Assignment 1: Effect of Financial Crisis on Corporate Profits

Dana DiVincenzo
6/13/21

Introduction

The stock market crash of 2007-2008 was considered to be the most severe financial crisis since the Great Depression. This crisis had long lasting impacts on the economy, and according the the Wall Street Journal, the total financial losses from this have been estimated to be $15 trillion.

In this assignment, I look at the financial data of firms in the years leading up to, during, and following the crash. I answer three questions to better understand the effects of the crisis on these firms and I look at if those firms were able to fully recover.

Question 1: What is the financial effect of the crisis on these firms, in terms of percent change of net income, from pre-crisis to crisis?

```
library(ggplot2)
library(dplyr)
library(readr)
library(knitr)

# reading in the file
df <- read.csv("/Users/dana/Downloads/Compustat 1990-2015 Lots.csv")
# dropping rows where tic col value is blank
df <- df[df$tic != "",]
```

```
# finds the total net income for each company each year and puts that value into a new d
f
total_ni_df <- df %>% group_by(tic, fyear) %>% summarise(total_ni = sum(ni))

# takes the total net income for each company in the years 2004, 2005, and 2006, and fin
ds the average of them all
# puts those values into a df
pre_crisis <- total_ni_df %>% group_by(tic) %>% filter(fyear==2004|fyear==2005|fyear==20
06) %>%
  summarise(pre_crisis_avg=mean(total_ni))

# takes the total net income for each company in the years 2007 and 2008 and finds the a
verage of the two
# puts those values into a df
crisis <- total_ni_df %>% group_by(tic) %>% filter(fyear==2007|fyear==2008) %>%
  summarise(post_crisis_avg=mean(total_ni))
```

```
# combining the pre and post crisis data frames horizontally
pre_and_post <- merge(pre_crisis, crisis)

# dropping the rows with na or 0 values in the pre_crisis_avg column
pre_and_post <- pre_and_post[!is.na(pre_and_post$pre_crisis_avg),]
pre_and_post <- pre_and_post[!pre_and_post$pre_crisis_avg == 0,]

# dropping the rows with na or 0 values in the post_crisis_avg column
pre_and_post <- pre_and_post[!is.na(pre_and_post$post_crisis_avg),]
pre_and_post <- pre_and_post[!pre_and_post$post_crisis_avg == 0,]

# calculating the percent difference between the pre and post crisis averages and making
it a new column
pre_and_post$percent <- (pre_and_post$post_crisis_avg - pre_and_post$pre_crisis_avg) / a
bs(pre_and_post$pre_crisis_avg) * 100

pre_and_post <- as.data.frame(pre_and_post)
```

```
# creating a five number summary of the values in the percent col
min_percent <- min(pre_and_post$percent)
first_quartile <- quantile(pre_and_post$percent, 0.25)
median_percent <- median(pre_and_post$percent)
third_quartile <-quantile(pre_and_post$percent, 0.75)
max_percent <- max(pre_and_post$percent)

five_number_summary_df <- data.frame(min_percent, first_quartile, median_percent, third_
quartile, max_percent)
names(five_number_summary_df) <- c("Minimum", "First Quartile", "Median", "Third Quartil
e", "Maximum")
row.names(five_number_summary_df) <- "Percent Change"
kable(five_number_summary_df, align = "lllll", caption = "Five Number Summary for Percen
t Change in Firms Average Net Income Pre-Crisis to Crisis.")
```

Five Number Summary for Percent Change in Firms Average Net Income Pre-Crisis to Crisis.

	Minimum	First Quartile	Median	Third Quartile	Maximum
Percent Change	-1.888716e+19	-124.0176	-2.588286	68.80517	141800

Question 1 Analysis

The above table shows us the five number summary for the percent change in firms average net income, from the pre-crisis period of 2004-2006, to the crisis period of 2007-2008. We can first see that the range in these values is very large, as the minimum value shows a huge percent decrease, and the highest value shows a large (but not as significant as the minimum value) percent increase. Because these low numbers show such a large percent decrease, it is likely that firms towards the bottom of the range had their net income go towards or below 0 in the crisis years.

The median change in this time was -2.588%, which tells us that the majority of firms had a lower net income during the crisis period than they did before the crisis period. This makes complete sense given a basic understanding of the crisis. However, -2.588% is not as drastic as a decrease as I would have expected

considering the severity of the crisis. Given that the minimum values here are so great, if we calculated the average percent change for the same time period, it would likely be significantly lower than the median shown here.

Question 2: What are the 10 firms with the largest and smallest changes in net income, respective to percent and absolute change?

```
# sorting the df by absolute value of the percent col, largest at the top
pre_and_post <- pre_and_post %>%
   arrange(desc(abs(pre_and_post$percent)))

# now need to capture the top 10 values in the percent column and their tics, put it in
 a table
tic_and_percent_cols <- c(1, 4)
top_10_percent_change <- pre_and_post[1:10, tic_and_percent_cols]

names(top_10_percent_change) <- c("tic", "Percent Change (%)")
kable(top_10_percent_change, align = "ll", caption = "Top 10 Firms with the Greatest Per
cent Change in Average Net Income Pre-Crisis to Crisis")
```

Top 10 Firms with the Greatest Percent Change in Average Net Income Pre-Crisis to Crisis

tic	Percent Change (%)
PGUS	-1.888716e+19
STAQ	1.418000e+05
NHLD	-1.338432e+05
ENTN	-9.898880e+04
0419B	-9.586731e+04
RAE	-6.982766e+04
DEK	-6.970000e+04
CBEY	6.288500e+04
3WCPSF	-6.169096e+04
FSCI	6.083968e+04

```
# sorting the df by absolute value of the percent col, smallest at the top
pre_and_post <- pre_and_post %>%
   arrange((abs(pre_and_post$percent)))

# now need to capture the bottom 10 values in the percent column and their tics, put it
 in a table
bottom_10_percent_change <- pre_and_post[1:10, tic_and_percent_cols]

names(bottom_10_percent_change) <- c("tic", "Percent Change (%)")
kable(bottom_10_percent_change, align = "ll", caption = "Top 10 Firms with the Least Per
cent Change in Average Net Income Pre-Crisis to Crisis")
```

Top 10 Firms with the Least Percent Change in Average Net Income Pre-Crisis to Crisis

tic	Percent Change (%)
UNT	-0.0113028
GXP	-0.0426977
SKH	-0.0585080
MCO	-0.0660996
CEHC	0.0895736
KHD.Z	0.1079447
GAXIQ	-0.1333778
VLCM	-0.1415043
ETCC	-0.1535743
RUSHA	0.1732778

```
# calculating the absolute change between the pre and post crisis averages and adding th
at as a column to the df
pre_and_post$abs_change <- (pre_and_post$post_crisis_avg - pre_and_post$pre_crisis_avg)
```

```
# sorting the df by absolute value of the abs_change col, largest at the top
pre_and_post <- pre_and_post %>%
   arrange(desc(abs(pre_and_post$abs_change)))

# now need to capture the top 10 values in the absolute change column and their tics, pu
t it in a table
tic_and_abs_cols <- c(1, 5)
top_10_abs_change <- pre_and_post[1:10, tic_and_abs_cols]

names(top_10_abs_change) <- c("tic", "Absolute Change ($)")
kable(top_10_abs_change, align = "ll", caption = "Top 10 Firms with the Greatest Absolut
e Change in Average Net Income Pre-Crisis to Crisis")
```

Top 10 Firms with the Greatest Absolute Change in Average Net Income Pre-Crisis to Crisis

tic	Absolute Change ($)
GM	-31549.000
VOD	29826.196
UBS	-21383.185
S	-16888.667
COP	-14956.167

tic	Absolute Change ($)
OGZPY	12306.667
XOM	9261.667
TWX	-8786.500
BHP	7782.333
BBL	7660.667

```
# sorting the df by absolute value of the abs_change col, smallest at the top
pre_and_post <- pre_and_post %>%
   arrange((abs(pre_and_post$abs_change)))

# now need to capture the bottom 10 values in the absolute column and their tics, put it
in a table
bottom_10_abs_change <- pre_and_post[1:10, tic_and_abs_cols]

names(bottom_10_abs_change) <- c("tic", "Absolute Change ($)")
kable(bottom_10_abs_change, align = "ll", caption = "Top 10 Firms with the Least Absolut
e Change in Average Net Income Pre-Crisis to Crisis")
```

Top 10 Firms with the Least Absolute Change in Average Net Income Pre-Crisis to Crisis

tic	Absolute Change ($)
ASOE	-0.0003333
CEHC	0.0008333
ANML.1	0.0011667
VODG	0.0011667
IHT	-0.0013333
ORRMF	0.0023333
AXRX	-0.0035000
OCTI	-0.0036667
TRNS	-0.0045000
ITDN	-0.0045000

Question 2 Analysis

The four tables in the previous section show us the top 10 firms with the largest and smallest changes in the pre-crisis to crisis period, respective to both percent change and absolute change. Because these tables show change, and not gains only or losses only, each table contains both positive and negative values. In both the percent and absolute tables showing the *greatest* changes, the firms associated with a positive number did very well in the crisis, as they had a huge positive change. In those same tables, the firms with negative numbers did the worst, as they had the largest negative changes. In both the percent and absolute tables showing the *smallest*

changes, however, the firms with both positive and negative numbers did similarly. The values here are so small that a change in either the positive or negative direction would likely not be a huge impact on a firm. One important thing to note is that the values in the largest percent change table are much bigger than the values in the largest absolute change table, because the percent change values show the relative change in firm performance rather than the actual change.

Question 3: How long did it take for firms to recover from the crisis?

```
# takes the total net income for each company in the years 2004, 2005, and 2006, and fin
ds the max of them all
# puts those values into a df
pre_max <- total_ni_df %>% group_by(tic) %>% filter(fyear==2004|fyear==2005|fyear==2006)
%>%
  summarise(pre_crisis_max=max(total_ni))

# dropping NAs
pre_max <- pre_max[!is.na(pre_max$pre_crisis_max),]
```

```
# grouping the total net income dataframe by year and firm, then filtering so we only se
lect the years 2009 - 2014
# making a new column from this calld breach_years, which tells us if the total net inco
me for that year is greater
# than the highest net income in the pre-crisis years, expressed as a boolean (TRUE = ?,
FALSE = ?)
breach_df <- total_ni_df %>% group_by(tic, fyear) %>%
  filter(fyear==2009|fyear==2010|fyear==2011|fyear==2012|fyear==2013|fyear==2014) %>%
  summarise(breach_years=(pre_max$pre_crisis_max[pre_max$tic == tic] <
              total_ni_df$total_ni[which(total_ni_df$tic == tic & total_ni_df$fyear == f
year)]))
```

```
recover_df <- data.frame()

# looping through each unique firm
for (tic in unique(pre_max$tic)) {
  YearsRecover <- NA
  # for each year and firm combo, if breach year is TRUE and if breach year is not 0 or
 NA,
  # we set the variable YearsRecover to be the difference between that year and 2008
  for (year in 2009:2014) {
    breach_year <- breach_df$breach_years[which(breach_df$tic==tic & breach_df$fyear ==
 year)]
    if (length(breach_year) != 0 && !is.na(breach_year) && breach_year) {
      YearsRecover <- (year - 2008)
      break
    }
  }
  # binds the empty data frame with the YearsRecover variable and associated tic
  recover_df <- rbind(recover_df, data.frame(tic, YearsRecover))
}
```

```
# dropping all NA values
removed_NA_YR <- na.omit(recover_df)
```

```
# finding the percent NA in the YearsRecover column, = 57.05692
percent_NA <- length(which(is.na(recover_df$YearsRecover))) / nrow(recover_df) * 100

# finding percent for each year
percent1 <- length(which(recover_df$YearsRecover == 1)) / nrow(recover_df) * 100
percent2 <- length(which(recover_df$YearsRecover == 2)) / nrow(recover_df) * 100
percent3 <- length(which(recover_df$YearsRecover == 3)) / nrow(recover_df) * 100
percent4 <- length(which(recover_df$YearsRecover == 4)) / nrow(recover_df) * 100
percent5 <- length(which(recover_df$YearsRecover == 5)) / nrow(recover_df) * 100
percent6 <- length(which(recover_df$YearsRecover == 6)) / nrow(recover_df) * 100
```

```
# finding the average of the YearsRecover col
mean_YR <- mean(removed_NA_YR$YearsRecover)

# creating a table which holds NA percent, percent by year, and mean
YR_table <- data.frame(percent_NA, percent1, percent2, percent3, percent4, percent5, per
cent6, mean_YR)
names(YR_table) <- c("Percent Which Did Not Recover by 2014",
                     "Percent Recovered in One Year", "Percent Recovered in Two Years",
                     "Percent Recovered in Three Years", "Percent Recovered in Four Year
s",
                     "Percent Recovered in Five Years", "Percent Recovered in Six Years"
,
                     "Average Recovery Time of Firms Which Recovered by 2014 (in years)"
)
YR_table_t <- as.data.frame(t(YR_table))
names(YR_table_t) <- ""
kable(YR_table_t, align = "ll", caption = "Summary of Firms Recovery from Crisis, in Yea
rs 2009-2014")
```

Summary of Firms Recovery from Crisis, in Years 2009-2014

Percent Which Did Not Recover by 2014	57.056923
Percent Recovered in One Year	21.933831
Percent Recovered in Two Years	9.602317
Percent Recovered in Three Years	4.990531
Percent Recovered in Four Years	2.840593
Percent Recovered in Five Years	2.116520
Percent Recovered in Six Years	1.459285
Average Recovery Time of Firms Which Recovered by 2014 (in years)	2.021530

Question 3 Analysis

The above table gives us a summary of the recovery times of firms after the financial crisis. If a firm was able to reach it's maximum net income from the pre-crisis period of 2004-2006, we can consider it fully recovered. We cut off our analysis here in the year 2014, so if a firm was not fully recovered by then, we assigned it a NA value. If the firm recovered by 2014, we recorded the length of time in years it took to recover.

We can see that over half of the firms, 57.06%, did not recover by 2014. This is significant and shows the long-lasting effects of this crisis. However, of the firms that did recover, a majority of them did so within 1 year. As time continues, we can see that less and less firms recover in that year. This tells me that the longer the firm is taking to recover, the less likely it is to recover at all. The last line in the table tells us that the average recovery time of the firms that did recover by 2014, in years, was 2.02.

Conclusions

The stock market crash of 2007-2008 is infamous, and the results of this report reflect why. During this time, the average company net income was lower than in previous years, and some firms saw significant losses and massive percent decreases. The firms that were performing poorly were falling at a much greater rate than the rate at which the firms doing well were gaining. However, some firms did perform better in the crisis years, as seen in the analysis of questions 1 and 2, and others were able to bounce back quickly, as seen in question 3. Unfortunately, many firms took 6+ years to recover, and it is likely that a significant percentage of them never recovered, as recovery seemed to become less likely as time from the crisis increased. We know now that the market was able to eventually recover and even thrive, up until the COVID19 pandemic in March of 2020.