FIN 557 Project Report

Impact of Covid on Bankruptcy

Group17: Jason Chang Chien, Wei-Yu Chen, Ching-Che Chi, Pranav Ganapathy Subramanian, Gamanash Kommineni, Ana Telia, Sabelo Zwane

Problem Statement

1. Analyzing the financial data from WRDS during the FY 2017-2023 to assess the impact of COVID on bankruptcy
2. Analyzing z-score to evaluate its performance for predicting the bankruptcy

Introduction

In this analysis, we will be examining the impact of the COVID-19 pandemic on the bankruptcy of firms in the United States for the years 2017 to 2023. Specifically, we will be looking at which states have been the most affected by COVID-19 in terms of bankruptcy rates and which industries have been hit the hardest. By analyzing this data, we hope to gain a better understanding of the long-term economic effects of the pandemic and how it has affected different sectors of the economy in different regions of the country. We are also calculating a metric called Z-Score using the required variables and checking how precise the Z-Score value is indicating whether a company is safe or on-alert.

Data sets and sources

- **Data1 - Bankruptcy data (WRDS)**: https://wrds-www.wharton.upenn.edu/pages/get-data/audit-analytics/corporate-legal/bankruptcy-notification/?no_login_redirect=True
- **Data2 - Financial Statements (WRDS):** https://wrds-www.wharton.upenn.edu/pages/get-data/compustat-capital-iq-standard-poors/compustat/north-america-daily/fundamentals-annual/
- **Data3 - Share Price and Shares Outstanding (WRDS):** https://wrds-www.wharton.upenn.edu/pages/get-data/center-research-security-prices-crsp/annual-update/stock-security-files/daily-stock-file/

Methodology

We use SAS studio for our analysis, R for regression analysis and since the data was downloaded from WRDS in SAS format, no preprocessing was necessary. The data was analyzed by year, industry and state.

Problem Statement I

1. Total Bankruptcy count for each year from 2017-2023:

```
/*Calculate the total bankruptcy count for each year 2017-2023*/
proc sql;
select YEAR(BANK_BEGIN_DATE) as YEAR,
       count(COMPANY_FKEY) as bankruptcy_count_year
from data
where LOC_STATE_COUNTRY='USA'
group by calculated YEAR;
quit;
```

YEAR_bankrupted	bankruptcy_count_year
2017	52
2018	38
2019	54
2020	83
2021	19
2022	25
2023	7

In this code, we use the PROC SQL procedure to perform a SQL query on the data table where we have the bankruptcy data. We use the COUNT function to count the number of distinct company key values in the table that meet the condition specified in the WHERE clause: specifically, we only count firms that filed for bankruptcy in the USA. We give the resulting count an alias of bankruptcy_count_year using the AS keyword. Finally, we end the query with the QUIT statement.

From the output, it is clear that during Covid more number of firms filed for bankruptcy as we hypothesized.

2. Calculate the bankruptcy count for each SIC each year from 2017-2023:

```
/*Calculate the bankruptcy count for each SIC each year 2017-2023*/
proc sql;
create table SIC_rank as
select substr(SIC_CODE_FKEY,1,3) as SIC,
       YEAR(BANK_BEGIN_DATE) as YEAR,
       count(COMPANY_FKEY) as bankruptcy_count_year,
       SIC_CODE_DESCRIP
from data
where LOC_STATE_COUNTRY='USA'
group by calculated SIC, calculated YEAR
order by calculated YEAR, bankruptcy_count_year desc;
quit;

/*Print only the most bankrupcy observations for every year*/
data SIC;
set SIC_rank;
by YEAR;
if first.YEAR;
run;

proc print data=SIC;
run;
```

Obs	SIC	YEAR_bankrupted	bankruptcy_count_year	SIC_CODE_DESCRIP
1	131	2017	4	Crude Petroleum and Natural Gas
2	131	2018	9	Crude Petroleum and Natural Gas
3	131	2019	11	Crude Petroleum and Natural Gas
4	131	2020	15	Crude Petroleum and Natural Gas
5	590	2021	4	Retail-Miscellaneous Retail
6	283	2022	6	Pharmaceutical Preparations
7	283	2023	2	Biological Products, Except Diagnostic Substances

Overall, this SAS code performs data aggregation, filtering, and sorting operations to extract meaningful insights about what kind of industries were impact during covid from the input data set.

3. Calculate which state has highest bankruptcy percentage for each year from 2017-2023 :

Obs	State	YEAR_bankrupted	bankruptcy_count_state	bankruptcy_count_year	bankruptcy_pct
1	TX	2017	15	52	28.85%
2	TX	2018	9	38	23.68%
3	TX	2019	19	54	35.19%
4	TX	2020	27	83	32.53%
5	MO	2021	4	19	21.05%
6	CA	2022	5	25	20.00%
7	CA	2023	2	7	28.57%

Results:

- 2017 to 2020: Texas was the most impacted state, which accounted for about 24% to 35%
- 2021: Missouri was the most impacted state, which accounted for 21%
- 2022 to 2023: California was the most impacted state, which accounted for 29%

4. Calculate percentage of number of firms bankrupt to total number of firm percentage for year 2020:

State/Province	Data Year - Fiscal	Frequency Count	Ratio_firm_bankrupt_pct
OK	2020	4	5.6
TN	2020	4	4.4
NV	2020	4	4.4
TX	2020	27	4.1
LA	2020	1	2.3
CO	2020	5	2.3
FL	2020	7	2.1
WI	2020	2	1.9
OH	2020	3	1.5
NJ	2020	3	1.1
MO	2020	1	1.1
AZ	2020	1	0.9
CT	2020	1	0.7

State/Province	Data Year - Fiscal	Frequency Count	Ratio_firm_bankrupt_pct
NC	2020	1	0.7
PA	2020	2	0.5
NY	2020	7	0.5
VA	2020	1	0.5
CA	2020	7	0.4
IL	2020	2	0.2

Results:

- For 2020, Texas not only has the greatest percentage of corporations that filed for bankruptcy but also the ratio of companies that went bankrupt to the total number of companies is among top 5 within the country (4.1% of publicly traded companies with in texas)

5. Linear regression between COVID-19 and bankruptcy by State in 2020

In this part, we export the SAS state_bank table which we created in the previous section, then we select the number of bankruptcy which YEAR equals to 2020 and add a new variable called Covid_case which is the number of COVID-19 cases in each state in 2020 from CDC official website.

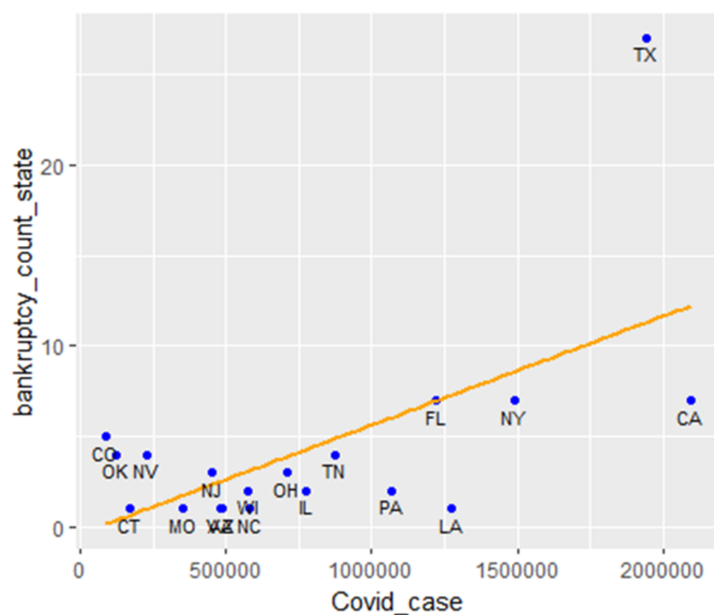
```
Call:
lm(formula = bankruptcy_count_state ~ Covid_case, data = state_bank)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2737 -1.8603 -0.8948  0.4923 15.6994

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.753e-01  1.883e+00  -0.199   0.8444
Covid_case   6.013e-06  1.930e-06   3.115   0.0063 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.828 on 17 degrees of freedom
Multiple R-squared:  0.3634,    Adjusted R-squared:  0.3259
F-statistic: 9.704 on 1 and 17 DF,  p-value: 0.006297
```

We run the linear regression to analyze the relationship between the number of COVID-19 cases in each state in 2020 and the number of bankruptcy cases in each state in 2020. The result shows that the p-value for the Covid_case variable is 0.0063, suggesting that there is a strong evidence of statistically significant relationship between COVID-19 cases and bankruptcy cases.



In this scatter plot, we found that there is a positive relationship between COVID-19 cases and bankruptcy number in each state. As the number of Covid cases increases in a state, the number of firms that went bankrupt also tends to increase.

Problem Statement II

In this part, we have to evaluate the performance of the Z-score. Z-score is a weighted average accounting ratios showed as following:

Z-Score = A x 3.3 + B x 0.99 + C x 0.6 + D x 1.2 + E x 1.4

Where,

- A=EBIT/Total Assets
- B=Net Sales /Total Assets
- C=Market Value of Equity / Total Liabilities
- D=Working Capital/Total Assets
- E=Retained Earnings /Total Assets

Preparation for Z-score calculation

To calculate the Z-score for each company, we need the above variables.

(1) First, we have to merge the financial data with bankruptcy data. We use the bankruptcy data (data) as the left table to join the financial statement data (data2) as the right table to perform a left join, while using Ticker as the key for two tables. The merged table is named “merge”.

```
/* Join data and data2 as merge*/
proc sql;
create table merge as
select *,
       YEAR(d.BANK_BEGIN_DATE) as YEAR_bankrupted
from data d left join
     data2 d2
on d.BEST_EDGAR_TICKER=d2.TIC
where d2.TIC is not null and
      d.BEST_EDGAR_TICKER is not null;
quit;

proc print data=merge;
run;
```

(2) Second, to calculate the C in Z-score which is Market Value of Equity / Total Liabilities. We have to extract the data from data3 and calculate the market value of equity (share price * shares outstanding). We averaged the monthly share price and shares outstanding to use as the annual data. Furthermore, we perform a proc sort procedure to eliminate duplicate rows. The final table is named “market2”.

(3) Based on the two tables, we then merge the “Merge” table with “Market2” table. We use “Market2” as the left table to join the “Merge” as the right table to perform a left join, while using Ticker and Year as the keys for two tables. The merged table is named “merge2”.

```
/* Join merge and market table as merge2*/
proc sql;
create table merge2 as
select m.YEAR_bankrupted, m2.*,
       m.EBIT, m.AT,
       m.SALE, m.LT,
       m.WCAP, m.RE
from market2 m2 left join
     merge m
     on m2.YEAR=m.FYEAR and
     m2.TICKER=m.TIC
where m.FYEAR is not null and
      m.TIC is not null and
      m.EBIT is not null and
      m.AT is not null and
      m.SALE is not null and
      m.LT is not null and
      m.WCAP is not null and
      m.RE is not null;
quit;

proc print data=merge2 (obs=50);
run;
```

The output table merge2:

Obs	YEAR_bankrupted	YEAR	COMNAM	TICKER	average_stock_price	average_shrout	market_value	EBIT	AT	SALE	LT	WCAP	RE
1	2017	2017	ADAMS RESOURCES & ENERGY INC	AE	40.515	4218.00	170892.27	1.5020	282.7040	1322.0600	135.5850	116.0870	135.0040
2	2019	2017	AVADEL PHARMACEUTICALS PLC	AVDL	9.646	40686.33	392453.59	59.6110	253.2770	171.8030	167.6970	37.4360	-285.9510
3	2020	2017	CHESAPEAKE ENERGY CORP	CHK	4.830	905443.50	4373292.13	1519.0000	12425.0000	9496.0000	12797.0000	-831.0000	-16582.0000
4	2017	2017	CIVITAS SOLUTIONS INC	CIVI	18.067	37324.75	674333.82	26.3110	830.3710	192.1240	142.0370	-36.0660	-5.0200
5	2017	2017	CUMULUS MEDIA INC	CMLS	0.481	29226.00	14045.43	142.4560	2027.3190	1135.6620	2723.4340	357.4630	-2093.5540
6	2020	2017	CALIFORNIA RESOURCES CORP	CRC	13.171	42573.00	560721.89	68.0000	6207.0000	2006.0000	6927.0000	-249.0000	-5693.0000
7	2020	2017	DIAMOND OFFSHORE DRILLING INC	DO	14.697	137209.58	2016637.84	226.8380	6250.5700	1485.7460	2476.3090	663.3650	1964.4920
8	2018	2017	FIRSTENERGY CORP	FE	31.328	442488.58	13862429.92	2891.0000	42257.0000	14012.0000	38332.0000	-969.0000	-6120.0000
9	2020	2017	GULFPORT ENERGY CORP	GPOR	14.930	180911.42	2701007.46	548.3660	5807.7520	1320.3030	2706.1380	-221.3560	-1316.4670
10	2020	2017	HELIOS & MATHESON ANALYTICS INC	HMNY	5.575	8940.42	49842.82	-49.0680	164.0330	10.4420	164.6260	-107.0980	-189.5990
11	2019	2017	LIVE VENTURES INC	LIVE	13.912	2015.25	28037.15	18.4480	128.5950	152.0610	95.0110	-10.8930	-28.5760
12	2020	2017	MALLINCKRODT PLC	MNK	40.073	100412.83	4023876.92	508.6000	15280.9000	3221.6000	8758.9000	1234.8000	2575.7000
13	2019	2017	P G & E CORP	PCG	63.317	511266.83	32371711.79	3282.0000	68012.0000	17135.0000	48540.0000	-848.0000	6588.0000
14	2017	2017	TIDEWATER INC NEW	TDW	11.511	36208.92	416790.58	-69.8730	1746.1800	330.1220	724.2360	625.0520	-39.4130
15	2017	2017	TIDEWATER INC	TDW	11.511	36208.92	416790.58	-69.8730	1746.1800	330.1220	724.2360	625.0520	-39.4130

Consequently, we derive the variables that are needed when calculating the Z-score (A,B,C,D,E) from the table.

Z-score Calculation and Evaluation

After collecting the necessary data, we created and calculated Variables needed for the Z-score. We created a new table zscore including the following Variables: EBIT/AT, SALE/AT, market value/LT/1000, WCAP/LT and RE/AT. We calculated Z score and stored the outcome in a new table called zscore2. We ordered the result by Ticker and Year.

```
/* Create Variables for Z-score */
proc sql;
create table zscore as
select *,
       EBIT/AT as A,
       SALE/AT as B,
       market_value/LT/1000 as C,
       WCAP/LT as D,
       RE/AT as E
from merge2;
quit;

/* Calculate the Z-score for each company that got bankrupted*/
proc sql;
create table zscore2 as
select YEAR, TICKER, YEAR_bankrupted, COMNAM, A, B, C, D, E,
       3.3*A+0.99*B+0.6*C+1.2*D+1.4*E as z_score,
       3.3*A+0.99*B+0.6*C+1.2*D+1.4*E as z_score_category
from zscore
order by TICKER, YEAR;
quit;
```

We created a format zscorefmt to assign descriptions accordingly.

```
/* format Z-score with according Interpretation */
proc format;
    value zscorefmt  low-<1.8   = 'Probability of Financial distress is very high'
                     1.8-<2.7   = 'Good chances of going bankrupt within 2 years'
                     2.7-<2.99  = 'On Alert'
                     3-high     = 'Safe';
run;

proc print data=zscore2;
    var YEAR TICKER YEAR_bankrupted COMNAM z_score z_score_category;
    format  z_score_category zscorefmt.;
run;
```

Finally, we created a table zscore3 and checked if the Z-score prediction was consistent with reality in the year the firms went bankrupt. We filtered the year of bankruptcy using having YEAR=YEAR_bankrupted option.

```
proc sql;
create table zscore3 as
select YEAR, TICKER, YEAR_bankrupted, COMNAM, z_score,
       z_score_category format zscorefmt.
from zscore2
group by TICKER
having YEAR=YEAR_bankrupted;
quit;

proc print data=zscore3;
run;
```

Obs	YEAR	TICKER	YEAR_bankrupted	COMNAM	z_score	z_score_category
1	2017	AE	2017	ADAMS RESOURCES & ENERGY INC	7.09949	Safe
2	2019	AVDL	2019	AVADEL PHARMACEUTICALS PLC	-3.03505	Probability of Financial distress is very high
3	2020	CHK	2020	CHESAPEAKE ENERGY CORP	-8.45721	Probability of Financial distress is very high
4	2017	CIVI	2017	CIVITAS SOLUTIONS INC	2.86901	On Alert
5	2017	CMLS	2017	CUMULUS MEDIA INC	-0.49868	Probability of Financial distress is very high
6	2020	CRC	2020	CALIFORNIA RESOURCES CORP	-1.43959	Probability of Financial distress is very high
7	2020	DEN	2020	DENBURY INC	-0.88550	Probability of Financial distress is very high
8	2020	DO	2020	DIAMOND OFFSHORE DRILLING INC	0.25051	Probability of Financial distress is very high
9	2018	FE	2018	FIRSTENERGY CORP	0.56245	Probability of Financial distress is very high
10	2020	GPOR	2020	GULFPORT ENERGY CORP	-3.88882	Probability of Financial distress is very high
11	2020	GTX	2020	GARRETT MOTION INC	0.30203	Probability of Financial distress is very high
12	2019	LIVE	2019	LIVE VENTURES INC	1.71266	Probability of Financial distress is very high
13	2020	MNK	2020	MALLINCKRODT PLC	-0.31351	Probability of Financial distress is very high
14	2019	PCG	2019	P G & E CORP	0.27611	Probability of Financial distress is very high
15	2017	TDW	2017	TIDEWATER INC	1.40447	Probability of Financial distress is very high
16	2017	TDW	2017	TIDEWATER INC NEW	1.40447	Probability of Financial distress is very high
17	2019	TPX	2019	TEMPUR SEALY INTERNATIONAL INC	3.07852	Safe
18	2019	TREE	2019	LENDINGTREE INC	5.33972	Safe
19	2020	TT	2020	TRANE TECHNOLOGIES PLC	3.22398	Safe
20	2020	VAL	2020	VALARIS PLC	-0.41223	Probability of Financial distress is very high
21	2020	WLL	2020	WHITING PETROLEUM CORP NEW	-6.28725	Probability of Financial distress is very high

Conclusion

After processing and analyzing the datasets, we conclude that Covid-19 had a negative impact on firm performance and pandemics have increased the number of firms, which have declared bankruptcy. This result is consistent with our expectations. We found out that the industry impacted the most was the Crude Petroleum and Natural Gas industries. Furthermore, we conclude that Texas was the state impacted the most heavily. In order to straighten our argument, that the number of bankruptcies has increased due to the pandemics, we have checked the correlation between the total number of Covid-19 cases in 2020 and the number of bankruptcies and found out that there is a significant positive relation. Moreover, Texas is a clear outlier. Furthermore, we have calculated the ratio of bankrupt firms to the total number of firms in each state, which shows that for instance in Texas 4.1% of all publicly traded companies went bankrupt in 2020.

In the second part of our analysis, we have estimated the Z-score for firms, which went bankrupt. We have concentrated on the year of bankruptcy of each firm and checked what was the status of these firms according to the Z-score in the year the firms went bankrupt. According to our analysis, Z-score is consistent with the reality in 16 out of 21. In 1 case it is close to consistency and in 4 cases the Z-score is deviating from reality. Since the sample size is small we can not estimate statistically how precise the Z-score is. However, we believe Z-score performs quite well.