Tarea 2 - Medidas de liquidez

Jordi Vanrell Forteza

26/5/2021

La tarea requiere escoger aleatoriamente tres archivos de cada uno de los tres grupos provistos. Los grupos se corresponden con acciones de capitalización pequeña (nombrados *Stock20xLOB.txt*), media (*Stock10xLOB.txt*) y grande (*StockxLOB.txt*). Los ficheros contienen datos para tres semanas del libro de órdenes (LOB) de activos negociados en el NASDAQ. La primera semana es del 7 al 11 de abril de 2008, la segunda (la semana de la crisis) es la del 15 al 19 de septiembre y la tercera es la del 12 al 16 de febrero de 2010. Se descargan todos y se meten en un mismo directorio. Se seleccionan aleatoriamente tres números entre 1 y 5 que marcarán la elección de las acciones en los tres grupos. Se leen con un bucle y se almacenan los nueve en una lista.

# Se escogen 3 acciones de cada grupo (al azar)  
set.seed(603) # semilla para completa reproducibilidad  
idx <- as.character(sample(1:5, size = 3, replace = F))  
idx <- c(paste0("", idx), paste0("10", idx), paste0("20", idx))  
idx2 <- list(idx[1:3], idx[4:6], idx[7:9])  
  
LOB <- vector(mode = "list", length = length(idx))  
for (i in idx){ # Bucle de lectura de los archivos; se almacenan en lista  
 LOB[[which(idx==i)]] <- read.table(paste0("Stock", i, "LOB.txt"), header = T)  
 rm(i)  
}

Dicha lista consiste en 9 tablas de datos ordenadas según el índice de posición del vector idx: 5, 3, 1, 105, 103, 101, 205, 203 y 201.

# 1. Programario para el cálculo de las medidas de liquidez.

require(tidyverse) # Paquete para la manipulación de de data frames  
require(lubridate) # Paquete para la manipulación de fechas

En este punto se requiere el desarrollo de un programa:

1. para calcular la horquilla relativa media ponderada en puntos básicos para *round trips* de 100, 500 y 1000 acciones considerando la profundidad total (suma de la mostrada y la oculta).
2. para calcular el logaritmo de la profundidad total acumulada en euros en el lado *ask* del LOB a distancias de 1, 5, 10 y 20 *ticks* (0.01 $).

Se arma un bucle dentro del cual se procederá al cálculo de todas las magnitudes indicadas, por minuto y para cada uno de los activos. Las funciones internas calculan, por orden, el precio medio ponderado en el *ask* y *bid* y, con este, la horquilla media ponderada para cada *round trip,* y después la profundidad acumulada del *ask* para cada distancia. Los datos para las horquillas y las profundidades se acumulan separadamente en dos listas, con una tabla por activo en cada una de ellas.

# Subtabla 1 Horquilla, tabla\_i\_ext (9 subtablas)  
tabla\_i\_ext <- vector(mode = "list", length = length(LOB))  
# Subtabla 2 Profundidad, tabla\_i\_ext2 (9 subtablas)  
tabla\_i\_ext2 <- vector(mode = "list", length = length(LOB))  
  
for (d in 1:length(LOB)){  
 LOB[[d]] <- LOB[[d]] %>%  
 mutate(vol = hvol + dvol, # se suman el volumen mostrado y el oculto  
 datime = ymd(date) + hms("09:30:00") + time - 34200) # fecha y hora  
 minute <- unique(LOB[[d]]$datime) # se extrae un vector con los minutos  
 subtabla\_i\_ext <- data.frame() # inicializa data.frames depositarios  
 subtabla\_i\_ext2 <- data.frame()  
 for (t in 1:length(minute)){  
 LOB\_q <- LOB[[d]] %>%  
 filter(datime == minute[t]) # filtra LOB según minuto  
 # Cálculo de los puntos medios:  
 q <- (LOB\_q$quote[LOB\_q$sign == 1] + LOB\_q$quote[LOB\_q$sign == -1])/2  
 ab <- function(part, rt\_amt){  
 # Función de cálculo del precio medio ponderado ask/bid  
 # INPUTS:  
 # part debe tomar valores -1 (ask) o 1 (bid)  
 # rt\_amt es la cantidad de acciones del round trip.  
 # OUTPUTS:  
 # q\_rt: el precio medio ask/bid del roundtrip.  
 LOB\_partmin <- LOB[[d]] %>%  
 # filtra por minuto y signo de la transacción  
 filter(datime == minute[t] & sign\*part > 0) %>%  
 arrange(abs(sign)) # ordena según prioridad de las operaciones  
 vol <- LOB\_partmin$vol # se vectorizan los volúmenes...  
 quote <- LOB\_partmin$quote # ...y los precios  
 # inicializa un vector para almacenar los volúmenes del round-trip:  
 vol\_rt <- c()   
 for (i in 1:length(vol)){  
 # El bucle almacena en el vector anterior la cantidad de acciones  
 # del roundtrip a cada precio diferente. El 0 en la func max  
 # asegura que no tomará valores negativos cuando el roundtrip se  
 # haya completado.  
 vol\_rt[i] <- max(min(rt\_amt-sum(vol[0:(i-1)]), vol[i]), 0)  
 }  
 # Si el roundtrip no se ha completado devuelve NA.  
 q\_rt <- ifelse(sum(vol\_rt) < rt\_amt, NA, sum(vol\_rt\*quote)/sum(vol\_rt))  
 return(q\_rt)  
 }  
 s <- function(rt\_amt){  
 # Función de cálculo de la horquilla relativa media ponderada  
 # INPUTS:  
 # rt\_amt: ídem que en función ab  
 # OUTPUTS: horquilla relativa media ponderada en puntos básicos  
 (log(ab(-1, rt\_amt)) - log(ab(1, rt\_amt)))\*10000  
 }  
 a\_tick <- function(part, ticks, tick){  
 # Función de cálculo de la profundidad:  
 # INPUTS:  
 # part ídem que en func ab; para ask es -1  
 # ticks: número de ticks  
 # tick: magnitud de cada tick (0.01 $)  
 # OUTPUT:  
 # logaritmo de las unidades monetarias totales.  
 LOB\_partmin <- LOB[[d]] %>%  
 filter(datime == minute[t] & sign\*part > 0) %>%  
 arrange(abs(sign))  
 vol <- LOB\_partmin$vol  
 quote <- LOB\_partmin$quote # 5 líneas son análogas en la func ab.  
 vol\_tick <- c()   
 for (i in 1:length(vol)){  
 # El bucle busca todos precios menores o iguales a q+(tick\*ticks)  
 # y devuelve volúmenes asociados a estos. En caso de no existir  
 # ningún volumen devuelve 0.  
 vol\_tick[i] <- ifelse(quote[i] <= q+(tick\*ticks), vol[i], 0)  
 }  
 # La función devuelve el log de las unidades monetarias totales,  
 # resultado de multiplicar volumen de acciones hasta ese precio por  
 # sus respectivos precios. El 0 en la func max censura los -Inf que   
 # resultan de aplicar log cuando no existe volumen a ese tick.  
 return(max(log(sum(vol\_tick\*quote)), 0))  
 }  
 # en cell3 se almacenan resultados de aplicar la función s de cálculo  
 # de la horquilla con los parámetros deseados.  
 cell3 <- data.frame(rt\_100 = s(100), rt\_500 = s(500), rt\_1000 = s(1000))  
 # Acumula resultados por minuto:  
 subtabla\_i\_ext <- rbind(subtabla\_i\_ext, cell3)  
 # En cell4 se almacenan resultados de la función a\_tick de cálculo  
 # de la profundidad con los parámetros deseados.  
 cell4 <- data.frame(tick\_1 = a\_tick(-1, 1, .01),  
 tick\_5 = a\_tick(-1, 5, .01),  
 tick\_10 = a\_tick(-1, 10, .01),  
 tick\_20 = a\_tick(-1, 20, .01))  
 # Acumula los resultados por minuto:  
 subtabla\_i\_ext2 <- rbind(subtabla\_i\_ext2, cell4)  
 }  
 # Se almacenan tablas resultados en lista, convenientemente renombradas  
 # las filas.  
 tabla\_i\_ext[[d]] <- subtabla\_i\_ext  
 rownames(tabla\_i\_ext[[d]]) <- minute  
 tabla\_i\_ext2[[d]] <- subtabla\_i\_ext2  
 rownames(tabla\_i\_ext2[[d]]) <- minute  
 cat(paste0("Stock",idx[d]," completado. ")) # Muestra prog en pantalla.  
 rm(d, t, ab, a\_tick, s, q, minute, cell3, cell4,   
 subtabla\_i\_ext, subtabla\_i\_ext2, LOB\_q)  
}

Tras efectuar los cálculos se han observado algunos valores negativos (o ceros) en las horquillas de algunos minutos. Buscando los datos originales se ha observado que estas se corresponden con momentos puntuales en las que el mejor *ask* se situaba por debajo del mejor *bid* (o al mismo nivel). Estas representan situaciones transitorias anómalas en las que no se había consolidado el LOB del toto y, por ello, se prescinde de dichas observaciones (en total 4 observaciones entre los 9 activos).

for (i in 1:length(tabla\_i\_ext)){  
 # Con este bucle se limpia tabla tabla\_i\_ext de todos los registros  
 # que contengan algún valor negativo de horquilla. Se conservan NA.  
 tabla\_i\_ext[[i]] <- tabla\_i\_ext[[i]] %>%   
 filter((rt\_100 > 0 & rt\_500 > 0 & rt\_1000 > 0) |   
 (rt\_100 > 0 & rt\_500 > 0 & is.na(rt\_1000)) |   
 (rt\_100 > 0 & is.na(rt\_500) & rt\_1000 > 0) |   
 (is.na(rt\_100) & rt\_500 > 0 & rt\_1000 > 0) |  
 (rt\_100 > 0 & is.na(rt\_500) & is.na(rt\_1000)) |  
 (is.na(rt\_100) & rt\_500 > 0 & is.na(rt\_1000)) |  
 (is.na(rt\_100) & is.na(rt\_500) & rt\_1000 > 0) |  
 (is.na(rt\_100) & is.na(rt\_500) & is.na(rt\_1000)))  
 rm(i)  
}

# 2. Medias de las medidas de liquidez y significatividad de las diferencias

En este apartado se piden:

* Los promedios de las medias diarias de las medidas de liquidez calculadas en 1 desglosadas por activo y cruzadas por submuestras según grado de capitalización bursátil.
* La realización de contrastes estadísticos que permitan valorar la significatividad de las diferencias diarias entre submuestras según capitalización.
* Análisis de las diferencias en las medidas entre la semana de la crisis y las otras dos.

## 2.1. Promedios de las medidas de liquidez por activo y submuestras.

En primera instancia, y siguiendo las indicaciones del enunciado, a través de un bucle, se hacen las medias diarias de las medidas de liquidez calculadas en 1. Para ello, para cada par de tablas correspondientes a un mismo activo, se genera una variable para la fecha en formato día/mes/año, se agrupan los datos por fecha y se calculan las medias. Los resultados se almacenan en un par de listas análogas a las del apartado 1.

Posteriormente se construye una función que promedia las medias diarias de las medidas de liquidez tomando como parámetros la primera y la última posición de los activos que pretenden agregarse tal como se ordenan en el vector idx. Luego se aplica la función en bucle de forma que, al acabar con todos los activos de una submuestra, calcula los valores para la submuestra en conjunto.

# En estas listas se almacenan las medias diarias de horquillas y profundidades.  
tabla\_i\_sum <- vector(mode = "list", length = length(LOB))  
tabla\_i\_sum2 <- vector(mode = "list", length = length(LOB))  
  
for (i in 1:length(LOB)){  
 # Se genera la variable date (fecha) para cada una de los activos.  
 tabla\_i\_ext[[i]]$date <- as.Date(rownames(tabla\_i\_ext[[i]]))  
 tabla\_i\_ext2[[i]]$date <- as.Date(rownames(tabla\_i\_ext2[[i]]))  
 tabla\_i\_sum[[i]] <- tabla\_i\_ext[[i]] %>%  
 group\_by(date) %>% # agrupa por fecha (día) y genera las medias  
 summarise(mean\_s\_rt\_100 = mean(rt\_100, na.rm = T),  
 mean\_s\_rt\_500 = mean(rt\_500, na.rm = T),  
 mean\_s\_rt\_1000 = mean(rt\_1000, na.rm = T))  
 tabla\_i\_sum2[[i]] <- tabla\_i\_ext2[[i]] %>%  
 group\_by(date) %>% # idem.  
 summarise(mean\_tick\_1 = mean(tick\_1, na.rm = T),  
 mean\_tick\_5 = mean(tick\_5, na.rm = T),  
 mean\_tick\_10 = mean(tick\_10, na.rm = T),  
 mean\_tick\_20 = mean(tick\_20, na.rm = T))  
 rm(i)  
}  
  
summarizer <- function(first, last){  
 # Función que resume la información de las medias diarias:  
 # INPUTS:  
 # first: orden del primer activo a considerar  
 # last: orden del último activo a considerar  
 # Cuando first == last recoge solamente el de ese activo.  
 # OUTPUTS:  
 # sum: data.frame con las medias de las medidas de liquidez  
 sum <- data.frame()  
 for (i in first:last){  
 summary <- merge(tabla\_i\_sum[[i]], tabla\_i\_sum2[[i]])  
 sum <- rbind(sum, summary)  
 }  
 sum <- sum %>% # Se resume la información de los activos considerados  
 summarise(mean\_s\_rt\_100 = round(mean(mean\_s\_rt\_100), 3),  
 mean\_s\_rt\_500 = round(mean(mean\_s\_rt\_500), 3),  
 mean\_s\_rt\_1000 = round(mean(mean\_s\_rt\_1000), 3),  
 mean\_tick\_1 = round(mean(mean\_tick\_1), 3),  
 mean\_tick\_5 = round(mean(mean\_tick\_5), 3),  
 mean\_tick\_10 = round(mean(mean\_tick\_10), 3),  
 mean\_tick\_20 = round(mean(mean\_tick\_20), 3))  
 rownames(sum) <- ifelse(first == last, paste0("Stock",idx[first]),   
 paste0("Stocks",idx[first],"-",idx[last]))  
 return(sum)  
}  
  
tabla\_i <- data.frame()  
stops <- c(3, 6, 9)  
for (i in 1:length(LOB)){ # Bucle para generar tabla\_i  
 tabla\_i <- rbind(tabla\_i, summarizer(i,i))  
 if(i %in% stops){   
 # Tras los activos #3, #6, #9 (últimos de su grado de capitalización)  
 # el bucle genera, además, los estadísticos de los activos agregados  
 # (por grado de capitalización)  
 tabla\_i <- rbind(tabla\_i, summarizer(i-2,i))  
 }  
 rm(i)  
}

En la Tabla I se disponen los promedios tal como se han especificado. En las filas se disponen los activos y sus agregaciones por submuestras, en las columnas se listan las medidas de liquidez.

# Por cuestiones de compilación la tabla (esta y las siguientes) se han generado de forma separada, de acuerdo con el código que sigue. Este es un código que NO puede compilarse en Rmd. Con posterioridad se han editado en Word para hacerlas más legibles.  
library(officer)  
library(flextable)  
library(magrittr)  
# Para exportar tablas a formato Word  
ft <- flextable(data = tabla\_i %>% add\_rownames()) %>%   
 theme\_zebra %>% autofit  
ft <- set\_caption(ft, caption = "Tabla I", style = "Table Caption")  
# Crea un archivo temp  
tmp <- tempfile(fileext = ".docx")  
# Crea un documento docx  
read\_docx() %>% body\_add\_flextable(ft) %>% print(target = tmp)  
browseURL(tmp) # abre el documento

**Tabla I**

|  | **Horquillas** | | | **Profundidad** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **100 ac.** | **500 ac.** | **1000 ac.** | **1 tick** | **5 ticks** | **10 ticks** | **20 ticks** |
| Stock5 | 6.518 | 6.608 | 6.730 | 11.958 | 14.017 | 14.496 | 14.504 |
| Stock3 | 3.524 | 4.390 | 5.621 | 9.675 | 12.751 | 13.452 | 13.513 |
| Stock1 | 5.724 | 6.493 | 7.571 | 10.119 | 12.707 | 13.245 | 13.297 |
| **Stocks5-1** | **5.255** | **5.830** | **6.641** | **10.584** | **13.158** | **13.731** | **13.772** |
| Stock105 | 10.648 | 14.928 | 20.500 | 8.976 | 12.107 | 12.646 | 12.909 |
| Stock103 | 12.401 | 15.578 | 22.545 | 8.481 | 10.962 | 11.572 | 11.767 |
| Stock101 | 11.139 | 15.348 | 21.992 | 7.949 | 11.072 | 11.662 | 11.974 |
| **Stocks105-101** | **11.396** | **15.284** | **21.679** | **8.469** | **11.380** | **11.960** | **12.217** |
| Stock205 | 32.163 | 63.231 | 101.849 | 2.584 | 7.972 | 9.279 | 10.242 |
| Stock203 | 19.143 | 33.918 | 52.556 | 4.152 | 9.616 | 10.454 | 10.929 |
| Stock201 | 40.228 | 90.244 | 154.763 | 3.569 | 8.637 | 9.354 | 10.142 |
| **Stocks205-201** | **30.511** | **62.464** | **103.056** | **3.435** | **8.742** | **9.696** | **10.438** |

De la Tabla I se desprende que las horquillas son tanto mayores cuanto más pequeño es el grado de capitalización y que la relación es en sentido contrario con respecto de la profundidad. Estas dos tendencias se traducen en grados de liquidez mayores cuanto más grande es el grado de capitalización bursátil.

## 2.2. Significatividad de las diferencias entre submuestras

En segundo lugar, se pide la realización de contrastes de significatividad de las diferencias en los datos de medias diarias de las medidas de liquidez entre las submuestras de activos según el grado de capitalización. Antes de efectuar los contrastes, por tanto, es necesario sacar las medias diarias de las medidas de liquidez para cada submuestra. Los datos se almacenan en una lista de 3 tablas, una por submuestra.

# Se almacenan en lista los estadísticos de cada submuestra.  
tabla\_i\_cs <- vector(mode = "list", length = length(LOB)/length(stops))  
  
for (i in 1:length(idx2)){  
 agg <- data.frame()  
 agg2 <- data.frame()  
 for (j in 1:length(idx2[[i]])){  
 # El bucle agrega según numeración de subconjuntos de idx2  
 agg <- rbind(agg, tabla\_i\_sum[[which(idx == idx2[[i]][j])]])  
 agg2 <- rbind(agg2, tabla\_i\_sum2[[which(idx == idx2[[i]][j])]])  
 }  
 agg <- agg %>%  
 group\_by(date) %>%  
 # una vez agregados los subconjuntos agrupa por fecha (día) y calcula  
 # las medias.  
 summarise(mean\_s\_rt\_100 = mean(mean\_s\_rt\_100),  
 mean\_s\_rt\_500 = mean(mean\_s\_rt\_500),  
 mean\_s\_rt\_1000 = mean(mean\_s\_rt\_1000))  
 agg2 <- agg2 %>% # idem  
 group\_by(date) %>%  
 summarise(mean\_tick\_1 = mean(mean\_tick\_1),  
 mean\_tick\_5 = mean(mean\_tick\_5),  
 mean\_tick\_10 = mean(mean\_tick\_10),  
 mean\_tick\_20 = mean(mean\_tick\_20))  
 # se juntan las tablas de horquillas y profundidad:  
 tabla\_i\_cs[[i]] <- merge(agg, agg2)  
 rm(i, j, agg, agg2)  
}

Como hay activos con tres grados de capitalización se aplica el *rank-sum test de Wilcoxon* por pares (es decir, con todas las combinaciones posibles de 2 submuestras). Esta prueba contrasta la hipótesis nula de que las medianas de dos distribuciones de valores son iguales contra la alternativa, que postula lo contrario. Se toma el nivel de significación del 5%, con lo que un p-valor del contraste por debajo de 0.05 implica que se rechaza la hipótesis nula de igualdad de medianas.

var <- colnames(tabla\_i\_cs[[1]])[2:8]  
# 3 combinaciones posibles de los 3 subconjuntos:  
pairs <- list(c(1, 2), c(1, 3), c(2, 3))  
cap <- c("Gran", "Med", "Peq")  
  
# inicializa data.frame con columna de 7 NAs (7 es el número de medidas)  
# Solo es necesaria para dar dimensión al data.frame.  
Wilcox\_p <- data.frame(rep(NA, 7))   
for (i in 1:length(pairs)){  
 Wilcox\_sub <- data.frame()  
 for (j in 1:length(var)){  
 # Para cada combinación por pares de los 3 subconjuntos calcula  
 # p-valor según la prueba de Wilcoxon.  
 x <- tabla\_i\_cs[[pairs[[i]][1]]][, var[j]]  
 y <- tabla\_i\_cs[[pairs[[i]][2]]][, var[j]]  
 pvalue <- wilcox.test(x, y, alternative = "two.sided", paired = F)$p.value  
 pvalue <- data.frame(pvalue) # lo mete en data.frame.  
 Wilcox\_sub <- rbind(Wilcox\_sub, pvalue) # agrega los valores  
 }  
 # Bautiza la combinación de subconjuntos:  
 colnames(Wilcox\_sub) <- paste0("Cap-", cap[pairs[[i]][1]], "-", cap[pairs[[i]][2]])  
 Wilcox\_p <- cbind(Wilcox\_p, Wilcox\_sub) # guarda los resultados.  
 rm(i, j, x, y, Wilcox\_sub, pvalue)  
}  
Wilcox\_p <- Wilcox\_p[,2:4] # descarta la columna accesoria de NAs   
# Se nombran filas para conservar la correspondencia con los p-valores.  
rownames(Wilcox\_p) <- var

Los resultados (p-valores) se muestran en la tabla siguiente:

# Para exportar tablas a formato Word  
ft <- flextable(data = Wilcox\_p %>% add\_rownames()) %>%   
 theme\_zebra %>% autofit  
ft <- set\_caption(ft, caption = "Tabla P – p-valores del contraste de Wilcoxon sobre las medidas de liquidez según grado de capitalización",   
 style = "Table Caption")  
# Crea un archivo temp  
tmp <- tempfile(fileext = ".docx")  
# Crea un documento docx  
read\_docx() %>% body\_add\_flextable(ft) %>% print(target = tmp)  
browseURL(tmp) # abre el documento

**Tabla P - p-valores del contraste de Wilcoxon sobre las medidas de liquidez según grado de capitalización.**

|  | **Gran vs Med** | **Gran vs Peq** | **Med vs Peq** |
| --- | --- | --- | --- |
| Horq. 100 acc. | 1.29·10-8 | 1.29·10-8 | 1.29·10-8 |
| Horq. 500 acc. | 1.29·10-8 | 1.29·10-8 | 1.29·10-8 |
| Horq. 1000 acc. | 5.8·10-7 | 1.29·10-8 | 1.29·10-8 |
| Prof. 1 tick | 8.64·10-3 | 1.29·10-8 | 1.29·10-8 |
| Prof. 5 ticks | 1.19·10-3 | 1.29·10-8 | 1.29·10-8 |
| Prof. 10 ticks | 2.64·10-4 | 1.29·10-8 | 1.29·10-8 |
| Prof. 20 ticks | 3.23·10-4 | 1.29·10-8 | 2.57·10-8 |

Todos los p-valores de la tabla P son menores que 0.05, por lo que se rechaza la hipótesis nula de igualdad de medianas. Por lo tanto, puede decirse que las tres submuestras de activos son distintas entre sí.

## 2.3. ¿Cómo afectó a la liquidez la semana de la crisis?

Igual que en 2.2. se contrastó la igualdad de medianas entre grados de capitalización sin discernir semana, puede contrastarse también el efecto que la crisis tuvo sobre las medidas de liquidez. Como paso previo se reorganizan las observaciones diarias según las dimensiones semana y capitalización (9 grupos) y luego se contrasta la igualdad de medianas para cada submuestra según capitalización.

tabla\_i\_cs\_week <- matrix(vector(mode = "list", length = 1), nrow = 3, ncol = 3)  
# Esta vez se acumulan las tablas en una "matriz de listas de tablas"  
# En las filas coloca los activos por capitalización  
# En las columnas coloca los activos por semana  
for (i in 1:length(tabla\_i\_cs)){  
 # Se genera una variable para conocer la semana del registro.  
 # Es conveniente usar isoweek para semanas lunes-domingo.  
 # Se agrega el año al final; 38-2008 es la semana 38 de 2008.  
 tabla\_i\_cs[[i]]$week <- paste0(isoweek(tabla\_i\_cs[[i]]$date), "-", year(tabla\_i\_cs[[i]]$date))  
 weeks <- unique(tabla\_i\_cs[[i]]$week)  
 for (w in 1:length(weeks)){  
 # Convierte los 3 data.frames de las medias de los subconjuntos  
 # en 9, dividiéndolos según semanas.  
 group\_week <- tabla\_i\_cs[[i]] %>%  
 filter(week == weeks[w])  
 tabla\_i\_cs\_week[i,w][[1]] <- group\_week  
 }  
 rm(i, w, group\_week)  
}  
  
tabla\_i\_week\_wilcox <- vector(mode = "list", length = length(cap))  
  
for (c in 1:length(cap)){ # Por cada grado de capitalización...  
 subtabla\_i\_week\_wilcox <- data.frame(rep(NA, 7))  
 for (i in 1:length(pairs)){ # Por cada una de las parejas de semanas...  
 Wilcox\_sub <- data.frame()  
 for (j in 1:length(var)){ # Por cada variable...  
 # Análogo al bucle ante-anterior  
 x <- tabla\_i\_cs\_week[c, pairs[[i]][1]][[1]][, var[j]]  
 y <- tabla\_i\_cs\_week[c, pairs[[i]][2]][[1]][, var[j]]  
 pvalue <- wilcox.test(x, y, alternative = "two.sided", paired = F)$p.value  
 pvalue <- data.frame(pvalue)  
 Wilcox\_sub <- rbind(Wilcox\_sub, pvalue)  
 }  
 colnames(Wilcox\_sub) <- paste0(weeks[pairs[[i]][1]], "\_vs\_", weeks[pairs[[i]][2]])  
 subtabla\_i\_week\_wilcox <- cbind(subtabla\_i\_week\_wilcox, Wilcox\_sub)  
 }  
 subtabla\_i\_week\_wilcox <- subtabla\_i\_week\_wilcox[,2:4]  
 rownames(subtabla\_i\_week\_wilcox) <- var  
 tabla\_i\_week\_wilcox[[c]] <- subtabla\_i\_week\_wilcox  
 rm(c, i, j, x, y, pvalue, subtabla\_i\_week\_wilcox, Wilcox\_sub)  
}

Los resultados (p-valores) se muestran en las tablas siguientes:

# Para exportar tablas a formato Word  
ft <- flextable(data = tabla\_i\_week\_wilcox[[1]] %>% add\_rownames()) %>%   
 theme\_zebra %>% autofit  
ft <- set\_caption(ft, caption = "Tabla Q.1 - Activos de Gran Capitalización",   
 style = "Table Caption")  
# Crea un archivo temp  
tmp <- tempfile(fileext = ".docx")  
# Crea un documento docx  
read\_docx() %>% body\_add\_flextable(ft) %>% print(target = tmp)  
browseURL(tmp) # abre el documento

**Tabla Q.1. – Activos de Gran Capitalización**

|  | **15/2008 vs 38/2008** | **15/2008 vs 8/2010** | **38/2008 vs 8/2010** |
| --- | --- | --- | --- |
| Horq. 100 ac. | 0.0079 | 0.0079 | 0.0079 |
| Horq. 500 ac. | 0.0079 | 0.0079 | 0.0952 |
| Horq. 1000 ac. | 0.0079 | 0.0079 | 0.0079 |
| Prof. 1 tick | 0.0079 | 0.0079 | 0.0079 |
| Prof. 5 ticks | 0.0079 | 0.0079 | 0.0079 |
| Prof. 10 ticks | 0.0079 | 0.0079 | 0.0079 |
| Prof. 20 ticks | 0.0079 | 0.0079 | 0.0079 |

# Para exportar tablas a formato Word  
ft <- flextable(data = tabla\_i\_week\_wilcox[[2]] %>% add\_rownames()) %>%   
 theme\_zebra %>% autofit  
ft <- set\_caption(ft, caption = "Tabla Q.2 - Act. de Capitalización media",   
 style = "Table Caption")  
# Crea un archivo temp  
tmp <- tempfile(fileext = ".docx")  
# Crea un documento docx  
read\_docx() %>% body\_add\_flextable(ft) %>% print(target = tmp)  
browseURL(tmp) # abre el documento

**Tabla Q.2. – Activos de Capitalización media**

|  | **15/2008 vs 38/2008** | **15/2008 vs 8/2010** | **38/2008 vs 8/2010** |
| --- | --- | --- | --- |
| Horq. 100 ac. | 0.0079 | 0.0079 | 0.0079 |
| Horq. 500 ac. | 0.0079 | 0.0079 | 0.0079 |
| Horq. 1000 ac. | 0.0079 | 0.0079 | 0.0079 |
| Prof. 1 tick | 0.0079 | 0.0079 | 0.0079 |
| Prof. 5 ticks | 0.0079 | 0.0079 | 0.0079 |
| Prof. 10 ticks | 0.0079 | 0.0079 | 0.0079 |
| Prof. 20 ticks | 0.0079 | 0.0079 | 0.0079 |

# Para exportar tablas a formato Word  
ft <- flextable(data = tabla\_i\_week\_wilcox[[3]] %>% add\_rownames()) %>%   
 theme\_zebra %>% autofit  
ft <- set\_caption(ft, caption = "Tabla Q.3 - Act. de Capitalización pequeña",   
 style = "Table Caption")  
# Crea un archivo temp  
tmp <- tempfile(fileext = ".docx")  
# Crea un documento docx  
read\_docx() %>% body\_add\_flextable(ft) %>% print(target = tmp)  
browseURL(tmp) # abre el documento

**Tabla Q.3. – Activos de Capitalización pequeña**

|  | **15/2008 vs 38/2008** | **15/2008 vs 8/2010** | **38/2008 vs 8/2010** |
| --- | --- | --- | --- |
| Horq. 100 ac. | 0.0079 | 0.1508 | 0.0079 |
| Horq. 500 ac. | 0.0079 | 0.0079 | 0.0079 |
| Horq. 1000 ac. | 0.0159 | 0.0079 | 0.0079 |
| Prof. 1 tick | 0.0079 | 0.0079 | 0.0079 |
| Prof. 5 ticks | 0.0079 | 0.0317 | 0.0079 |
| Prof. 10 ticks | 0.0079 | 0.2222 | 0.0079 |
| Prof. 20 ticks | 0.1508 | 0.4206 | 0.0317 |

Como se observa, para los activos de todo tipo de capitalización las semanas resultan en general muy diferentes entre ellas. Sin embargo, la semana de la crisis es muy diferente a las demás excepto para la profundidad de 20 *ticks* (para los activos de capitalización pequeña, tabla Q.3.) y para la horquilla de un *round trip* de 500 acciones en los activos de gran capitalización (tabla Q.1.).

# 3. Evolución intradía de las medidas de liquidez

En este apartado se piden:

* Las medias diarias de la horquilla relativa media para *round-trips* de 100 y 1000 acciones y la profundidad media por el lado del *ask* a 5 y 20 *ticks* del punto medio para cada activo e intervalo de 15 minutos.
* Análisis de las diferencias en las medidas entre la semana de la crisis y las otras dos, según los parámetros del apartado anterior.

## 3.1. Medidas de liquidez media por activo e intervalo de 15 minutos

La liquidez media por activo e intervalo de 15 minutos puede calcularse a partir de los datos por minuto calculados en el apartado 1. El procedimiento consiste en generar una variable que recoja el cuarto de hora superior más próximo del minuto (todavía sin agrupar los diferentes días) y, tras agrupar según esta, calcular la media de los 15 minutos. Con posterioridad, y para cada activo, se extrae de esta la hora y el minuto (excluyendo día, mes y año) y, tras agrupar de nuevo, se recalculan las medias diarias por intervalo.

# Se inicializa una lista donde se almacenarán los resultados  
tabla\_i\_sumquart <- vector(mode = "list", length = length(tabla\_i\_ext))  
  
for (i in 1:length(tabla\_i\_ext)){  
 rt\_15 <- tabla\_i\_ext[[i]] %>%  
 # Se calculan horquillas para round-trips de 100 y 1000 como la media  
 # de los datos por minuto. Para ello se genera variable que redondea  
 # observación (minuto, información presente en nombre de las filas)   
 # a su cuarto de hora superior más próximo.  
 mutate(ceil\_date\_15 =   
 ceiling\_date(ymd\_hms(rownames(tabla\_i\_ext[[i]])),   
 unit = "15 minutes")) %>%  
 group\_by(ceil\_date\_15) %>% # se agrupa por cuarto de hora redondeado  
 summarise(mean\_rt\_100 = mean(rt\_100, na.rm = T), # cálculo de medias  
 mean\_rt\_1000 = mean(rt\_1000, na.rm = T))  
 tick\_15 <- tabla\_i\_ext2[[i]] %>%  
 # Mismoprocedimiento para las medidas de profundidad  
 mutate(ceil\_date\_15 = ceiling\_date(ymd\_hms(rownames(tabla\_i\_ext2[[i]])),   
 unit = "15 minutes")) %>%  
 group\_by(ceil\_date\_15) %>%  
 summarise(mean\_tick\_5 = mean(tick\_5),  
 mean\_tick\_20 = mean(tick\_20))  
 tabla\_i\_sumquart[[i]] <- merge(rt\_15, tick\_15)  
 rm(i, rt\_15, tick\_15)  
}  
  
# Se inicializa una lista donde se almacenarán los resultados  
tabla\_i\_quartday <- vector(mode = "list", length = length(tabla\_i\_sumquart))  
  
for (i in 1:length(tabla\_i\_sumquart)){  
 resume <- tabla\_i\_sumquart[[i]] %>%  
 # Se genera una variable hora:minuto extrayendo la información con  
 # funciones de lubridate y agregándolas como concatenaciones.  
 mutate(ceil\_quarter = hm(paste0(hour(ceil\_date\_15), ":",  
 minute(ceil\_date\_15)))) %>%  
 group\_by(ceil\_quarter) %>% # se agrupa por hora:minuto  
 # Luego calcula las medias de los grupos.  
 summarise(horq\_med100 = mean(mean\_rt\_100),  
 horq\_med1000 = mean(mean\_rt\_1000),  
 profund\_med5 = mean(mean\_tick\_5),  
 profund\_med20 = mean(mean\_tick\_20))  
 tabla\_i\_quartday[[i]] <- resume # se almacenan resultados.  
 rm(i, resume)  
}

Para representar la información gráficamente se hace necesario convertir a numérica la variable del cuarto de hora superior.

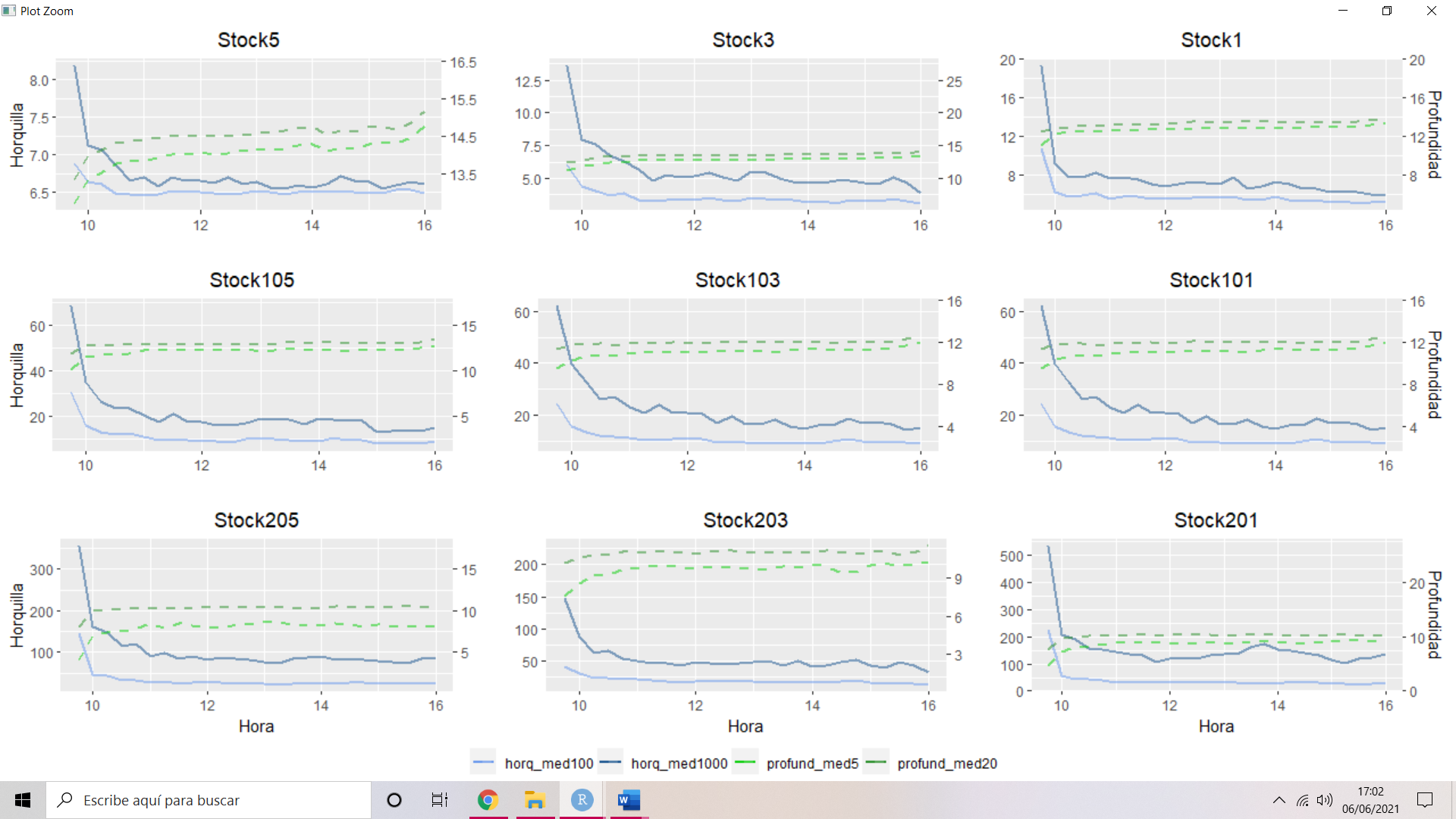
require(cowplot)  
fill\_color <- c("cornflowerblue", "dodgerblue4", "green3", "forestgreen")  
for (i in 1:length(tabla\_i\_quartday)){  
 tabla\_i\_quartday[[i]]$ceil\_quarter <-  
 as.numeric(tabla\_i\_quartday[[i]]$ceil\_quarter, "hours")  
 rm(i)  
}

Como la información completa recorre cuatro medidas de nueve activos, se ha optado por representarla en una composición de nueve gráficos, en cada uno de los cuales se incluye la evolución intradía de sus cuatro medidas de liquidez (cuatro líneas). Por lo que respecta al código, se define una función marco para cada gráfico y luego se pule la representación de cada uno de ellos en la composición.

sub\_graf <- function(data, trans){  
 ggplot(data = data) +  
 geom\_line(aes(x = ceil\_quarter, y = horq\_med100, color = "horq\_med100"),   
 alpha = .5, linetype = 1, size = .75) +  
 geom\_line(aes(x = ceil\_quarter, y = horq\_med1000, color = "horq\_med1000"),   
 alpha = .5, linetype = 1, size = .75) +  
 geom\_line(aes(x = ceil\_quarter, y = profund\_med5/trans,   
 color = "profund\_med5"),   
 alpha = .5, linetype = 2, size = .75) +  
 geom\_line(aes(x = ceil\_quarter, y = profund\_med20/trans,   
 color = "profund\_med20"),   
 alpha = .5, linetype = 2, size = .75) +  
 scale\_colour\_manual("", breaks = c("horq\_med100", "horq\_med1000",  
 "profund\_med5", "profund\_med20"),  
 values = c(fill\_color[1], fill\_color[2],   
 fill\_color[3], fill\_color[4])) +  
 theme(plot.title = element\_text(hjust = .5)) +  
 scale\_y\_continuous(sec.axis = sec\_axis(trans = ~.\*trans)) +  
 ylab("Horquilla") +  
 xlab("Hora")  
}  
  
legend <- get\_legend(sub\_graf(tabla\_i\_quartday[[1]], 1) + theme(legend.position = "bottom"))  
  
plot\_grid(plot\_grid(sub\_graf(tabla\_i\_quartday[[1]], 2) + xlab("") +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[1])),   
 sub\_graf(tabla\_i\_quartday[[2]], 2) + xlab("") +   
 ylab("") + theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[2])),   
 sub\_graf(tabla\_i\_quartday[[3]], 1) + xlab("") +   
 ylab("") + scale\_y\_continuous(sec.axis =   
 sec\_axis(trans = ~.\*1, name = "Profundidad")) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[3])),  
 sub\_graf(tabla\_i\_quartday[[4]], .25) + xlab("") +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[4])),   
 sub\_graf(tabla\_i\_quartday[[6]], .25) + xlab("") +   
 ylab("") + theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[5])),   
 sub\_graf(tabla\_i\_quartday[[6]], .25) + xlab("") +   
 ylab("") + scale\_y\_continuous(sec.axis =   
 sec\_axis(trans = ~.\*.25, name = "Profundidad")) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[6])),  
 sub\_graf(tabla\_i\_quartday[[7]], .05) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[7])),   
 sub\_graf(tabla\_i\_quartday[[8]], .05) +   
 ylab("") + theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[8])),   
 sub\_graf(tabla\_i\_quartday[[9]], .05) + ylab("") +   
 scale\_y\_continuous(sec.axis = sec\_axis(trans = ~.\*.05, name = "Profundidad")) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Stock",idx[9])),   
 nrow = 3), legend, nrow = 2, rel\_heights = c(0.95, 0.05))

En la Figura 1 se aprecia la evolución de las medidas de liquidez para los 9 activos. Los gráficos para cada uno de ellos se organizan por grados de capitalización en las filas. En cada gráfico, la variable tiempo está en el eje de las X; las medidas de horquilla están escaladas según el eje Y izquierdo y las de profundidad según el Y derecho.

En general se aprecia que la profundidad relativa es menor cuanto menor es el grado de capitalización. En la evolución intradía se aprecia, en general, un valor más bajo al principio de la jornada y una tendencia ligeramente ascendente hacia el final de esta. En cuanto a la horquilla relativa media, los valores son tanto más altos cuanto más grande es el grado de capitalización. En la evolución intradía la horquilla empieza ancha al inicio de la jornada y, tras bajar rápidamente durante la mañana, se estabiliza hasta el final de la sesión.

**Figura 1**

## 3.2. ¿Cómo afectó la semana de la crisis a la evolución de la liquidez intradía?

En este subapartado se pretende comprobar, de manera visual, cómo afectó la crisis a la evolución intradía de las medidas de liquidez. Para ello se efectúa un ejercicio similar al de 3.1., pero con la información de los activos reorganizada según semana y grado de capitalización. Los datos de partida son los datos originales agrupados por cuarto de hora, distinguiendo días.

for (i in 1:length(tabla\_i\_sumquart)){ # Genera var week para cada activo  
 tabla\_i\_sumquart[[i]] <- tabla\_i\_sumquart[[i]] %>%  
 # Se genera las variables para la semana y hora:minuto con  
 # funciones de lubridate.  
 mutate(week = paste0(isoweek(tabla\_i\_sumquart[[i]]$ceil\_date\_15), "-",   
 year(tabla\_i\_sumquart[[i]]$ceil\_date\_15)),  
 ceil\_quarter = hm(paste0(hour(ceil\_date\_15), ":",  
 minute(ceil\_date\_15))),  
 ceil\_quarter = as.numeric(ceil\_quarter, "hours"))  
 rm(i)  
}

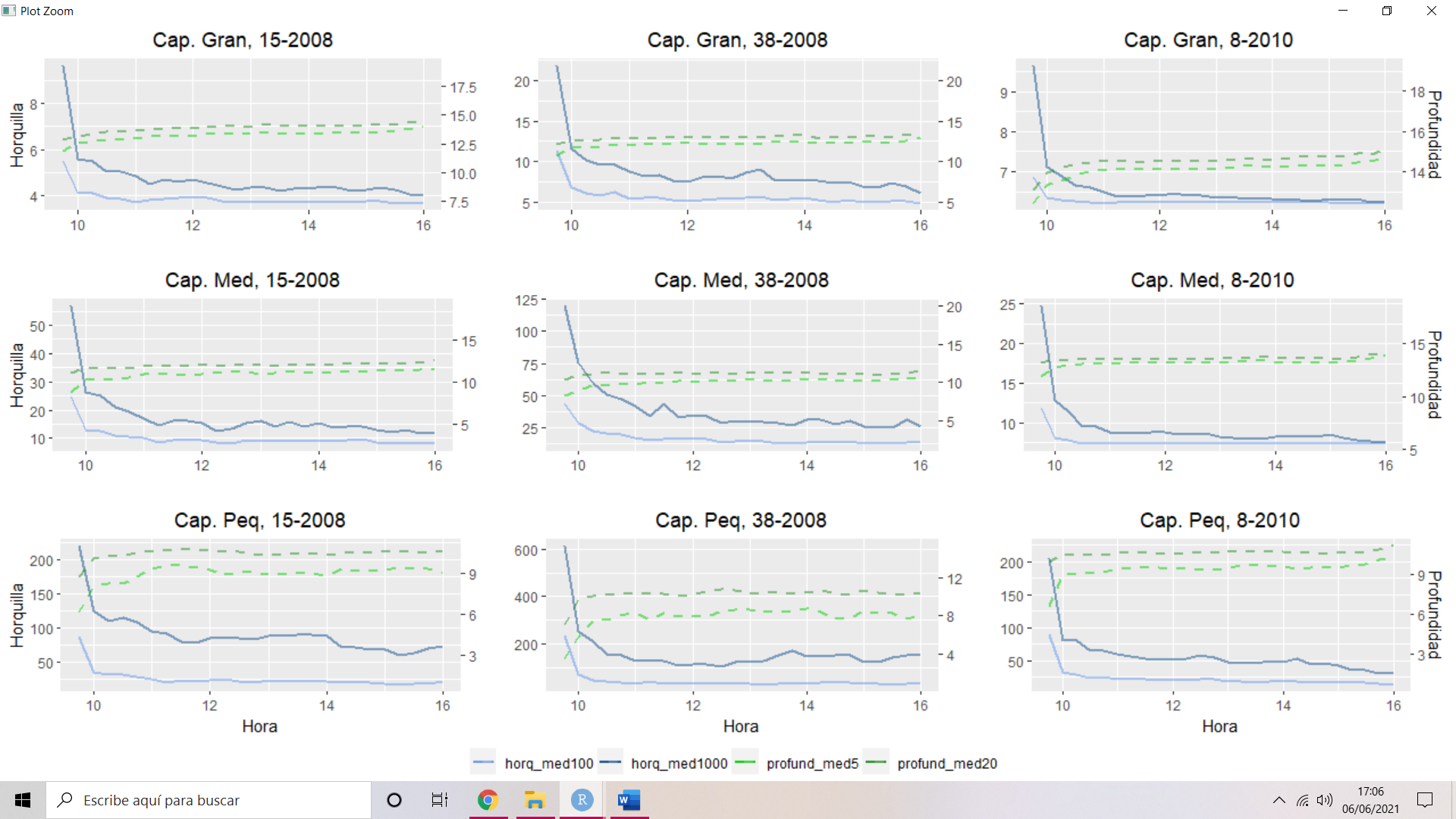
A partir de estos, y contando con las variables semana y hora:minuto generadas con lubridate, se procede a reorganizar la información según semana y hora:minuto.

# Se arma matriz de listas con los grados de capitalización en las filas   
# y semanas en las columnas  
tabla\_i\_weekhour <- matrix(vector(mode = "list", length = 1),   
 nrow = 3, ncol = 3)  
for (i in 1:length(idx2)){ # Por cada grado de capitalización  
 for (w in 1:length(weeks)){ # Por cada semana  
 agg <- data.frame()  
 for (j in 1:length(idx2[[i]])){ # Por cada activo  
 # Filtra activos y semana  
 agg\_med <- tabla\_i\_sumquart[[which(idx == idx2[[i]][j])]] %>%  
 filter(week == weeks[w])  
 agg <- rbind(agg, agg\_med)  
 }  
 agg <- agg %>%  
 group\_by(ceil\_quarter) %>% # Se agrupa por cuarto de hora  
 # Se sacan las medias  
 summarise(horq\_med100 = mean(mean\_rt\_100),  
 horq\_med1000 = mean(mean\_rt\_1000),  
 profund\_med5 = mean(mean\_tick\_5),  
 profund\_med20 = mean(mean\_tick\_20))  
 tabla\_i\_weekhour[i,w][[1]] <- agg # Se agrega a posición en matriz  
 }  
 rm(i, j, w, agg\_med, agg)  
}

Luego se representa gráficamente, de forma análoga a como se ha hecho en el apartado 3.1. La función básica se ha definido también en aquel apartado.

plot\_grid(plot\_grid(sub\_graf(tabla\_i\_weekhour[1, 1][[1]], 2) + xlab("") +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[1], ", ", weeks[1])),   
 sub\_graf(tabla\_i\_weekhour[1, 2][[1]], 1) + xlab("") +   
 ylab("") + theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[1], ", ", weeks[2])),   
 sub\_graf(tabla\_i\_weekhour[1, 3][[1]], 2) + xlab("") +  
 ylab("") + scale\_y\_continuous(sec.axis = sec\_axis(trans = ~.\*2, name = "Profundidad")) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[1], ", ", weeks[3])),  
 sub\_graf(tabla\_i\_weekhour[2, 1][[1]], (1/3)) + xlab("") +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[2], ", ", weeks[1])),   
 sub\_graf(tabla\_i\_weekhour[2, 2][[1]], (1/6)) + xlab("") +   
 ylab("") + theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[2], ", ", weeks[2])),   
 sub\_graf(tabla\_i\_weekhour[2, 3][[1]], .75) + xlab("") +   
 ylab("") + scale\_y\_continuous(sec.axis = sec\_axis(trans = ~.\*.75, name = "Profundidad")) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[2], ", ", weeks[3])),  
 sub\_graf(tabla\_i\_weekhour[3, 1][[1]], .05) +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[3], ", ", weeks[1])),   
 sub\_graf(tabla\_i\_weekhour[3, 2][[1]], .025) + ylab("") +   
 theme(legend.position = "none") +   
 ggtitle(paste0("Cap. ", cap[3], ", ", weeks[2])),   
 sub\_graf(tabla\_i\_weekhour[3, 3][[1]], .05) + ylab("") +   
 scale\_y\_continuous(sec.axis = sec\_axis(trans = ~.\*.05, name = "Profundidad")) +   
 ggtitle(paste0("Cap. ", cap[3], ", ", weeks[3])) +  
 theme(legend.position = "none"),   
 nrow = 3), legend, nrow = 2, rel\_heights = c(0.95, 0.05))

**Figura 2**



De acuerdo con lo observado, se aprecian diferencias tanto en la submuestra de gran capitalización como en las demás. En las submuestras de capitalización grande y media las horquillas son ligeramente mayores; la diferencia se exacerba en la submuestra de capitalización pequeña. En cuanto a la profundidad, también se aprecian diferencias en la profundidad en las submuestras de capitalización media y pequeña: la semana de la crisis la profundidad era menor. Con todo lo anterior puede decirse que la semana de la crisis se incrementó la iliquidez del mercado tanto más como menor era el grado de capitalización.

# 4. La importancia del volumen oculto en la provisión de liquidez

En este apartado se intentará responder a:

* ¿Qué proporción de la profundidad total del LOB está oculta?
* ¿Hubo más volumen oculto durante la semana de la crisis?, ¿Hay diferencias en el uso de volumen oculto en función de la capitalización bursátil?
* ¿Tienden a ocultarse más o menos cerca del punto medio?

## 4.1. Proporción del volumen oculto en la profundidad total

Para medir esta magnitud por activo se procede al cálculo de la profundidad total por las partes *ask* y *bid*, tanto del volumen total como del oculto. Las dos ratios del volumen oculto sobre el total (*ask* y *bid*) representan los porcentajes del volumen oculto en ese minuto en particular. Los datos por minuto se almacenan en una tabla y, con posterioridad, se calcula la media de la profundidad en cada lado del LOB. El promedio de estas dos medias representa la importancia del volumen oculto para ese activo en particular.

# Subtabla Profunidad, tabla\_ii\_sub1 (9 subtablas)  
tabla\_ii\_sub1 <- vector(mode = "list", length = length(LOB))  
  
for (d in 1:length(LOB)){  
 minute <- unique(LOB[[d]]$datime) # se extrae un vector con los minutos  
 subtabla\_ii\_sub1 <- data.frame() # se inicializa data.frame depositario  
 for (t in 1:length(minute)){  
 total\_deep <- function(part){  
 # part es el parámetro para ask y bid; para ask es -1  
 LOB\_partmin <- LOB[[d]] %>%  
 filter(datime == minute[t] & sign\*part > 0)  
 deep\_v <- sum(LOB\_partmin$vol\*LOB\_partmin$quote)  
 deep\_hv <- sum(LOB\_partmin$hvol\*LOB\_partmin$quote)  
 return(deep\_hv/deep\_v)  
 }  
 # En cell se almacenan resultados de la función total\_deep de cálculo  
 # de la profundidad total.  
 cell2 <- data.frame(deep\_a = total\_deep(-1), deep\_b = total\_deep(1))  
 # Acumula los resultados por minuto:  
 subtabla\_ii\_sub1 <- rbind(subtabla\_ii\_sub1, cell2)  
 }  
 # Se almacenan tablas resultados en lista, convenientemente renombradas  
 # las filas.  
 tabla\_ii\_sub1[[d]] <- subtabla\_ii\_sub1  
 rownames(tabla\_ii\_sub1[[d]]) <- minute  
 cat(paste0("Stock",idx[d]," completado. ")) # Muestra prog en pantalla.  
 rm(d, t, minute, cell2, subtabla\_ii\_sub1, total\_deep)  
}

# Resumen de las tablas de tabla\_ii\_sub1  
tabla\_ii\_part1 <- data.frame()  
  
for (d in 1:length(tabla\_ii\_sub1)){  
 ask <- mean(tabla\_ii\_sub1[[d]]$deep\_a)  
 bid <- mean(tabla\_ii\_sub1[[d]]$deep\_b)  
 propf\_oculta <- data.frame(propf\_oculta = (ask+bid)/2)  
 tabla\_ii\_part1 <- rbind(tabla\_ii\_part1, propf\_oculta)  
 rm(ask, bid, d, propf\_oculta)  
}  
rownames(tabla\_ii\_part1) <- paste0("Stock", idx)

## 4.2. El volumen oculto durante la semana de la crisis

Para determinar la importancia del volumen oculto durante esa semana en particular (en contraposición a las otras dos) se considera suficiente obtener la ratio del volumen oculto sobre el volumen total para cada semana y activo.

# Subtabla para semanas, tabla\_ii\_sub2 (9 subtablas)  
tabla\_ii\_sub2 <- vector(mode = "list", length = length(LOB))  
  
for (d in 1:length(LOB)){  
 minute <- unique(LOB[[d]]$datime) # se extrae un vector con los minutos  
 subtabla\_ii\_sub2 <- data.frame() # se inicializa data.frame depositario  
 for (t in 1:length(minute)){  
 LOB\_partmin <- LOB[[d]] %>%  
 filter(datime == minute[t]) # Filtra LOB por minuto  
 occ\_vol <- sum(LOB\_partmin$hvol)/sum(LOB\_partmin$vol)  
 cell <- data.frame(Vol\_ocult = occ\_vol,  
 minute = minute[t])  
 # Acumula los resultados por minuto:  
 subtabla\_ii\_sub2 <- rbind(subtabla\_ii\_sub2, cell)  
 }  
 # Se almacenan las tablas resultados en lista,   
 # convenientemente renombradas las filas.  
 tabla\_ii\_sub2[[d]] <- subtabla\_ii\_sub2  
 rownames(tabla\_ii\_sub2[[d]]) <- minute  
 cat(paste0("Stock",idx[d]," completado. ")) # Muestra prog en pantalla.  
 rm(d, t, minute, cell, occ\_vol, subtabla\_ii\_sub2, LOB\_partmin)  
}

# Resumen de las tablas de tabla\_ii\_sub2  
tabla\_ii\_part2 <- data.frame()  
  
for (d in 1:length(tabla\_ii\_sub2)){  
 # Se generala variable semana  
 tabla\_ii\_sub2[[d]]$week <- paste0(isoweek(tabla\_ii\_sub2[[d]]$minute), "-",   
 year(tabla\_ii\_sub2[[d]]$minute))  
 weeks <- unique(tabla\_ii\_sub2[[d]]$week)  
 subtabla\_ii\_sub3 <- data.frame(aux = rep(NA, 1)) # Inicializa tabla  
 for (w in weeks){ # Por cada semana  
 tabla\_ii\_sub\_w <- tabla\_ii\_sub2[[d]] %>%  
 filter(week == w)   
 occ\_vol <- mean(tabla\_ii\_sub\_w$Vol\_ocult) # Media del vol oculto  
 Vol\_ocult <- data.frame(occ\_vol) # Transforma en data.frame  
 subtabla\_ii\_sub3 <- cbind(subtabla\_ii\_sub3, Vol\_ocult) # Agrega a df  
 }  
 tabla\_ii\_part2 <- rbind(tabla\_ii\_part2, subtabla\_ii\_sub3[,2:4])  
 rm(d, occ\_vol, Vol\_ocult, subtabla\_ii\_sub3, tabla\_ii\_sub\_w, w)  
}  
rownames(tabla\_ii\_part2) <- paste0("Stock", idx)  
colnames(tabla\_ii\_part2) <- paste0("Semana\_", weeks)

## 4.3. Distancia media del volumen oculto al precio eficiente

Para conocer el grado en que el volumen oculto se posiciona más o menos cerca del precio eficiente primero cabe definir cómo va a medirse la distancia. Podría medirse en ticks, pero se ha considerado más conveniente medirlo en términos relativos, como porcentaje del precio de la acción.

Para obtener esta medida de distancia es necesario obtener el precio eficiente para cada minuto del LOB y los precios medios ponderados en el *ask* y el *bid* del volumen oculto, estos últimos a partir de los vectores de precios y volúmenes ocultos. Las distancias relativas del volumen oculto se calculan como el valor absoluto de la resta entre los precios medios ponderados *ask* y *bid* y el precio eficiente. Finalmente se calcula la media de ambas distancias ponderadas por el volumen oculto en cada caso.

Los datos por minuto y activo se almacenan en tablas. Luego se agregan los datos por semanas por si pudieran vislumbrarse diferencias entre ellas.

# Subtabla para distancia, tabla\_ii\_sub3 (9 subtablas)  
tabla\_ii\_sub3 <- vector(mode = "list", length = length(LOB))  
  
for (d in 1:length(LOB)){  
 minute <- unique(LOB[[d]]$datime) # se extrae un vector con los minutos  
 subtabla\_ii\_sub3 <- data.frame() # Inicializan data.frames depositarios  
 for (t in 1:length(minute)){  
 LOB\_partmin <- LOB[[d]] %>%  
 # filtra por minuto  
 filter(datime == minute[t])  
 # Cálculo de los puntos medios:  
 q <- (LOB\_partmin$quote[LOB\_partmin$sign == 1] +   
 LOB\_partmin$quote[LOB\_partmin$sign == -1])/2  
 ask\_hvol <- LOB\_partmin$hvol[LOB\_partmin$sign < 0]   
 # se vectorizan volúmenes...  
 bid\_hvol <- LOB\_partmin$hvol[LOB\_partmin$sign > 0]  
 ask\_quote <- LOB\_partmin$quote[LOB\_partmin$sign < 0] # ...y precios  
 bid\_quote <- LOB\_partmin$quote[LOB\_partmin$sign > 0]  
 ask\_meanq <- ifelse(sum(ask\_hvol) == 0, 0,   
 sum(ask\_hvol\*ask\_quote)/sum(ask\_hvol))  
 bid\_meanq <- ifelse(sum(bid\_hvol) == 0, 0,   
 sum(bid\_hvol\*bid\_quote)/sum(bid\_hvol))  
 ask\_reldist <- ifelse(ask\_meanq == 0, 0, abs(q - ask\_meanq)/ask\_meanq)  
 bid\_reldist <- ifelse(bid\_meanq == 0, 0, abs(q - bid\_meanq)/bid\_meanq)  
 pmean\_reldist <- ifelse(sum(ask\_hvol) == 0 & sum(bid\_hvol) == 0, NA,  
 ((ask\_reldist\*sum(ask\_hvol)) +   
 (bid\_reldist\*sum(bid\_hvol)))/  
 (sum(ask\_hvol)+sum(bid\_hvol)))  
 # en cell se almacenan los resultados  
 cell <- data.frame(reldist = pmean\_reldist)  
 # Acumula los resultados por minuto:  
 subtabla\_ii\_sub3 <- rbind(subtabla\_ii\_sub3, cell)  
 }  
 # Se almacenan tablas resultados en lista, convenientemente renombradas  
 # las filas.  
 cat(paste0("Stock", idx[d]," completado. ")) # Muestra prog en pantalla.  
 tabla\_ii\_sub3[[d]] <- subtabla\_ii\_sub3  
 rownames(tabla\_ii\_sub3[[d]]) <- minute  
 rm(d, t, q, ask\_hvol, bid\_hvol, ask\_quote, bid\_quote, ask\_meanq, bid\_meanq,   
 ask\_reldist, bid\_reldist, pmean\_reldist, minute, cell, subtabla\_ii\_sub3,   
 LOB\_partmin)  
}

# Se separa por semana  
tabla\_ii\_part3 <- data.frame()  
  
for (d in 1:length(tabla\_ii\_sub3)){ # Por cada activo  
 tabla\_ii\_sub3[[d]]$week <-   
 paste0(isoweek(ymd\_hms(rownames(tabla\_ii\_sub3[[d]]))), "-",   
 year(ymd\_hms(rownames(tabla\_ii\_sub3[[d]]))))  
 row <- data.frame(aux = rep(NA, 1)) # se inicializa data.frame  
 for (w in 1:length(weeks)){ # Por cada semana  
 tabla\_week <- tabla\_ii\_sub3[[d]] %>%  
 # Filtra por semana  
 filter(week == weeks[w])  
 mean\_reldist <- mean(tabla\_week$reldist, na.rm = T) # media de dist.  
 mean\_reldist <- data.frame(mean\_reldist)  
 row <- cbind(row, mean\_reldist) # agrega celda a la fila  
 }  
 row <- row[,2:4]  
 colnames(row) <- paste0("Dist\_q\_", weeks)  
 tabla\_ii\_part3 <- rbind(tabla\_ii\_part3, row)  
 rm(d, row, w, tabla\_week, mean\_reldist)  
}  
rownames(tabla\_ii\_part3) <- paste0("Stock",idx)

Por último, se agregan las tres partes que integran Tabla II y se sacan las medias de todas las medidas de los activos por grado de capitalización. Los resultados se presentan como porcentajes.

# Agregación de resultados: Tabla II  
tabla\_ii <- cbind(tabla\_ii\_part1, tabla\_ii\_part2)  
tabla\_ii <- cbind(tabla\_ii, tabla\_ii\_part3)  
  
# Se sacan las medias por grado de capitalización  
columns\_t2 <- colnames(tabla\_ii)  
  
tabla\_ii\_transformed <- data.frame()  
for (i in 1:length(idx2)){  
 tabla\_ii\_transfer <-   
 tabla\_ii[which(idx == idx2[[i]][1]):which(idx == idx2[[i]][3]),]  
 mean\_third <-   
 data.frame(lapply(tabla\_ii[which(idx == idx2[[i]][1]) :   
 which(idx == idx2[[i]][3]),][, columns\_t2], mean))  
 colnames(mean\_third) <- columns\_t2  
 rownames(mean\_third) <- paste0("Stock", idx[which(idx == idx2[[i]][1])],   
 "-", idx[which(idx == idx2[[i]][3])])  
 tabla\_ii\_transfer <- rbind(tabla\_ii\_transfer, mean\_third)  
 tabla\_ii\_transformed <- rbind(tabla\_ii\_transformed, tabla\_ii\_transfer)  
}  
  
require(scales)  
tabla\_ii\_transformed[, columns\_t2] <-   
 lapply(tabla\_ii\_transformed[, columns\_t2], percent\_format(accuracy = .01))

# Para exportar tablas a formato Word  
ft <- flextable(data = tabla\_ii\_transformed %>% add\_rownames()) %>%   
 theme\_zebra %>% autofit  
ft <- set\_caption(ft, caption = "Tabla II",   
 style = "Table Caption")  
# Crea un archivo temp  
tmp <- tempfile(fileext = ".docx")  
# Crea un documento docx  
read\_docx() %>% body\_add\_flextable(ft) %>% print(target = tmp)  
browseURL(tmp) # abre el documento

**Tabla II**

|  | **Vol. oculto**  **/ Total** | **Volumen oculto /**  **Total, por semana** | | | **Distancia al p. medio (q),**  **por semana (% de q)** | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **15/2008** | **38/2008** | **8/2010** | **15/2008** | **38/2008** | **8/2010** |
| Stock5 | 12.53% | 16.35% | 17.36% | 8.14% | 0.23% | 0.21% | 0.30% |
| Stock3 | 12.82% | 14.34% | 16.90% | 11.17% | 0.14% | 0.10% | 0.14% |
| Stock1 | 14.26% | 28.69% | 13.86% | 4.69% | 0.11% | 0.15% | 0.31% |
| **Stock5-1** | **13.21%** | **19.79%** | **16.04%** | **8.00%** | **0.16%** | **0.15%** | **0.25%** |
| Stock105 | 25.04% | 42.11% | 33.16% | 15.39% | 0.24% | 0.89% | 0.22% |
| Stock103 | 15.37% | 15.01% | 25.57% | 5.50% | 0.35% | 1.17% | 0.37% |
| Stock101 | 20.29% | 21.10% | 25.27% | 14.26% | 0.30% | 0.67% | 0.58% |
| **Stock105-101** | **20.23%** | **26.08%** | **28.00%** | **11.71%** | **0.30%** | **0.91%** | **0.39%** |
| Stock205 | 43.96% | 61.01% | 40.14% | 38.98% | 2.14% | 1.26% | 1.22% |
| Stock203 | 35.42% | 27.08% | 57.28% | 29.24% | 0.46% | 1.11% | 1.28% |
| Stock201 | 38.70% | 54.72% | 47.14% | 21.58% | 3.07% | 2.16% | 1.52% |
| **Stock205-201** | **39.36%** | **47.60%** | **48.19%** | **29.93%** | **1.89%** | **1.51%** | **1.34%** |

De los resultados de la tabla II se desprende, por un lado, que el volumen oculto es tanto más importante relativamente como menor sea el grado de capitalización bursátil. Está alrededor del 13% en los activos de gran capitalización, del 20% en los activos de capitalización media y casi en el 40% para los de capitalización pequeña. Por otro, se detecta más volumen oculto en la negociación durante la semana de la crisis; comparada con la semana 8/2010 la diferencia es grande, pero es menor respecto de la 15/2008 (incluso se mueve en sentido contrario para los activos de gran capitalización). En cuanto a la distancia al precio eficiente a la que se colocan los *traders* que ocultan su volumen, se observa que se ocultan tanto más cerca cuanto más grande sea el grado de capitalización; el efecto de la crisis no parece afectar excesivamente a este indicador: apenas se aprecia una diferencia significativa para los activos de capitalización media.

# 5. Conclusiones

De lo anterior puede concluirse que las horquillas ponderadas medias son más estrechas cuanto más alto es el grado de capitalización. También puede decirse que estas empiezan más anchas al inicio de la sesión bursátil, pero que bajan durante la mañana y permanecen estables o disminuyen hacia el final de la jornada. Además, se concluye que la crisis en general ensanchó las horquillas, exacerbando el efecto cuanto más pequeño era el grado de capitalización bursátil. Complementariamente se observa una disminución de la profundidad relativa con el grado de capitalización, al inicio de la sesión y una tendencia estable o tendente al aumento hacia el final de la jornada. Respecto de estas medidas, la evidencia sugiere que la crisis se tradujo en un grado menor de profundidad relativa, especialmente por lo que respecta a los activos de capitalización media y pequeña.

Con todo, puede decirse que existen mayores grados de liquidez en los activos de gran capitalización que en los de capitalización media y mayor en estos últimos que en los de capitalización pequeña. Además, en general, los mercados son más ilíquidos al inicio de la sesión bursátil, esto se atenúa durante la jornada y tiende a aumentar hacia el final de esta.

Las tres semanas analizadas son distintas entre sí, pero la semana de la crisis pareció traducirse en una disminución de la liquidez del mercado, con efectos tanto mayores cuanto menor era el grado de capitalización bursátil.

Por último, cabe señalar que el volumen oculto es relativamente más importante en la negociación de los activos de menos capitalización. Cuanto más grande es el grado de capitalización más cerca del precio eficiente se ocultan los *traders*. Aunque la crisis supuso un aumento del volumen oculto en la negociación (en especial para los activos de capitalización media y pequeña), la diferencia de distancia al precio eficiente a la que se ocultan no se ha observado demasiado significativa excepto por lo que respecta a los activos de capitalización media.