Effect of 2007-08 Financial Crisis on Corporate Profits

Khizar Tahir

Financial Crisis of 2007-2008

The Financial Crisis of 2007 - 2008 was caused primarily by the bursting of the housing market bubble and excessive risk-taking by banks [1]. It has often been addressed as the worst financial crisis since the Wall Street Crash of 1929, which led to the Great Depression.

The Compustat 1990-2015 Lots dataset provides 25 years of financial data from public companies to give users a big picture view of individual firms annual financial results, which can be analyzed to summarize performance over a period of time using several ratios and metrics.

The question we look to answer with our analysis is whether or not the financial crisis adversely affected the profitability of firms. We will do this primarily by utilizing the company ticker (tic), fiscal year (fyear) and net income (ni) variables within the dataset.

For the duration of this analysis, we will define our crisis timeline as follows:

Pre-crisis is defined as years 2004 - 2006

Crisis is defined as years 2007 - 2008

Post- Crisis is defined as years 2009 - 2014

Our analysis begins with loading the required packages.

For this assignment, we are using dplyr and tidyverse for our dataframes and scales for formatting. We then read our data in using read_csv.

First, we analyze the effect the crisis had on firms in financial terms by comparing average annual net income decreases for firms during the crisis compared to pre-crisis years.

Section 1.1: Calculate pre-crisis and crisis averages for each firm and combine them

Our approach to this was to create pre-crisis and crisis average functions that calculated the mean net income for both timelines.

We utilize the dplyr package to chain commands and group the data by the unique ticker of firms. Then, we filter the data to keep it only for the applicable time period, whether it be pre-crisis (pre-crisis_avg) or crisis (crisis_avg).

Then, we calculate the mean net income in those time periods, summarized as “pre-mean” for the pre-crisis period and “crisis_mean” for the crisis periods.

Finally, we create a dataframe “j” (for joining) and combine the two dataframes into a single one. A full_join ensures all the values from both dataframes are included. We omit an NA values this has.

```
pre_crisis_avg <- df %>% group_by(tic) %>%
  filter (fyear == 2004 | fyear == 2005 | fyear == 2006) %>%
  summarize(pre_mean = mean(ni, na.rm = TRUE)) %>%
  arrange(pre_mean) %>% drop_na()

crisis_avg <- df %>% group_by(tic) %>%
  filter (fyear == 2007 | fyear == 2008) %>%
  summarize(crisis_mean = mean(ni, na.rm = TRUE)) %>%
  drop_na()

j <- full_join(pre_crisis_avg,crisis_avg) %>% na.omit()

print(head(j))
```

```
## # A tibble: 6 x 3
##   tic   pre_mean crisis_mean
##   <chr>    <dbl>       <dbl>
## 1 VOD    -21014.       8812.
## 2 CBS     -7630.      -5213.
## 3 DAL     -5073       -3655
## 4 DLPH    -4191.        -14
## 5 CPN     -3982.       1352.
## 6 GM      -3247      -34796
```

The results of these averages show that there are 10,881 firms for which we can calculate the “pre-mean”, while there are only 9,606 firms for which we can calculate the “crisis_mean”. Because these numbers differ, we know there are firms for which only one period’s data is available only. An na.omit() on j removes such values and leaves us with only those firms (8,618) for which both periods of data are available making comparison accurate.

Section 1.2: Calculate percent by which annual net income decreases for firms during crisis years compared to pre-crisis

Here, we calculate pre-crisis and post crisis mean net income for ALL the firms combined.

Then, we calculate the percentage change “crisis_effect” using the formula ((crisis_mean/pre_mean) - 1)*100.

```
all_firms_pre_mean <- mean(pre_crisis_avg$pre_mean)
all_firms_crisis_mean <- mean(crisis_avg$crisis_mean)

crisis_effect <- percent(round((all_firms_crisis_mean - all_firms_pre_mean)/all_firms_pre_mean*100, 3),
 suffix = "%", scale = 1)

print(c("Percentage change in Net income:", crisis_effect))
```

```
## [1] "Percentage change in Net income:" "-16%"
```

Conclusion Part 1: There was an approximate decrease of 16% in firms’ net income from pre-crisis to crisis period.

Next, we utilize the results derived from Part 1 to visualize the 10 firms with the largest and smallest changes respectively, using both absolute change and percent change.

In this part, we want to know the magnitude of change experienced by firms. We are not concerned about who had a positive change or negative.

Section 2.1: effects based on absolute change

We first want to show the absolute change experienced by firms. We mutate our joined table “j” to insert a column “abs_dif” that calculates the absolute difference between crisis_mean and pre_crisis mean.

We arranged the data in a descending order of absolute difference. The 10 rows of head give us the companies that faced the greatest change, positive or negative. The 10 rows of tail give us the companies that faced the least change due to crisis.

```
f2_1 <- j %>%
  mutate(abs_dif = format(abs(crisis_mean - pre_mean), nsmall = 2, big.mark = ",", scientific = FALSE))
 %>%
  arrange(desc(abs_dif)) %>%
  select(tic, abs_dif) %>% drop_na()

print(head(f2_1, n=10))
```

```
## # A tibble: 10 x 2
##    tic   abs_dif
##    <chr> <chr>
##  1 AIG   57,999.1666666667
##  2 FNMA  35,502.8333333333
##  3 C     33,091.1666666667
##  4 GM    31,549.0000000000
##  5 VOD   29,826.1956666667
##  6 FMCC  29,032.5000000000
##  7 BAC2  23,378.1666666667
##  8 UBS   21,383.1850000000
##  9 RBS   19,924.7005000000
## 10 S     16,888.6666666667
```

```
print(tail(f2_1, n=10))
```

```
## # A tibble: 10 x 2
##    tic    abs_dif
##    <chr>  <chr>
##  1 IHT    "     0.0013333333"
##  2 VODG   "     0.0011666667"
##  3 ANML.1 "     0.0011666667"
##  4 CEHC   "     0.0008333333"
##  5 ASOE   "     0.0003333333"
##  6 AMRB   "     0.0001666667"
##  7 MDJT.1 "     0.0000000000"
##  8 2151B  "     0.0000000000"
##  9 4779B  "     0.0000000000"
## 10 FRO3   "     0.0000000000"
```

Results of Section 2.1

These results show that there is a great variability in the data. Some firms experienced huge effects, other very very minimal.

Section 2.2: effects based on absolute percentage change

Here, we do the same thing as above but arrange and show the absolute percentage difference each firm experienced.

```
f2_2 <- j %>%
  filter(pre_mean != 0) %>%
  mutate(abs_dif = format(abs(crisis_mean - pre_mean), nsmall = 2, big.mark = ",", scientific = FALSE))
 %>%
  mutate(pcnt_dif = format((abs(crisis_mean - pre_mean)/abs(pre_mean)*100), big.mark = ",", scientific =
FALSE)) %>%
  arrange(desc(pcnt_dif)) %>%
  select(tic, pcnt_dif) %>% na.omit()

print(head(f2_2, n=10))
```

```
## # A tibble: 10 x 2
##    tic    pcnt_dif
##    <chr>  <chr>
##  1 PGUS   "18,887,155,576,056,430,592.000000000"
##  2 STAQ   "                   141,800.000000000"
##  3 NHLD   "                   133,843.181818181"
##  4 ENTN   "                    98,988.800000000"
##  5 0419B  "                    95,867.307692308"
##  6 RAE    "                    69,827.659574468"
##  7 DEK    "                    69,700.000000000"
##  8 CBEY   "                    62,884.999999999"
##  9 3WCPSF "                    61,690.957446809"
## 10 FSCI   "                    60,839.682539682"
```

```
print(tail(f2_2, n=10))
```

```
## # A tibble: 10 x 2
##    tic    pcnt_dif
##    <chr>  <chr>
##  1 GAXIQ  "                         0.133377793"
##  2 KHD.Z  "                         0.107944732"
##  3 WASH   "                         0.107426979"
##  4 CEHC   "                         0.089573630"
##  5 MCO    "                         0.066099552"
##  6 SKH    "                         0.058508045"
##  7 GXP    "                         0.042697732"
##  8 UNT    "                         0.011302761"
##  9 AMRB   "                         0.002077016"
## 10 MDJT.1 "                         0.000000000"
```

Section 2.3: Results

It is worth noting that percentage changes are rather misleading. The biggest changes are for firms that had an extremely small net income in pre_crisis period and had some income in the crisis period. The absolute change is small, but the percentage change is HUGE since the denominator (pre_mean) is extremely small.

this work can be extended to remove very small values of pre-crisis mean, so percentage change can reflect what absolute changes are reflecting. We did not do it as it may be outside of scope of this project.

Next steps could also be to see what type of companies had the biggest changes. Were they in the same industry? Could use the Fama French 3.

Lastly, we determine how long it took for firms to recover from the crisis by comparing firms highest net incomes pre-crisis and post-crisis.

The first few statements below detail the algorithm developed used to evaluate this question, with analysis based on the results of the algorithm. Recovery in terms of this specific analysis can be defined as a firm “breaching” or, in other words, reaching and or beating it’s maximum net income score from the years of 2004 through 2006 at some point during the years between 2008 and 2014.

First, here is the overall algorithm used to evaluate the amount of recovery years for firms:

```
data4 <- df %>%
  select(tic, fyear, ni) %>%
  group_by(tic) %>%
  arrange(fyear, desc(ni)) %>%
  mutate(UN.ID = paste(tic, fyear, sep = "$$")) %>%
  group_by(UN.ID) %>%
  filter(row_number()==1) %>%
  filter(fyear >= 2004 & fyear <= 2014) %>%
  mutate(code = case_when(fyear <= 2006 ~ "pre-crisis",
                          fyear == 2007 ~ "crisis",
                          fyear == 2008 ~ "crisis",
                          fyear >= 2009 ~ "post-crisis")) %>%
  group_by(tic, code) %>%
  drop_na(ni) %>%
  ungroup() %>%
  filter(!code == "crisis") %>%
  group_by(tic, code) %>%
  mutate(maxni = max(ni)) %>%
  mutate(pre_max = case_when(code == "pre-crisis" ~ maxni)) %>%
  group_by(tic) %>%
  fill(pre_max) %>%
  mutate(sub_year = 2008) %>%
  filter(code == "post-crisis") %>%
  mutate(test = case_when(ni > pre_max ~ 1)) %>%
  filter(!is.na(pre_max)) %>%
  filter(!is.na(ni)) %>%
  group_by(tic) %>%
  arrange(desc(test), fyear) %>%
  filter(row_number()==1) %>%
  mutate(YearsRecover = ifelse(is.na(test), NA, fyear - sub_year)) %>%
  select(tic, YearsRecover)
```

Before analyzing any data with the algorithm, the first part of the dplyr statement includes cleaning the data.

```
data4 <- df %>%
  select(tic, fyear, ni) %>%
  group_by(tic) %>%
  arrange(fyear, desc(ni)) %>%
  mutate(UN.ID = paste(tic, fyear, sep = "$$")) %>%
  group_by(UN.ID) %>%
  filter(row_number()==1)
```

An example of why this cleaning is required is given below. Notice the duplicate fyear rows, with one duplicate containing a net income value and the other containing an NA as the net income:

```
##      tic fyear    ni
## 998  AET  1990 614.1
## 999  AET  1990    NA
## 1000 AET  1991    NA
## 1001 AET  1991 505.2
## 1002 AET  1992  56.0
## 1003 AET  1992    NA
```

Digging into the actual algorithm, our approach was to first filter on pre-crisis and post-crisis time periods and determine the max net income during the pre-crisis years. We then used the “group_by” function from the dplyr package and the “fill” function from the tidyr package to assign every ticker the max net income from the pre-crisis years for that ticker. Therefore, when the algorithm then filters the dataset to focus only on the post-crisis years, the pre-crisis years’ max net income is still available by ticker as a separate column.

```
  filter(fyear >= 2004 & fyear <= 2014) %>%
  mutate(code = case_when(fyear <= 2006 ~ "pre-crisis",
                          fyear == 2007 ~ "crisis",
                          fyear == 2008 ~ "crisis",
                          fyear >= 2009 ~ "post-crisis")) %>%
  group_by(tic, code) %>%
  drop_na(ni) %>%
  ungroup() %>%
  filter(!code == "crisis") %>%
  group_by(tic, code) %>%
  mutate(maxni = max(ni)) %>%
  mutate(pre_max = case_when(code == "pre-crisis" ~ maxni)) %>%
  group_by(tic) %>%
  fill(pre_max) %>%
  mutate(sub_year = 2008) %>%
  filter(code == "post-crisis")
```

The actual “decision making” of the algorithm then occurs, via a new column named “test”. This column associates a one with every year in the post-crisis years where net income was greater than the maximum net income from the pre-crisis years for each ticker. Using this test column, the algorithm then determines how long it took a company to breach that maximum net income again post-crisis by identifying the first fiscal year in which the net income in the post-crisis years exceeded the pre-crisis maximum. Then, using 2008 as the starting point, the algorithim defines the “YearsRecovered” variable.

```
  mutate(test = case_when(ni > pre_max ~ 1)) %>%
  filter(!is.na(pre_max)) %>%
  filter(!is.na(ni)) %>%
  group_by(tic) %>%
  arrange(desc(test), fyear) %>%
  filter(row_number()==1) %>%
  mutate(YearsRecover = ifelse(is.na(test), NA, fyear - sub_year)) %>%
  select(tic, fyear, ni, pre_max, YearsRecover)
```

An example of the resulting dataframe (arranged) from the algorithm is given below:

```
## # A tibble: 5 x 2
## # Groups:   tic [5]
##   tic   YearsRecover
##   <chr>        <dbl>
## 1 GM               1
## 2 OGZPY            1
## 3 CHL              1
## 4 PBR              1
## 5 BCS              1
```

A small test for a firm that could not recover (ADCT) is below:

```
print(data4$YearsRecover[which(data4$tic == 'ADCT')])
```

```
## [1] NA
```

The results of algorithm are summarized into the following table:

```
## # A tibble: 7 x 2
## # Groups:   YearsRecover [7]
##   YearsRecover Number.of.Firms
##          <dbl>           <dbl>
## 1            1            2318
## 2            2            1045
## 3            3             570
## 4            4             351
## 5            5             262
## 6            6             189
## 7           NA            2636
```

From this table, we derive some basic statistics about the data set related to the financial crisis in the United States in 2008, given by this simple print statement:

```
## [1] "The percentage of firms that recovered between 2008 and 2014 was: 64%. The average number of yea
rs from 2008 that firms took to recover was: 2.1 years. Finally, the standard deviaiton of years recover
ed for firms that did indeed recover was: 1.43 years."
```

Based on these summary statistics, we can conclude that, on average, it tooks firms a little over two years to recover from the crisis, if they did in fact end up recovering in the first place. Of the firms that had both pre-crisis and post-crisis years data, 2,636 or 36 percent did not recover before 2014. The standard deviation of the amount of recovery years was 1.43 years, meaning about 70 percent of the firms that did recover took between 0.67 years and 3.53 years to recover.

In conclusion, did the crisis affect corporate profits?

The case for the financial market crisis in the United States between the years of 2007 and 2008 as being a difficult time for firms and ultimately having an overall negative effect on the economy is strengthened by observing that, for firms with pre-crisis and post-crisis years data, about 36 percent of firms never reached a recovery point before 2014. Subsequently, the average recovery period for firms that did recover was 2.1 years with a standard deviation of 1.43 years. This data, coupled with information from answers from #1 and #2 suggest that by-in-large, struggled to recoup their previous positions in the market by reaching the same maximum net income levels in the post-crisis years that they had once achieved in the pre-crisis years.

A future expansion on this analysis would be to include the analysis of the similarities and differences between firms with similar net income percent changes before and after the crisis, as well as equal recovery times. This analysis may give insight into competitive advantages that certain firms and industries possed going into and ultimately surving a financial crisis, or slump in general.

Sources:

[1] 'Financial Crisis of 2007-2008' (2020) *Wikipedia* . Available at:
https://en.wikipedia.org/wiki/Financial_crisis_of_2007%E2%80%932008
(https://en.wikipedia.org/wiki/Financial_crisis_of_2007%E2%80%932008)
(Accessed: 15 September 2020)