

Discriminant Analysis of Financial Sectors

Zhitong Wang

Hangzhou Runzhou Investment Management Limited Corporation, Hangzhou, 310051, China

Abstract. The Machine learning has been wide used and trained to do jobs to analyze companies' performance based on the financial data, and the portfolio managers have widely accepted the use the machine learning as a tool to assist company analysis and investment decision making. As a result of the test performed above, the discriminant test, investors can train the model to distinguish performance of the companies, whether it will perform at the top 25% of the sector or bottom 25% of the sector, from the financial/banking sector based on the financial ratios of net income/revenue and total debt/total liabilities or EBIT/revenue and total debt/total liabilities.

Keywords: Machine Learning; Financial Data; Performance Analysis.

1. Introduction

In the financial world, when people are managing their own portfolio of investments, whether the companies in the investment portfolio or the companies that investors are about to invest in will make money in the future concerns investors the most. When investors invest money into the companies' stocks, what they are expecting are the value increase of the company which leads to higher stock prices and hence the increase of value of their portfolios. Therefore, a plenty of valuation models is developed upon the need and the progress of the industry. From traditional approaches such as P/E ratio analysis, Discounted Cash Flow Estimations, and Comparable Company Analysis to Modern Approaches and to Machine Learnings, the methods of valuations are evolving and progressing and the investors are using these methods to maximize their returns and do their best to minimize the risk.

Even though the valuation models may vary, the data which these valuation models incorporate are from the same area, which is financial statements. Financial Statements of the companies gives investors an overview of the company financially, and the data inside the financial statements reflect the financial health of the companies, and give insightful summary of the revenue, cash flow and the performance of the company (Investopedia, 2022). To be specific, the balance sheet of the company describes what the company's assets are, the liabilities, or what they owe, are and the stockholders' ownership, which is equity, is during the time when the financial statement is released; also the income statements gives the reader a throughout report on how much does the company earn, what are the gross profit and the net profit as well as the costs and expenses of the company over the past year; the cash flow statement reflects how the cash flows in and out of the company, whether the cash is used in operating activities or investing activities, and the generated cash can also being used to expand the company or pay the dividends to shareholders. Therefore, incorporating financial statements into the model helps investors to analyze companies and make good judgement of whether the companies will give investors satisfactory outcome in the future.

In the passage, I will one of the analysis models called discriminant analysis to analysis the performance of the companies I choose from the Financial/Banking sector, using different financial data and their ratios.

2. Procedures

First of all, 10 companies are selected from the Financial/Banking sector. The companies chosen and their simple description are listed as follows:

Berkshire Hathaway Inc. – It is one of the largest companies in the US whose CEO is the well-known Warren Buffett. Berkshire Hathaway Inc. expands its business and investment to insurance, utility and energy, manufacturing, service and retailing as well as financial and financial services. It now owns equities from companies all over the world (Wikipedia, 2022).

JP Morgan Chase & Co. – It is a multinational investment bank and financial services holding company which based in the United States of America. It is the largest bank in the United States and the largest bank by market capitalization in the world. Its business includes investment banking, asset management, private banking, private wealth management, commercial banking, and treasury services (Wikipedia, 2022).

Visa Inc. – It is a multinational financial service corporation based in California, the main business of the company is to facilitate electronic funds transfers around the world through visa branded visa debit cards, credit cards and prepaid cards issued by different banks. Now it is the world's second-largest card payment company in the world, one place behind China Union Pay (Wikipedia, 2022).

Bank of America Corporation – It is an American multinational investment bank and financial service holding company. It is the second largest banking company in the United States. Its primary business includes commercial banking, wealth management, and investment banking (Wikipedia, 2022).

Wells Fargo & Company – It is a multinational financial service company based in the United States. Its business includes wealth and investment management and consumer banking and lending (Wikipedia, 2022).

Bank of China – It is the Chinese government owned commercial bank, and it is the fourth largest bank in the world. It provides commercial banking, investment banking and personal banking business to its customers (Wikipedia, 2022).

American Express – It is an American multinational corporation in payment card services. It provides transfer business, investment banking and card services to the customers (Wikipedia, 2022).

Goldman Sachs – It is a multinational financial service and investment banking company based in the United States. It offers various business such as investment banking, asset management (Wikipedia, 2022).

Morgan Stanley – It is a multinational financial service and investment management company based in the United States. Its primary business are institutional securities, wealth management and investment management (Wikipedia, 2022).

UBS Group AG – It is a multinational financial service and investment bank based in Switzerland. It is the largest private bank in the world and its business include wealth management, asset management, and investment banking services for private, corporate, and institutional clients.

After knowing the target companies, the historical data of the companies' stock prices performances are downloaded from Yahoo Finance.com. The historical data includes close prices of each companies' stocks, on a daily bases, from November 10, 2015 to November 10, 2021. All the data was included in Excel Spread sheets, and the total return from the selected period of time is calculated using the formula $\text{Return} = (\text{End Date Price} / \text{Start Date Price}) - 1$. After calculating the stock returns over the 5 years, the 10 companies are divided in to two groups by their performances. The first group includes top 25% performance based on their returns (named "Top 25%") and the second group includes bottom 25% performance (named "Bottom 25%"). Hence two groups of 3 companies are generated.

Furthermore, the financial data for the companies are chosen to be tested in the discriminant analysis. These data includes companies' Revenue, which represents the total gain for each company during the past year and reflects the company's business outcome during the past year; company's EBIT is selected, which is the earnings before interests and tax, which is a representation of how company truly gain after the costs, expenses, depreciations and amortizations; company's net income is also selected since it reflects how much income made during the year is belonged to the equity holders; total debt is selected, which is the debt that company owes for the past year; total asset is selected to show the company's asset; and lastly total liabilities is included to reflects company's obligations and debts.

These financial data can be transferred into different ratios, the analysis includes net income margins (net income/revenue), total debt/total liabilities, EBIT margins (EBIT/Revenue), total debt/total asset.

After calculating the ratios above for each company, the outcome of the Fisher's Discriminant Function $Z = aX_1 + bX_2$ is calculated by putting each pair of ratios into X_1 and X_2 and calculate the Mahalanobis Distance based on the mean difference and the pooled variance between two groups of the companies using excel solver to get the optimizations of the a and b in the function, therefore we get the result of discriminant functions for each company based on each pair of the ratios. The cutoff point, the average between the means of the discriminant result of the two groups, is calculated. Finally, the discriminant results for each company are compared to the cutoff point, if the results are lower than the cutoff point, the company is marked as "top" as it is the top 25% of the performance, and if the results are higher than the cutoff point, then the company is marked as "bottom" as it is the top 25% of the performance.

After running several times with different pairs of ratios, the best two ratios are determined by the discriminant test if the "top" marks and the "bottom" marks are best corresponding with the actual performance of the company.

3. Results

First of all, returns of the companies are surprisingly varied: Berkshire Hathaway returns 114.2% of growth from 2015 to 2021, the JP Morgan returns 91.7% of growth, Visa returns 144.0% of growth, which is the highest of all, Bank of America returns 104.6% of growth, Wells Fargo returns -17.8%, Bank of China returns -26.2%, which is the last place of all 10 companies' returns, American Express returns 98.8%, Goldman Sachs returns 0.4%, Morgan Stanley returns -6.7%, and UBS returns -15.4%. Of all the returns, we are able to group Berkshire Hathaway (+114.2%), Visa (+144.0%), and Bank of America (+104.6%) to the "top 25%" group (highlighted as yellow), and Wells Fargo (-17.8%), Bank of China (-26.2%), and UBS (-15.4%) to the "bottom 25%" group (highlighted as blue). (Figure. 1)

	Return
BRK-B	114.2%
JPM	91.7%
V	144.0%
BAC	104.6%
WFC	-17.8%
BACHY	-26.2%
AXP	98.8%
GS-PJ	0.4%
MS-PF	-6.7%
UBS	-15.4%

Figure 1. The returns of the Companies selected, 11/10/2015 to 11/10/2021

The Net Income/Revenue and the Total Debt/Total Liabilities produce "a" in the discriminant function is equal to 0.6874557005, the "b" in the discriminant function is equal to 0.726225324, the Mahalanobis Distance is equal to 5.126536472, and the cutoff point is equal to 0.316126257. Finally the outcome returns two "top" representing Berkshire Hathaway and Visa are correctly in the category of "top 25%", and Wells Fargo, Bank of China and UBS are correctly in the category of "bottom 25%", with only one error of Bank of America being classed as "bottom 25%". (Figure. 2)

DICRIMINANT ANALYSIS							Mahalanobis Distance
Mahalanobis Distance							5.126536472
				Mean Gp1 - Mean Gp2	0.07327		
				Pooled Variance	0.001047	cutoff point	0.316126257
	1.0000004	a	b				
		0.687457005	0.726225324				
				Mean =	0.352761		
		Net Income/Revenue	Total Debt/Total Liabilities	Variance =	0.001405		
BRK-B	1	0.253203284	0.257431498	Top 25%	0.36102	top	
V	2	0.510723916	0.047261036	Top 25%	0.385423	top	
BAC	3	0.342901709	0.104803394	Top 25%	0.311841	bottom	
WFC	1	0.258064516	0.098925572	Bottom 25%	0.249251	bottom	
BACHY	2	0.358493315	0.063640745	Bottom 25%	0.292666	bottom	
UBS	3	0.212698594	0.207010169	Bottom 25%	0.296557	bottom	
				Mean =	0.279491		
				Variance =	0.00069		

Figure 2. Net/Income versus Total Debt/Total Liabilities

The Net Income/Revenue versus Total Debt/Total Asset comes with the outcome of “a” equal to 0.911920525, “b” equal to 0.410366882, Mahalanobis Distance equal to 0.815612183, and cutoff point equal to 0.350112846. The outcome misplaced Berkshire Hathaway into the “bottom 25%” group and Bank of China into the “top 25%” group. (Figure. 3)

DICRIMINANT ANALYSIS							Mahalanobis Distance
Mahalanobis Distance							
				Mean Gp1 - Mean Gp2	0.100729		0.815612183
				Pooled Variance	0.01244	cutoff point	0.350112846
	1	0.911920525	0.410366882				
		a	b				
				Mean =	0.400477		
		Net Income/Revenue	Total Debt/Total Asset	Variance =	0.022752		
BRK-B	1	0.253203284	0.11917387	Top 25%	0.279806	bottom	
V	2	0.510723916	0.253052017	Top 25%	0.569584	top	
BAC	3	0.342901709	0.095873317	Top 25%	0.352042	top	
WFC	1	0.258064516	0.089271524	Bottom 25%	0.271968	bottom	
BACHY	2	0.358493315	0.06356041	Bottom 25%	0.353	top	
UBS	3	0.212698594	0.195706698	Bottom 25%	0.274276	bottom	
				Mean =	0.299748		
				Variance =	0.002128		

Figure 3. Net Income/Revenue versus Total Debt/Total Asset

The Net Income/Revenue versus EBIT/Revenue produce “a” equal to 1.0000000 and “b” equal to 0, the Mahalanobis Distance equal to 0.755762872 and the cutoff point equal to 0.322680902. The outcome misplaced Berkshire Hathaway into the “bottom 25%” group and Wells Fargo and UBS into the “top 25%” group. (Figure. 4)

DICRIMINANT ANALYSIS							Mahalanobis Distance
Mahalanobis Distance							
				Mean Gp1 - Mean Gp2	0.092524		0.755762872
				Pooled Variance	0.011327	cutoff point	0.322680902
	1.0000001	1.00000004	0				
				Mean =	0.368943		
		Net Income/Revenue	EBIT/Revenue	Variance =	0.017088		
BRK-B	1	0.253203284	0.326695541	Top 25%	0.253203	bottom	
V	2	0.510723916	0.687658162	Top 25%	0.510724	top	
BAC	3	0.342901709	0.381268726	Top 25%	0.342902	top	
WFC	1	0.258064516	0.367120216	Bottom 25%	0.258065	bottom	
BACHY	2	0.358493315	0.457918723	Bottom 25%	0.358493	top	
UBS	3	0.212698594	0.270515417	Bottom 25%	0.212699	bottom	
				Mean =	0.276419		
				Variance =	0.005567		

Figure 4. Net Income/Revenue versus EBIT/Revenue

The Total Debt/Total Asset versus EBIT/Revenue produce “a” equaled to 0.31172022 and “b” equaled to 0.950174432, the Mahalanobis Distance equaled to 0.471406846 and the cutoff point equaled to 0.436935836. The outcome misplaced Berkshire Hathaway into the “bottom 25%” group and misplaced Bank of China into the “top 25%” group. (Figure. 5)

DICRIMINANT ANALYSIS									
Mahalanobis Distance									Mahalanobis Distance
					Mean Gp1 - Mean Gp2	0.107462			0.471406846
					Pooled Variance	0.024497	cutoff point	0.436935836	
	1.0000009	a	b						
		0.31172022	0.950174432						
					Mean =	0.490667			
					Variance =	0.044279			
		Total Debt/Total Asset	EBIT/Revenue						
BRK-B	1	0.11917387	0.326695541		Top 25%	0.347567	bottom		
V	2	0.253052017	0.687658162		Top 25%	0.732277	top		
BAC	3	0.095873317	0.381268726		Top 25%	0.392157	bottom		
WFC	1	0.089271524	0.367120216		Bottom 25%	0.376656	bottom		
BACHY	2	0.06356041	0.457918723		Bottom 25%	0.454916	top		
UBS	3	0.195706698	0.270515417		Bottom 25%	0.318043	bottom		
					Mean =	0.383205			
					Variance =	0.004716			

Figure 5. Total Debt/Total Asset versus EBIT/Revenue

The Total Debt/Total Asset versus Total Debt/Total Liabilities produce “a” equaled to 0.969570893 and “b” equaled to 0.244810708, the Mahalanobis Distance equaled to 0.286831491 and the cutoff point equaled to 0.163752258. The outcome misplaced Bank of America into the “bottom 25%” group and misplaced UBS into the “top 25%” group. (Figure. 6)

DICRIMINANT ANALYSIS									
Mahalanobis Distance									Mahalanobis Distance
					Mean Gp1 - Mean Gp2	0.041898			0.286831491
					Pooled Variance	0.00612	cutoff point	0.163752258	
	1	a	b						
		0.969570893	0.244810708						
					Mean =	0.184701			
					Variance =	0.004811			
		Total Debt/Total Asset	Total Debt/Total Liabilities						
BRK-B	1	0.11917387	0.257431498		Top 25%	0.17857	top		
V	2	0.253052017	0.047261036		Top 25%	0.256922	top		
BAC	3	0.095873317	0.104803394		Top 25%	0.118613	bottom		
WFC	1	0.089271524	0.098925572		Bottom 25%	0.110773	bottom		
BACHY	2	0.06356041	0.063640745		Bottom 25%	0.077206	bottom		
UBS	3	0.195706698	0.207010169		Bottom 25%	0.24043	top		
					Mean =	0.142803			
					Variance =	0.00743			

Figure 6. Total Debt/Total Asset versus Total Debt/Total Liabilities

DICRIMINANT ANALYSIS									
Mahalanobis Distance									Mahalanobis Distance
					Mean Gp1 - Mean Gp2	0.067558			1.980317397
					Pooled Variance	0.002305	cutoff point	0.34195876	
	0.9999994	a	b						
		0.565731873	0.824588884						
					Mean =	0.375738			
					Variance =	0.004304			
		EBIT/Revenue	Total Debt/Total Liabilities						
BRK-B	1	0.326695541	0.257431498		Top 25%	0.397097	top		
V	2	0.687658162	0.047261036		Top 25%	0.428001	top		
BAC	3	0.381268726	0.104803394		Top 25%	0.302116	bottom		
WFC	1	0.367120216	0.098925572		Bottom 25%	0.289265	bottom		
BACHY	2	0.457918723	0.063640745		Bottom 25%	0.311537	bottom		
UBS	3	0.270515417	0.207010169		Bottom 25%	0.323737	bottom		
					Mean =	0.30818			
					Variance =	0.000306			

Figure 7. EBIT/Revenue versus Total Debt/Total Liabilities

At last, the EBIT/Revenue versus Total Debt/Total Liabilities produce “a” equaled to 0.565731873 and “b” equaled to 0.824588884, the Mahalanobis Distance equaled to 1.980317397 and the cutoff point equaled to 0.34195876. The outcome, as same as the first one, the outcome returns two “top”,

representing Berkshire Hathaway and Visa are correctly in the category of “top 25%”, and Wells Fargo, Bank of China and UBS are correctly in the category of “bottom 25%”, with only one error of Bank of America being classed as “bottom 25%”. (Figure. 7)

4. Analysis

The financial data that are choice are revenue, EBIT, net income, total debt, total asset, and total liabilities. The logic behind choosing these financial factors is crucial.

First of all, revenue reflects the total money gained from the operation of business during the time of that year. The importance of the revenue is obvious, it is the top line data in the financial statement, therefore all the investors focus on this financial data as their first step of the analysis (Investopedia, 2022).

Secondly, EBIT is also crucial when analyzing company's performance. EBIT stands for earnings before interest and tax. It is calculated as revenue minus cost of goods sold, selling, general and administrative expenses, research and development expenses, depreciation and amortization. It is used to analyze companies' performance excluding the affects brought by capital and tax (Investopedia, 2022). EBIT and net income are both selected since financial institutions may generate large amount of money from the interests, therefore when selecting data, EBIT and net income should both be considered.

Net Income is selected since it is the income that a company earns after excluding all the costs and expenses, as well as excluding the effect of the interests and tax. Net income belongs to all the shareholders, it means what the investors of the company will get after a year of company's operation, if the company decide to distribute all the net incomes. However, nearly all the companies do not distribute all the net incomes since there will be future plans for the companies, and to develop companies, money requires to be re-invested or re-used, and these money comes from the net incomes.

Total debt is the sum of the long-term and short-term debts of the companies. It reflects what companies are performing on the money they owe. And the total asset shows how much the company owns on their asset end.

The total liabilities are crucial for financial institutions since it not only include the debts of the companies, but for financial companies, it also includes the money that companies owes to their clients. The clients put their money to the financial companies for investments, savings, or wealth management, then this amount of money becomes liabilities to the companies since they need to pay back this money when the clients ask to or the by the end of wealth management. Therefore, apart from total debt, the total liabilities also are taken into account.

The output of the discriminant test shows that net incomes/revenue versus total debt/total liabilities and EBIT/revenue versus total debt/total liabilities both comes with the same outcome, that is correctly discriminates most of the companies, but only left with Bank of America falsely categorized as “bottom” 25%. Therefore, net incomes/revenue versus total debt/total liabilities and EBIT/revenue versus total debt/total liabilities both can be the best pairs that have the most discriminating power of all other ratios.

This analysis is a process of machine learning, when using the ratios of the known companies' performance, and determine the discriminant function based on the financial ratios of the known companies, in order to train the model to set a discriminant function. This discriminant test shows the boundary of the two groups of the companies, and the cutoff point is the boundary. The machine training set discriminant output of “a” and “b” for future analysis, the set numbers of the discriminant function represent the function of the companies' performance based on the financial ratios, and the model has been trained to estimate and categorize the company performance into different groups. Therefore, when doing future analysis, suppose a new company, investors are able to calculate possible performance of the target company by pulling out the financial ratio of either net income/revenue versus total debt/total liabilities or EBIT/revenue versus total debt/total liabilities, and then put these variables into the known function, and then investors are able to get the output of

the discriminant test and see whether the output is on the left side (less than cutoff point) or the right side (greater than cutoff point), and gives a good estimation of the target company's future performance, whether it is likely to perform at the top 25% or the bottom 25%, and base on the result we are able to make sound investment decisions.

5. Conclusion

Future analysis based on the test can be done to get the possible performance of the company. However, this test can be improved, since the training data from the test consists only 6 companies from the 10 total companies stock performance, 3 from the top 25% and 3 from the bottom 25%, so the sample number is still too small. The procedure can be improved by selecting more numbers of the companies and training more financial ratios possible to get a much more accurate result. Furthermore, the dividend, buybacks of the stock by the company, and stock split effect on the stock price is not considered when doing this analysis, so more accurate analysis should be able to consider all the effects and then calculate the performance of the stock for the companies.

References

- [1] Hayes, A. (2022, September 8). What is revenue? definition, formula, calculation, and example. Investopedia. Retrieved November 11, 2022, from <https://www.investopedia.com/terms/r/revenue.asp>.
- [2] Liberto, D. (2022, November 7). Total liabilities: Definition, types, and how to calculate. Investopedia. Retrieved November 11, 2022, from <https://www.investopedia.com/terms/t/total-liabilities.asp>.
- [3] Maverick, J. B. (2022, May 26). Why do shareholders need financial statements? Investopedia. Retrieved November 11, 2022, from <https://www.investopedia.com/ask/answers/032615/why-do-shareholders-need-financial-statements.asp#:~:text=Key%20Takeaways%201%20Financial%20statements%20provide%20a%20snapshot,a%20company%27s%20revenue%2C%20expenses%2C%20profitability%2C%20and%20debt.%20%E6%9B%B4%E5%A4%9A%E9%A1%B9%E7%9B%AE>.
- [4] Murphy, C. B. (2022, September 14). Earnings before interest and taxes (EBIT): How to calculate with example. Investopedia. Retrieved November 11, 2022, from <https://www.investopedia.com/terms/e/ebit.asp>.
- [5] Wikimedia Foundation. (2022, November 10). Berkshire Hathaway. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Berkshire_Hathaway.
- [6] Wikimedia Foundation. (2022, November 2). Bank of America. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Bank_of_America.
- [7] Wikimedia Foundation. (2022, November 2). Wells Fargo. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Wells_Fargo.
- [8] Wikimedia Foundation. (2022, November 6). JPMorgan Chase. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/JPMorgan_Chase.
- [9] Wikimedia Foundation. (2022, November 7). UBS. Wikipedia. Retrieved November 11, 2022, from <https://en.wikipedia.org/wiki/UBS>.
- [10] Wikimedia Foundation. (2022, November 8). Goldman Sachs. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Goldman_Sachs.
- [11] Wikimedia Foundation. (2022, October 25). Morgan Stanley. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Morgan_Stanley.
- [12] Wikimedia Foundation. (2022, October 29). Visa Inc.. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Visa_Inc.
- [13] Wikimedia Foundation. (2022, October 8). Bank of China. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/Bank_of_China.
- [14] Wikimedia Foundation. (2022, October 9). American Express. Wikipedia. Retrieved November 11, 2022, from https://en.wikipedia.org/wiki/American_Express.