Proceedings of the 2nd International Conference on Management Research and Economic Development
DOI: 10.54254/2754-1169/72/20240681

Predicting Stock Returns with Industry Concentration

Haolin Luo[1,a,*]

[1]*Beijing Foreign Studies University, Zizhuyuan Street, Beijing, China*
a. 1316234487@qq.com
**corresponding author*

Abstract: In order to predict stock returns more effectively, it has been proposed to use industry concentration as a benchmark. We want to verify whether this variable is significant today. In this paper, we first screen out some firms that may affect the results of the experiment, such as financial firms as well as some other small and micro firms. Then, we use the total assets and total sales of the firms separately to calculate their industry concentration. They are then divided into 10 portfolios from low to high based on their industry concentration. Their returns are then compared to their industry concentration to determine whether it is significant or not. The results show that industry concentration is not significant in predicting stock returns today. Possible reasons for this could be that the factor of industry concentration has been overused, or that the methodology of our calculations was not rigorous enough. In addition, the impact of the new crown epidemic cannot be ignored.

Keywords: stock return predict, industry concentration, portfolio create

1. Introduction

In today's society, stock return has always been a topic of great interest and the predict for stock return is an important and complex issue. In the field of finance, people use many variables to predict returns, and there are various approaches to solve this problem. The most commonly used method is portfolio sorting strategy.

The purpose of this paper is to use data from individual companies to process and use industry concentration to predict stock return. Industry concentration is the sum of the market shares (output value, production, sales, sales volume, number of employees, total assets, etc.) of the top N largest firms in an industry. It is a measure of the degree of concentration of the market structure of the whole industry, and is used to measure the difference in the number and relative size of firms, which is an important quantitative indicator of market power.

Previously, Hou and Robinson have used the industry concentration variable to study stock return, where they use the NYSE benchmark [1]. At the same time, I want to test whether this variable also has a significant effect on today's stocks and whether it can be used today. I hypothesize that industry concentration still has a significant effect on stock return.

In order to test this hypothesis, I follow the procedure of Hou et al. to see if I could also obtain significant results including the latest data in 2023 [1]. I collected the data of companies in the last 20 years and screened them, keeping only companies in NYSE, Nasdaq, Amex and removing companies in the financial sector. The results show that the stock return of the 10 portfolios classified by industry concentration is not significantly correlated with the industry concentration. The results

DOI: 10.54254/2754-1169/72/20240681

show that there is no significant correlation between stock return and industry concentration for the 10 portfolios classified by industry concentration.

The reason for this result may be that this factor is already known and no longer significant. Another possible explanation is that industry concentration is also not significant because of the epidemic that has set back the stock market in general in recent years.

The rest of the paper is structured as follows. The second part describes the collection and processing of the data, the third part shows the results obtained, the fourth part is a table, and the fifth part is an analysis of the reasons for such results.

This paper introduce the relationship between industry concentration and portfolio return prediction. Hou and Robinson studied the effect of industry concentration on average stock returns [1]. They constructed a new indicator of industry concentration based on data on U.S. public companies and found, after comparing it with the traditional Herfindahl index, that industry concentration was significantly and positively related to average stock returns, controlling for other risk factors, and under various groupings [1]. Previous studies of stock returns include Asness, Clifford, and Ross who used the Fama-French three-factor model to investigate the presence of industry factors in stock returns. After controlling for factors such as market, size, and book-to-market ratio, they find that stock returns are also affected by factors within the industry, such as relative value [2]. Moskowitz, Tobias and Mark propose a method for evaluating portfolio performance based on sector momentum strategies and apply it to the U.S. stock market. They find that momentum portfolios ranked according to past stock returns tend to achieve significant positive returns. This is due to co-movement between stocks and the industries to which they belong, which is present for a variety of characteristic groupings. They also find that momentum effects are related to economic cycles, interest rate changes, and stock volatility [3]. Cohen, Randolph, Christopher and Tuomo propose a new value metric, the value variance, which they find that it outperforms traditional value indicators, better predicts future stock market returns, and explains the correlation between the value factor (HML) and the size factor (SMB) in cross-sectional stock returns [4]. Hou investigates the role of information diffusion in explaining the lead-lag effect in stock returns. By using data on U.S. listed companies, he finds that information diffusion is one of the important reasons for the existence and variation of the lead-lag effect in stock returns [5].

In addition, Almazan, Andres and Carlos studied the determinants of capital structure dispersion in industries [6]. Mackay, Peter and Gordon examine the importance of industry on firms' financial and practical decisions, and they find that, in addition to standard industry fixed effects, financial structure depends on a firm's position in the industry. Financial structure, technology, and risk are jointly determined within the industry. These findings are consistent with recent industry equilibrium models of financial structure [7]. Hou and Robinson study the effect of industry concentration on average stock returns. They construct a novel industry concentration index based on data from U.S. public companies and find that industry concentration is significantly positively related to average stock returns after controlling for other risk factors and under various groupings [1]. Hou, Chen and Lu study the effect of industry concentration on replication studies of anomalies in finance and accounting. They collect 447 anomaly variables and examined the existing literature, finding that most of the anomalies do not hold under today's standards, and even for the successfully replicated anomalies, their economic effects are much smaller than initially reported. In short, capital markets are more efficient than previously recognized [5].

So I verify the usefulness of industry concentration by combining the methods they introduced. The result is not significant.

DOI: 10.54254/2754-1169/72/20240681

2. Data Processing

I perform data processing and analysis in this section. The monthly returns are from CRSP and the accounting information is from the annual and quarterly base files used to calculate the statistics. The sample is collected from January 2001 to January 2023. Financial companies are excluded.

Section 2.1 describes our data processing procedures. Section 2.2 provides a detailed analysis of the resulting results.

The formula for calculating industry concentration is as follows:

$$\frac{x}{y} \times 100\%$$

Where x represents the total sales or assets of the company and y represents the total SALES or ASSET of the industry of company.

2.1. Data Processing

In order to merge the company information in compustat with the monthly returns from CRSP, I need link file process the data obtained from compustat and CRSP, then merge them with link file. First, I download 20 years of GVKEY, Stock Exchange Code, Fiscal Year, Total Assets, Equity and Sales data for each company from the WRDS database. To facilitate the calculation and merge, I make sure that the same gvkey and LPERMNO correspond to only one company, remove the duplicate and extra data, and then screen out the companies using the same gvkey in the same year according to the fiscal year. Finally, the compustat data is expanded by 12 copies representing each month, and then the first month of the fiscal year is determined according to the fiscal year, so that the company information in the compustat can be merged with the monthly returns of CRSP. Make sure predictions are made only on ex-ante information. According to the total sales and total assets of each industry, two different industry concentrations are derived.

Next is the screening of different companies, in order to ensure the accuracy of the results, I only retain the NYSE, Nasdaq, Amex companies ordinary companies, and, delete the financial companies. Then I generated a log function that divided the price of each month by the price of the previous month. After that I made calculations based on every three months, every six months, every nine months, and every twelve months. Finally, I sort the log functions of each month and create 10 portfolios according to NYSE breakpoints, and with the total market equity of each portfolio for each month form the value weighted portfolios.

2.2. Analyzing the Results

From the table, I can see that whether it is January, March, June or December data, their results are not significant, which means that industry concentration and stock return are not correlated.

DOI: 10.54254/2754-1169/72/20240681

Table 1: Summary Statistics.

Panel A: Summary of Industry Concentration Measures								
	Mean	Median	STD	Max	Min	25%	75%	90%
Sales	3204.219	328.73	15100.05	569962	-1964.999	54.385	1564.944	5748.331
Assets	4257.758	405.171	20313.83	797769	0	88.31	1900	7282

Panel B: Summary of different portfolios						
portfolios	obs(sales)	mean(sales)	STD	obs(assets)	mean(assets)	STD
1	350,203	131.664	219.347	350,969	222.787	360.983
2	102,045	745.302	922.207	101,660	1113.191	1492.846
3	70,778	1390.016	1916.799	71,243	2020.893	3087.228
4	57,559	2106.171	3021.325	61,981	2929.985	4886.917
5	50,668	2898.535	4548.362	51,449	4657.701	8468.711
6	43,721	4434.907	6884.025	44,772	8152.99	15580.97
7	41,037	6721.275	14330.74	43,348	8582.689	20534.48
8	38,849	9528.493	19228.32	38,940	13135.95	30242.1
9	36,428	17413.5	37272.45	35,832	20111.03	47832.15
10	39,996	17527.8	47499.78	40,119	21238.77	63461.55

The sample includes data on companies listed on the New York Stock Exchange /AMEX/NASDA between January 2001 and January 2023, derived from the WRDS database. The securities contain CRSP monthly returns and COMPUSTAT annual data, and Panel A reports the maximum, minimum, standard deviation, mean, median and 25,75, 90th percentile of Sales and Assets for all companies. Panel B reports the average and total number of assets and sales after being divided into 10 portfolios.

Table 2: Monthly weighted average returns using total sales as the criterion for calculating industry concentration.

Panel A: Monthly value-weighted portfolio average return										
	1	2	3	4	5	6	7	8	9	10
ret	-0.007	-0.002	-0.001	-0.0007	0.0006	-0.0005	-0.0001	-0.00008	0.0003	-0.001
Panel B: Value weighted portfolio average return in three months										
	1	2	3	4	5	6	7	8	9	10
ret	0.022	0.019	0.020	0.017	0.023	0.014	0.013	0.014	0.013	0.014
Panel C: Value weighted portfolio average return in six months										
	1	2	3	4	5	6	7	8	9	10
ret	0.042	0.037	0.039	0.032	0.043	0.026	0.029	0.027	0.023	0.029
Panel D: Value weighted portfolio average return in nine months										
	1	2	3	4	5	6	7	8	9	10
ret	0.068	0.056	0.065	0.049	0.065	0.041	0.051	0.044	0.036	0.046
Panel E: Value weighted portfolio average return in twelve months										
	1	2	3	4	5	6	7	8	9	10
ret	0.015	0.014	0.014	0.012	0.015	0.009	0.011	0.011	0.010	0.013

I calculate industry concentration based on the total share of firms sold in the industry and divide it into ten portfolios by share.The firm data is derived from the WRDS database from January 1, 2001 to January 1, 2023. I use the WRDS database as the basis for calculating industry concentration. Then

DOI: 10.54254/2754-1169/72/20240681

their weighted average returns are calculated by different months. Panel A is calculated in one month, panel B is calculated in three months, panel C is calculated in six months, panel D is calculated in nine months, and panel E is calculated in twelve months.

Table 3: Monthly Weighted Average Returns Using Total Assets as the Calculation Criteria for Industry Concentration.

Panel A: Monthly value-weighted portfolio average return										
	1	2	3	4	5	6	7	8	9	10
ret	-0.006	-0.002	-0.001	-0.0009	-0.0001	-0.001	-0.002	0.003	-0.0002	-0.0007
Panel B: Value weighted portfolio average return in three months										
	1	2	3	4	5	6	7	8	9	10
ret	0.025	0.018	0.020	0.018	0.018	0.013	0.012	0.019	0.011	0.015
Panel C: Value weighted portfolio average return in six months										
	1	2	3	4	5	6	7	8	9	10
ret	0.046	0.036	0.041	0.037	0.033	0.026	0.021	0.035	0.022	0.030
Panel D: Value weighted portfolio average return in nine months										
	1	2	3	4	5	6	7	8	9	10
ret	0.075	0.057	0.066	0.056	0.051	0.041	0.036	0.055	0.038	0.048
Panel E: Value weighted portfolio average return in twelve months										
	1	2	3	4	5	6	7	8	9	10
ret	0.017	0.013	0.013	0.013	0.011	0.011	0.009	0.014	0.009	0.013

I calculate the industry concentration based on the total share of firms' assets in the industry and divide it into ten portfolios according to their share. firm data is obtained from the WRDS database from January 1, 2001 to January 1, 2023. Then their weighted average returns are calculated by different months. Panel A is calculated in one month, panel B is calculated in three months, panel C is calculated in six months, panel D is calculated in nine months, and panel E is calculated in twelve months.

DOI: 10.54254/2754-1169/72/20240681

Table 4: Monthly Average Returns Using Total Sales as the Criterion for Calculating industry concentration.

Panel A: Monthly equal-weighted portfolio average return										
	1	2	3	4	5	6	7	8	9	10
ret	-0.007	-0.002	-0.0007	-0.0004	0.0009	0.00005	0.002	0.0009	0.002	0.0001
Panel B: Equal-weighted portfolio average return in three months										
	1	2	3	4	5	6	7	8	9	10
ret	0.110	0.055	0.058	0.043	0.061	0.042	0.044	0.056	0.033	0.035
Panel C: Equal-weighted portfolio average return in six months										
	1	2	3	4	5	6	7	8	9	10
ret	0.192	0.106	0.102	0.081	0.096	0.078	0.088	0.099	0.062	0.069
Panel D: Equal-weighted portfolio average return in nine months										
	1	2	3	4	5	6	7	8	9	10
ret	0.270	0.159	0.153	0.125	0.134	0.113	0.129	0.151	0.093	0.104
Panel E: Equal-weighted portfolio average return in twelve months										
	1	2	3	4	5	6	7	8	9	10
ret	0.075	0.037	0.047	0.028	0.044	0.029	0.030	0.029	0.023	0.025

I calculate industry concentration based on the total share of firms sold in the industry and divide it into ten portfolios by share.The firm data is derived from the WRDS database from January 1, 2001 to January 1, 2023. I use the WRDS database as the basis for calculating industry concentration. Then their average returns are calculated by different months. Panel A is calculated in one month, panel B is calculated in three months, panel C is calculated in six months, panel D is calculated in nine months, and panel E is calculated in twelve months.

Table 5: Monthly Average Returns Using Total Assets as the Calculation Criteria for Industry Concentration.

Panel A: Monthly equal-weighted portfolio average return										
	1	2	3	4	5	6	7	8	9	10
ret	-0.006	-0.002	-0.002	-0.001	-0.0002	0.0005	0.000007	0.001	0.001	0.00004
Panel B: Equal-weighted portfolio average return in three months										
	1	2	3	4	5	6	7	8	9	10
ret	0.109	0.060	0.056	0.043	0.058	0.037	0.054	0.040	0.034	0.036
Panel C: Equal-weighted portfolio average return in six months										
	1	2	3	4	5	6	7	8	9	10
ret	0.191	0.111	0.104	0.079	0.095	0.068	0.100	0.076	0.064	0.071
Panel D: Equal-weighted portfolio average return in nice months										
	1	2	3	4	5	6	7	8	9	10
ret	0.269	0.166	0.154	0.123	0.126	0.100	0.151	0.116	0.097	0.107
Panel E: Equal-weighted portfolio average return in twelve months										
	1	2	3	4	5	6	7	8	9	10
ret	0.075	0.039	0.045	0.029	0.043	0.025	0.028	0.029	0.024	0.026

DOI: 10.54254/2754-1169/72/20240681

I calculate the industry concentration based on the total share of firms' assets in the industry and divide it into ten portfolios according to their share. firm data is obtained from the WRDS database from January 1, 2001 to January 1, 2023. Then their average returns are calculated by different months. Panel A is calculated in one month, panel B is calculated in three months, panel C is calculated in six months, panel D is calculated in nine months, and panel E is calculated in twelve months.

3. Conclusion

Based on the data in Tables 2 and 3, I categorize firms into ten portfolios using different measures of industry concentration (e.g., total sales and total assets). Their respective weighted average returns are then computed over different time frames. I find that, although the results are less significant, portfolios with higher industry concentration typically show higher average returns, while portfolios with lower industry concentration show lower average returns. This suggests that firms' average stock returns increase with industry concentration.

I also found that the performance of the average returns varied over different time periods. Tables 2, 3, 4 and 5 provide average returns in monthly, three-month, six-month, nine-month and twelve-month periods. Stock return performance may vary over different time frames. Average returns over long-term time frames may be more reflective of overall company performance, while average returns over short-term time frames may be more affected by market volatility.

Table 1 also provides statistical information on firms' sales and asset size, such as mean, median, maximum, and minimum values, as well as basic information about each Portfolio. These statistics can help to understand the size distribution of the firms in the sample, as well as a general picture of the sales and asset size of firms in the market.

In summary, the data in this paper illustrate the relationship between industry concentration and average returns, as well as investment performance over time. In the data processing stage, I use a more precise benchmark than in the previous literature, but the results are unsatisfactory. Possible reasons for this are that this benchmark has lost its value due to overuse by investors or that the results are not significant enough due to the effects during the epidemic. And according to Table 2 Table 4 and Table 3 Table 5, it can be found that the value of equal-weighted is more significant than value weighted. My suggestion for the direction of improvement of this paper is that the data during the epidemic period can be excluded, or due to the sic code I used to categorize the industries, and a different basis of industry categorization may be able to improve it. These findings can provide insights for investment decisions and stock portfolio construction.

References

[1] *Hou, K., & Robinson, D. T. (2006). Industry concentration and average stock returns. The journal of finance, 61(4), 1927-1956.*

[2] *Asness, C. S., & Stevens, R. (1996). Intra-industry and inter-industry variation in the cross-section of expected stock returns. Working paper, Goldman Sachs Asset Management.*

[3] *Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum?. The Journal of finance, 54(4), 1249-1290.*

[4] *Cohen, R. B., Polk, C., & Vuolteenaho, T. (2003). The value spread. The Journal of Finance, 58(2), 609-641.*

[5] *Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. The Review of financial studies, 33(5), 2019-2133.*

[6] *Almazan, A., & Molina, C. A. (2005). Intra-industry capital structure dispersion. Journal of Economics & Management Strategy, 14(2), 263-297.*

[7] *MacKay, P., & Phillips, G. M. (2005). How does industry affect firm financial structure?. The review of financial studies, 18(4), 1433-1466.*