FIN 557 Project Report
Stock Valuation Across Industries
4/25/23

By: Tzu-Mien Liu, Tung-Han Lee, Pratik Relekar, Amin Azimi, Hinduja Hanumanthappa, Gurpratap Singh Thind, Xiying Zhao

Introduction

The project involves examining how stocks perform in diverse industries and using this analysis to devise an investment strategy. As a part of this initiative, we selected 50 companies from five industries and evaluated whether each stock is undervalued, fairly valued, or overvalued in its respective industry. The primary goal is to identify the most undervalued stocks across all industries and create an investment portfolio using these stocks.

Data source

WRDS to collect annual and quarterly data for 10 companies in each of the 5 sectors spanning from January 2010 to April 2023.

Links -
https://wrds-www.wharton.upenn.edu/pages/get-data/compustat-capital-iq-standard-poors/compustat/north-america-daily/fundamentals-quarterly/

https://finance.yahoo.com/screener/predefined/undervalued_growth_stocks/

Research question

The primary goal of this project is to analyze 50 sample stocks from 5 different industries and develop an investment portfolio that is optimal.

Approach

Firstly, we categorized the 50 stocks into three groups according to their average Price-to-Book ratio (PB ratio) for the period spanning from 2015 to 2017. The PB ratio is calculated by dividing a share's market price to its book value.
Secondly, we determined the average returns of all companies in each of the three groups over the subsequent five-year period from 2018 to 2022.
Finally, we calculated the number of stocks in each group and their respective average returns within each industry for the 2018-2022 period.

Data Manipulation and Analysis

Analysis 1: Based on P/B ratio

1. Variables constructed

    - Outstanding: Calculate the number of shares outstanding by dividing the total market value by stock price.
    - PB_ratio: Calculate the P/B ratio by dividing the market Capitalization by Book value of Equity.
    - Year: Extract the year from DATADATE.
    - Quarter: Extract the quarter from DATADATE
    - Stock_group: Define whether the stock is value stock or growth stock based on its P/B ratio.

2. Data manipulation and techniques applied

    - Filter to only include raw data in 2015, 2016 and 2017.
    - Calculate the average PB ratio over a three year period.
    - We assume that the stocks in the bottom 25% of the PB ratio are "value stocks". We assume that the stocks in the top 25% of the PB ratio are "growth stocks". Others are defined as "general stocks".
    - Filter to only include raw data in 2018 to 2022.
    - Merge stock_group and stock price(after 2017) by company name.
    - Calculate the mean return of all companies in each group over the 5-year period.

/*Create new colums: Outstanding, PB_ratio, Year, QUARTER*/
/*Only select positive PB_ratio*/
Code -

```
data stock1;
   set project.stock;
   Outstanding = MKVALTQ/PRCCQ;
   PB_ratio = PRCCQ*Outstanding/CEQQ;
   Year = year(DATADATE);
   QUARTER = qtr(DATADATE);
   if PB_ratio > 0;
run;
```

/*Filter three years' data to calculate average PB_ratio*/
Code -

```
data stock2;
set stock1;
keep GVKEY DATADATE YEAR QUARTER SIC CONM PRCCQ PB_ratio;
where YEAR in (2015,2016,2017);
run;
```

/*Calculate three years average PB ratio of different companies*/
Code -

```
proc means data=stock2;
      var PB_ratio;
      class CONM;
      ways 1;
      output out=PBmean mean=AvgPB;
run;
```

			Analysis Variable : PB_ratio			
Company Name	N Obs	N	Mean	Std Dev	Minimum	Maximum
ABBOTT LABORATORIES	12	12	2.9212215	0.2304514	2.4572231	3.2206158
ADOBE INC	12	12	7.2900269	1.5495808	5.7274303	10.5379073
AMAZON.COM INC	12	12	19.6258849	2.0860711	15.9476317	23.7854296
AMERICAN EXPRESS CO	12	12	3.3990029	0.5277738	2.7135477	4.6802705
AMGEN INC	12	12	4.1137618	0.3869517	3.6127940	4.9756579
APPLE INC	12	12	5.1781704	0.6557187	4.0744626	6.1339082
BANK OF AMERICA CORP	12	12	0.8237361	0.2082835	0.5605840	1.2404111
BIOGEN INC	12	12	5.8930964	1.2122449	4.6465073	8.5133680
BRISTOL-MYERS SQUIBB CO	12	12	7.0269820	0.9095007	5.7716468	8.5231445
CAPITAL ONE FINANCIAL CORP	12	12	0.8705902	0.1240814	0.6678992	1.0896682
CHEVRON CORP	12	12	1.3143445	0.1880502	0.9583188	1.6097975
CISCO SYSTEMS INC	12	12	2.3914415	0.1665785	1.9744555	2.6044150
CITIGROUP INC	12	12	0.7726264	0.1375904	0.5792029	1.0536674
CONOCOPHILLIPS	12	12	1.6239784	0.2522495	1.2758025	2.1110009
COSTCO WHOLESALE CORP	12	12	6.2949281	0.7191762	5.5082083	7.9966510
CVS HEALTH CORP	12	12	2.6769663	0.3992317	1.9504656	3.1582363
DOLLAR TREE INC	12	12	4.5616321	1.8222483	2.9126214	8.4219275
EOG RESOURCES INC	12	12	3.5857949	0.5821383	2.7587687	4.5183884
EXXON MOBIL CORP	12	12	2.0091847	0.1447219	1.8129632	2.2785901
GILEAD SCIENCES INC	12	12	6.6484262	2.2192330	4.0763187	10.5887419
GOLDMAN SACHS GROUP INC	12	12	1.1203966	0.1804934	0.8412385	1.4075318
HALLIBURTON CO	12	12	3.5275058	1.2424112	1.8845065	5.1265934
HEWLETT PACKARD ENTERPRISE	9	9	0.9939345	0.1512933	0.7630990	1.1973220

/*Find p25 and p75 of companies' average PB_ratio*/
Code -

```
proc univariate data= PBmean;
        var AvgPB;
run;
```

Quantiles (Definition 5)	
Level	Quantile
100% Max	31.859395
99%	31.859395
95%	19.625885
90%	9.383415
75% Q3	6.092965
50% Median	3.527506
25% Q1	2.009185
10%	1.026918
5%	0.870590
1%	0.772626
0% Min	0.772626

/*Average PB_ratio p25 =2.009185 and p75 = 6.092965*/
/*We assume that the bottom 25% of the stocks in the PB ratio are "value stocks" */
/*We assume that the top 25% of the stocks in the PB ratio are "growth stocks" */
/*Others are "general stocks"*/
/*Divide all stocks into three groups(based on p25 and p75)*/
Code -

```
data stock4;
      set PBmean;
      length stock_group $14;
      keep  CONM AvgPB stock_group;
      where AvgPB is not missing;
      if AvgPB=<2.009185 then do;
    stock_group="value stock";
              output;
      end;
      else if AvgPB>6.092965 then do;
    stock_group="growth stock";
              output;
      end;
   else do;
      stock_group="general stock";
              output;
      end;
run;
```

Obs	CONM	AvgPB	stock_group
1	ABBOTT LABORATORIES	2.9212	general stock
2	ADOBE INC	7.2900	growth stock
3	AMAZON.COM INC	19.6259	growth stock
4	AMERICAN EXPRESS CO	3.3990	general stock
5	AMGEN INC	4.1138	general stock
6	APPLE INC	5.1782	general stock
7	BANK OF AMERICA CORP	0.8237	value stock
8	BIOGEN INC	5.8931	general stock
9	BRISTOL-MYERS SQUIBB CO	7.0270	growth stock
10	CAPITAL ONE FINANCIAL CORP	0.8706	value stock

/*Filter five years' data*/
Code -

```
data stock_after_2017;
set stock1;
keep GVKEY DATADATE YEAR QUARTER CONM PRCCQ;
where YEAR in (2018,2019,2020,2021,2022);
run;
```

/*Sort by company name*/
Code -

```
proc sort data=stock4;
by CONM;
run;
```

```
proc sort data=stock_after_2017;
by CONM;
run;
```

/*Merge stock group and stock price(after 2017) by company name*/
Code -

```
data stock_price;
merge stock4 stock_after_2017;
by CONM;
run;
```

```
proc sort data=stock_price;
by CONM;
run;
```

/*Get stock price in 2018 and 2022 to calculate stock return*/
Code -

```
data stock_return;
set stock_price;
```

by CONM;
if first.CONM = 1 then output;
else if last.CONM = 1 then output;
run;

/*Calculate stock return*/
Code -

```
data stock_return2;
        set stock_return;
        by CONM Year;
        PRCCQ2=lag(PRCCQ);
        if first.CONM then PRCCQ2=.;
        ret=(PRCCQ-PRCCQ2)/PRCCQ2;
run;
```

/*Keep non-missing return*/
Code -

```
data stock_return3;
set stock_return2;
where ret~=.;
run;
```

Obs	CONM	AvgPB	stock_group	GVKEY	DATADATE	PRCCQ	Year	QUARTER	PRCCQ2	ret
1	ABBOTT LABORATORIES	2.9212	general stock	001078	2022-12-31	109.7900	2022	4	59.92	0.83228
2	ADOBE INC	7.2900	growth stock	012540	2022-11-30	344.9300	2022	4	209.13	0.64936
3	AMAZON.COM INC	19.6259	growth stock	064768	2022-12-31	84.0000	2022	4	1447.34	-0.94196
4	AMERICAN EXPRESS CO	3.3990	general stock	001447	2022-12-31	147.7500	2022	4	93.28	0.58394
5	AMGEN INC	4.1138	general stock	001602	2022-12-31	262.6400	2022	4	170.48	0.54059
6	APPLE INC	5.1782	general stock	001690	2022-12-31	129.9300	2022	4	167.78	-0.22559
7	BANK OF AMERICA CORP	0.8237	value stock	007647	2022-12-31	33.1200	2022	4	29.99	0.10437
8	BIOGEN INC	5.8931	general stock	024468	2022-12-31	276.9200	2022	4	273.82	0.01132
9	BRISTOL-MYERS SQUIBB CO	7.0270	growth stock	002403	2022-12-31	71.9500	2022	4	63.25	0.13755
10	CAPITAL ONE FINANCIAL CORP	0.8706	value stock	030990	2022-12-31	92.9600	2022	4	95.82	-0.02985

/*Get average returns of all companies in each group*/
Code -

```
proc means data=stock_return3;
```

```
var ret;
class stock_group;
ways 1;
run;
```

The MEANS Procedure

			Analysis Variable : ret			
stock_group	N Obs	N	Mean	Std Dev	Minimum	Maximum
general stock	23	23	0.3933174	0.5601166	-0.4925115	1.6275885
growth stock	13	13	0.6933160	1.1253475	-0.9419625	3.7284477
value stock	13	13	0.2740693	0.3809828	-0.3299259	0.9902176

Key observations: Growth stocks had the highest return over the past five years, while value stocks had the lowest return of the three groups. This result does not meet our expectations. We speculate that the opposite trend was caused by the pandemic.

Analysis 2: Based on industry

In the previous analysis, we examined the average PB ratios of 50 stocks between 2015-2017 and their average returns over 5 years (2018-2022). Moving forward, we aim to incorporate the industry of the companies into our analysis and assess the stock performance across various industries.
1. Variables Constructed:
    - Sector: define the industry of company based on the company list(see Appendix)
2. Data Manipulation with Techniques Applied:
    - Add a new column named "Sector" to define industry using case…when… sentence in sql
    - Combine the data using merge sentence in SAS data step

 # Create a new column: "Sector" to define industry
 # In our case, we have five industries and each industry contains 10 companies
 # Define the sectors and assign TIC to them

Code -

```
proc sql;
create table sector_data as
select distinct TIC,CONM,
```

```
case
 when TIC in
('AAPL','MSFT','IBM','INTC','CSCO','ORCL','HPE','DELL','ADBE','TXN') then
'Technology'
 when TIC in ('XOM', 'CVX', 'COP', 'HAL','OXY','MPC','VLO','EOG','SLB','WMB')
then   'Energy'
 when TIC in ('WMT','AMZN','HD','COST','TGT','LOW','CVS','WBA','DLTR','TJX')
then 'Retail'
 when TIC in ('JNJ','PFE','MRK','BMY','LLY','ABT','AMGN','UNH','GILD','BIIB')
then  'Healthcare'
 when TIC in ('JPM','BAC','WFC','C','GS','AXP','MS','V','MA','COF') then 'Finance'
 end as Sector
 from stock1;
quit;
```

Check if all assigned well

Code -

```
proc sql;
select Sector, count(distinct TIC) as TIC_count
from sector_data
group by Sector;
quit;
```

Output -

Sector	TIC_count
Energy	10
Finance	10
Healthcare	10
Ratail	10
Technology	10

Sort stock1 and sector_data by CONM

Code -

```
proc sort data=stock1;
by CONM;
run;

proc sort data=sector_data;
by CONM;
run;
```

 # Merge the sector data with the main data

Code -

```
data main_data;
merge stock1 sector_data;
by CONM;
run;
```

3. Data Analysis:
   - The number of stocks of different sectors in each stock group (2015-2017)

 # Stock average performance in different industries

Code -

```
proc sql;
select stock_group, sector, count(*) as stock_number
from stock_group_sector
group by stock_group, sector
order by stock_group, stock_number desc;
quit;
```

Output -

stock_group	Sector	stock_number
general stock	Healthcare	7
general stock	Energy	5
general stock	Technology	5
general stock	Ratail	5
general stock	Finance	1
growth stock	Ratail	5
growth stock	Healthcare	3
growth stock	Technology	3
growth stock	Finance	2
value stock	Finance	7
value stock	Energy	5
value stock	Technology	1

Key observations: The retail sector boasts the highest number of growth stocks (5), while the finance sector has the highest number of value stocks (7). Based on this, we can predict that the finance sector is likely to outperform other sectors in the next five years since it has the largest concentration of value stocks.

- The average return ratio of each industry (2018-2022)

#Get average returns of each industry

Code -

```
proc means data=stock_industry;
var ret;
class Sector;
ways 1;
run;
```

Output -

The MEANS Procedure

			Analysis Variable : ret			
Sector	N Obs	N	Mean	Std Dev	Minimum	Maximum
Energy	10	10	0.3188949	0.3685617	-0.1747453	0.9902176
Finance	10	10	0.2996820	0.4260455	-0.3299259	0.9852135
Healthcare	10	10	0.8725559	1.1009737	0.0113213	3.7284477
Ratail	10	10	0.4125792	0.7871251	-0.9419625	1.8247774
Technology	10	10	0.3410446	0.6297978	-0.4925115	1.6275885

Key observations: The healthcare industry has the highest return ratio, almost twice as much as the second-highest industry, which contradicts our previous prediction based on the PB ratio. The healthcare industry has been significantly impacted by the pandemic, which could have contributed to its impressive performance. This reinforces our assumption that the COVID-19 pandemic has had a significant impact on the stock market.

Conclusion and Findings

In general, the stock price of value stocks should be the one with the most potential for growth. However, this trend has reversed in recent years. Over the past five years, growth stocks have outperformed value stocks, possibly due to the pandemic.

It is worth noting that the retail sector has a larger proportion of growth stocks, while the finance sector has more value stocks. Nonetheless, the healthcare industry has demonstrated the highest return ratio, providing evidence that COVID-19 has had a significant impact on the stock market.

Appendix - Company List

Companies	Sector
Apple Inc. - AAPL (1976)	Tech
Microsoft Corporation - MSFT (1975)	Tech
IBM - IBM (1911)	Tech
Intel Corporation - INTC (1968)	Tech
Cisco Systems, Inc. - CSCO (1984)	Tech
Oracle Corporation - ORCL (1977)	Tech
Hewlett Packard Enterprise - HPE (1939)	Tech
Dell Technologies Inc. - DELL (1984)	Tech
Adobe Inc. - ADBE (1982)	Tech
Texas Instruments Incorporated - TXN (1930)	Tech
Exxon Mobil Corporation - XOM (formerly Standard Oil Company) (1870)	Energy
Chevron Corporation - CVX (formerly Standard Oil Company of California) (1879)	Energy
ConocoPhillips - COP (2002)	Energy
Halliburton Company - HAL (1919)	Energy
Williams Company - WMB (1908)	Energy
Occidental Petroleum Corporation - OXY (1920)	Energy
Marathon Petroleum Corporation - MPC (1887)	Energy
Valero Energy Corporation - VLO (1980)	Energy
EOG Resources, Inc. - EOG (1999)	Energy
Schlumberger Limited - SLB (1926)	Energy
Walmart Inc. - WMT (1962)	Retail
Amazon.com, Inc. - AMZN (1994)	Retail
The Home Depot, Inc. - HD (1978)	Retail
Costco Wholesale Corporation - COST (1983)	Retail
Target Corporation - TGT (1902)	Retail
Lowe's Companies, Inc. - LOW (1946)	Retail
CVS Health Corporation - CVS (1963)	Retail
Walgreens Boots Alliance, Inc. - WBA (1901)	Retail
Dollar Tree, Inc. - DLTR (1986)	Retail
TJX Companies Inc. - TJX (1956)	Retail
Johnson & Johnson - JNJ (1886)	Healthcare
Pfizer Inc. - PFE (1849)	Healthcare
Merck & Co., Inc. - MRK (1891)	Healthcare
Bristol-Myers Squibb Company - BMY (1858)	Healthcare
Eli Lilly and Company - LLY (1876)	Healthcare
Abbott Laboratories - ABT (1888)	Healthcare
Amgen Inc. - AMGN (1980)	Healthcare
UnitedHealth Group Incorporated - UNH (1977)	Healthcare
Gilead Sciences, Inc. - GILD (1987)	Healthcare
Biogen Inc. - BIIB (1978)	Healthcare
JPMorgan Chase & Co. - JPM (2000)	Finance
Bank of America Corporation - BAC (1904)	Finance
Wells Fargo & Company - WFC (1852)	Finance
Citigroup Inc. - C (1812)	Finance
Goldman Sachs Group, Inc. - GS (1869)	Finance
American Express Company - AXP (1850)	Finance
Morgan Stanley - MS (1935)	Finance
Visa Inc. - V (1958)	Finance
Mastercard Incorporated - MA (1966)	Finance
Capital One Financial Corporation - COF (1988)	Finance