, yaxt="n",xlim=c(0,120),ylab = "Percent
age of total volume",xlab="Months to expiration",main = "Eurodollar(GE) 2017-
07-31 to 2017-09-01")
axis(2, at = seq(0, 12, by = 2),las=1)
abline(a=0, b=0) #zero line
```



(b) Two most active interest rate instruments: ZN, GE

Five most active non-rate instruments: ES, CL, NQ, GC, 6E

```
entropy_GE=crossprod(-volume_GE$v/100,log(volume_GE$v/100))
N_eff_GE=exp(entropy_GE)

volume_ZN<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`GE,date>=2017.07.31")
total_vol_ZN<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`ZN,
date>=2017.07.31")
volume_ZN$v = volume_ZN$v/ total_vol_ZN$v
entropy_ZN=crossprod(-volume_ZN$v/100,log(volume_ZN$v/100))
N_eff_ZN=exp(entropy_ZN)

h<-open_connection("172.19.10.167",7000)
inst_data_non<-execute(h,"`v xdesc select v:sum siz by inst:sym2inst[sym] fro
m trade where date>=2017.07.31")
volume_ES<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`ES,date>=2017.07.31")
total_vol_ES<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`ES,
date>=2017.07.31")
volume_ES$v = volume_ES$v/ total_vol_ES$v * 100 #percent
```

```
entropy_ES=crossprod(-volume_ES$v/100,log(volume_ES$v/100))
N_eff_ES=exp(entropy_ES)

volume_CL<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`CL,date>=2017.07.31")
total_vol_CL<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`CL,
date>=2017.07.31")
volume_CL$v = volume_CL$v/ total_vol_CL$v * 100 #percent
entropy_CL=crossprod(-volume_CL$v/100,log(volume_CL$v/100))
N_eff_CL=exp(entropy_CL)

volume_NQ<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`NQ,date>=2017.07.31")
total_vol_NQ<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`NQ,
date>=2017.07.31")
volume_NQ$v = volume_NQ$v/ total_vol_NQ$v * 100 #percent
entropy_NQ=crossprod(-volume_NQ$v/100,log(volume_NQ$v/100))
N_eff_NQ=exp(entropy_NQ)

volume_GC<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`GC,date>=2017.07.31")
total_vol_GC<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`GC,
date>=2017.07.31")
volume_GC$v = volume_GC$v/ total_vol_GC$v * 100 #percent
entropy_GC=crossprod(-volume_GC$v/100,log(volume_GC$v/100))
N_eff_GC=exp(entropy_GC)

volume_6E<-execute(h,"select v:sum siz by (expir:`int$ sym2expir[sym] - 211)
from trade where sym2inst[sym]=`6E,date>=2017.07.31")
total_vol_6E<-execute(h,"select v:sum siz from trade where sym2inst[sym]=`6E,
date>=2017.07.31")
volume_6E$v = volume_6E$v/ total_vol_6E$v * 100 #percent
entropy_6E=crossprod(-volume_6E$v/100,log(volume_6E$v/100))
N_eff_6E=exp(entropy_6E)
```

So we have:

effective_number_6E = 1.11

effective_number_CL = 2.78

effective_number_ES = 1.02

effective_number_GC = 1.14

effective_number_GE = 16

effective_number_NQ = 1.03

effective_number_ZN = 1

Of them GE has “interesting” structure in that it has large effective numbers: more than a few maturities are active at one time.

3 Robert-Rosenbaum parameter correlated with the same size ratio

Choose dates range: 2017.08.07 to 2017.08.11. Consider the 43 instruments without missing data.

```
rr = matrix(0,48,3)
h<-open_connection("172.19.10.167",6000)
#find the most active symbol
symbol<-execute(h,"{[d;inst] 1#`sumsiz xdesc select sumsiz:sum siz by sym fro
m trade where date=d,sym2inst[sym]=inst}[2017.08.07;`GE]")
print (symbol$sym)

## [1] "GEZ8"

symbol<-execute(h,"{[d;inst] 1#`sumsiz xdesc select sumsiz:sum siz by sym fro
m trade where date=d,sym2inst[sym]=inst}[2017.08.08;`GE]")
print (symbol$sym)

## [1] "GEZ7"

symbol<-execute(h,"{[d;inst] 1#`sumsiz xdesc select sumsiz:sum siz by sym fro
m trade where date=d,sym2inst[sym]=inst}[2017.08.09;`GE]")
print (symbol$sym)

## [1] "GEZ8"

symbol<-execute(h,"{[d;inst] 1#`sumsiz xdesc select sumsiz:sum siz by sym fro
m trade where date=d,sym2inst[sym]=inst}[2017.08.10;`GE]")
print (symbol$sym)

## [1] "GEZ8"

symbol<-execute(h,"{[d;inst] 1#`sumsiz xdesc select sumsiz:sum siz by sym fro
m trade where date=d,sym2inst[sym]=inst}[2017.08.11;`GE]")
print (symbol$sym)

## [1] "GEZ8"

q_t = matrix(0,5,1)
eta = matrix(0,5,1)

#2017.08.07
# average quote size
ave_quote<-execute(h,"{[d;s] select avg qsiz by sym from aj[`seq;select sym,s
eq from trade where date=d, sym=s; select sprd:(last ask)-last bid, qsiz:(las
t asiz)+last bsiz by seq from quote where date=d,sym=s,ask>bid] where not nul
l sprd,sprd>0} [2017.08.07;`GEZ8]")
# average trade size:
ave_trade<-execute(h,"{[d;s] select tsiz: avg siz by sym from trade where dat
```

```
e=d,sym=s}[2017.08.07;`GEZ8]")
qsiz_tsiz = ave_quote$qsiz/ave_trade$tsiz
q_t[1,1]=qsiz_tsiz

#number of reversions and continuations
cont_rev <-execute(h,"{[d;s] select n: count i by dp from select dp: deltas(s
ignum deltas prc), seq, prc, siz, aggr from (update diff:differ prc from sele
ct from trade where date=d,sym=s) where diff=1}[2017.08.07;`GEZ8]")
sum_rev = 0
sum_con = 0
for (i in 1:4){
  if (cont_rev$dp[i]==-2|cont_rev$dp[i]==2){
    sum_rev=sum_rev+cont_rev$n[i]
    }
  else if (cont_rev$dp[i]==0){
    sum_con=cont_rev$n[i]
    }
}
eta[1,1]=sum_con/sum_rev

#2017.08.08
# average quote size
ave_quote<-execute(h,"{[d;s] select avg qsiz by sym from aj[`seq;select sym,s
eq from trade where date=d, sym=s; select sprd:(last ask)-last bid, qsiz:(las
t asiz)+last bsiz by seq from quote where date=d,sym=s,ask>bid] where not nul
l sprd,sprd>0} [2017.08.08;`GEZ7]")
# average trade size:
ave_trade<-execute(h,"{[d;s] select tsiz: avg siz by sym from trade where dat
e=d,sym=s}[2017.08.08;`GEZ7]")
qsiz_tsiz = ave_quote$qsiz/ave_trade$tsiz
q_t[2,1]=qsiz_tsiz

#number of reversions and continuations
cont_rev <-execute(h,"{[d;s] select n: count i by dp from select dp: deltas(s
ignum deltas prc), seq, prc, siz, aggr from (update diff:differ prc from sele
ct from trade where date=d,sym=s) where diff=1}[2017.08.08;`GEZ7]")

sum_rev = 0
sum_con = 0
for (i in 1:4){
  if (cont_rev$dp[i]==-2|cont_rev$dp[i]==2){
    sum_rev=sum_rev+cont_rev$n[i]
    }
  else if (cont_rev$dp[i]==0){
    sum_con=cont_rev$n[i]
    }
}
eta[2,1]=sum_con/sum_rev
```

```
#2017.08.09
# average quote size
ave_quote<-execute(h,"{[d;s] select avg qsiz by sym from aj[`seq;select sym,s
eq from trade where date=d, sym=s; select sprd:(last ask)-last bid, qsiz:(las
t asiz)+last bsiz by seq from quote where date=d,sym=s,ask>bid] where not nul
l sprd,sprd>0} [2017.08.09;`GEZ8]")

# average trade size:
ave_trade<-execute(h,"{[d;s] select tsiz: avg siz by sym from trade where dat
e=d,sym=s}[2017.08.09;`GEZ8]")
qsiz_tsiz = ave_quote$qsiz/ave_trade$tsiz
q_t[3,1]=qsiz_tsiz

#number of reversions and continuations
cont_rev <-execute(h,"{[d;s] select n: count i by dp from select dp: deltas(s
ignum deltas prc), seq, prc, siz, aggr from (update diff:differ prc from sele
ct from trade where date=d,sym=s) where diff=1}[2017.08.09;`GEZ8]")

sum_rev = 0
sum_con = 0
for (i in 1:4){
  if (cont_rev$dp[i]==-2|cont_rev$dp[i]==2){
    sum_rev=sum_rev+cont_rev$n[i]
    }
  else if (cont_rev$dp[i]==0){
    sum_con=cont_rev$n[i]
    }
}

#2017.08.10
# average quote size
ave_quote<-execute(h,"{[d;s] select avg qsiz by sym from aj[`seq;select sym,s
eq from trade where date=d, sym=s; select sprd:(last ask)-last bid, qsiz:(las
t asiz)+last bsiz by seq from quote where date=d,sym=s,ask>bid] where not nul
l sprd,sprd>0} [2017.08.10;`GEZ8]")

# average trade size:
ave_trade<-execute(h,"{[d;s] select tsiz: avg siz by sym from trade where dat
e=d,sym=s}[2017.08.10;`GEZ8]")
qsiz_tsiz = ave_quote$qsiz/ave_trade$tsiz
q_t[4,1]=qsiz_tsiz

#number of reversions and continuations
cont_rev <-execute(h,"{[d;s] select n: count i by dp from select dp: deltas(s
ignum deltas prc), seq, prc, siz, aggr from (update diff:differ prc from sele
ct from trade where date=d,sym=s) where diff=1}[2017.08.10;`GEZ8]")

sum_rev = 0
sum_con = 0
```

```
for (i in 1:4){
  if (cont_rev$dp[i]==-2|cont_rev$dp[i]==2){
    sum_rev=sum_rev+cont_rev$n[i]
    }
  else if (cont_rev$dp[i]==0){
    sum_con=cont_rev$n[i]
    }
}

#2017.08.11
# average quote size
ave_quote<-execute(h,"{[d;s] select avg qsiz by sym from aj[`seq;select sym,s
eq from trade where date=d, sym=s; select sprd:(last ask)-last bid, qsiz:(las
t asiz)+last bsiz by seq from quote where date=d,sym=s,ask>bid] where not nul
l sprd,sprd>0} [2017.08.11;`GEZ8]")

# average trade size:
ave_trade<-execute(h,"{[d;s] select tsiz: avg siz by sym from trade where dat
e=d,sym=s}[2017.08.11;`GEZ8]")
qsiz_tsiz = ave_quote$qsiz/ave_trade$tsiz
q_t[5,1]=qsiz_tsiz

#number of reversions and continuations
cont_rev <-execute(h,"{[d;s] select n: count i by dp from select dp: deltas(s
ignum deltas prc), seq, prc, siz, aggr from (update diff:differ prc from sele
ct from trade where date=d,sym=s) where diff=1}[2017.08.11;`GEZ8]")

sum_rev = 0
sum_con = 0
for (i in 1:4){
  if (cont_rev$dp[i]==-2|cont_rev$dp[i]==2){
    sum_rev=sum_rev+cont_rev$n[i]
    }
  else if (cont_rev$dp[i]==0){
    sum_con=cont_rev$n[i]
    }
}

eta[5,1]=sum_con/sum_rev
q_t_ave = mean(q_t)
eta_ave = mean(eta)
rr[1,1]="GE"
rr[1,2]=q_t_ave
rr[1,3]=eta_ave
```

Similar procedure for the other instruments (code omitted)... And we have the following plot:

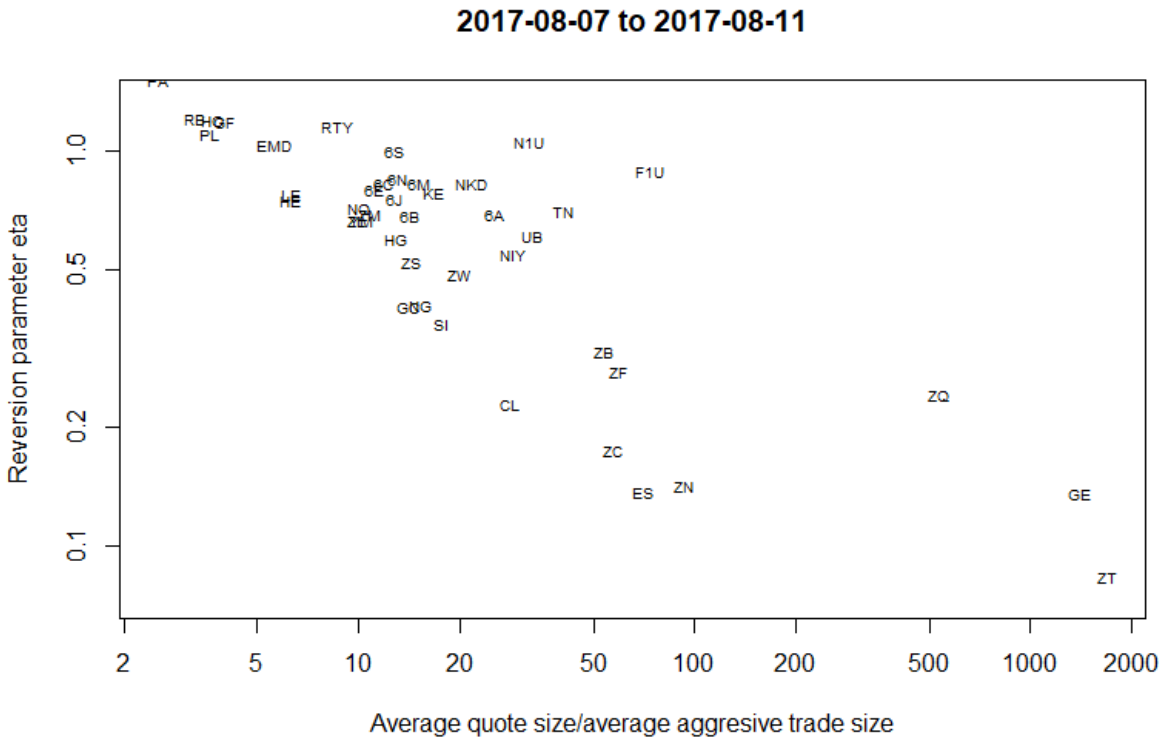
```
x = as.numeric(rr[c(1:3,6:8,10,12:38,40:48),2])
y = as.numeric(rr[c(1:3,6:8,10,12:38,40:48),3])
```

```
label = rr[c(1:3,6:8,10,12:38,40:48),1]
plot(x,y,log="xy",xlab = "Average quote size/average aggresive trade size", y
lab="Reversion parameter eta", main = "2017-08-07 to 2017-08-11",type="n")

## Warning in xy.coords(x, y, xlabel, ylabel, log): 42 x values <= 0 omitted
## from logarithmic plot

## Warning in xy.coords(x, y, xlabel, ylabel, log): 42 y values <= 0 omitted
## from logarithmic plot

text(x,y,labels=label, cex= 0.6, pos=3)
```



In the plot, ZQ, GE and ZR are on the lower right, because they have large *average quote size/average aggressive trade size* compared to the others. This is due to their special matching algorithm.

For GE, it uses Allocation(A). For ZQ and ZT, they use Split FIFO and Pro-rata(K). Those are all pro-rata algorithms, and pro-rata algo causes the average quote size to increase because traders need to post more size to get their desired size traded.