The 2022 International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/5/20220066

Can Hedge Fund Returns Be Predicted?

Lingyan Meng[1,a,*], Xiaohe He[2,b], Mingxuan Zhou[3,c] and Xintian Han[4,d]
[1]Raffles Institution, Singapore, 575954, Singapore
[2]Department of Mathematics, New York University, NYC, 10003, U.S
[3]Department of Mathematics, University of Waterloo, Waterloo, N2L 3G1, Canada
[4]Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK
a. menglingyan2014@gmail.com, b. xh2012@nyu.edu, c. m259zhou@uwaterloo.ca, d. Xintian.Han@warwick.ac.uk
Mingxuan Zhou and Xintian Han contributed equally and should be considered co-third authors.
**corresponding author*

Abstract: After the 2008 financial crisis, hedge funds regained their popularity. Investors naturally wonder whether it is possible to predict and explain hedge fund returns just as its constituents. To answer this question, we examined hedge fund performance of 14 strategies from 2000 to 2017 by separating them into 3 groups. After deriving a statistical model, we applied it to the period of 2017-2022 and examined the errors. We observed that most strategies have a positive risk-adjusted rate of return and the current period's returns have a positive relationship with the previous period's. We concluded that monthly return has too much randomness while 3 strategies' yearly returns in the middle quantile could be predicted. More historical return data can improve the accuracy of the model.

Keywords: hedge funds, financial market, predict

1. Introduction

Modern hedge funds, as commingled investment vehicles, combine funds from multiple individual investors in the attempt to outperform the market. Similar to the financial markets, in which they invest, hedge funds' returns fluctuate over their 70 years of history. It is in the interest of their participants to predict the returns and decide on the best time to enter and exit.

In this paper we carefully examined 14 different strategies and divided each strategy into three parts: top 25%, middle 50%, and bottom 25%. We tracked hedge fund performance variations after each of three groups. Apart from monthly data, we are also considering yearly returns.

2. Literature Search

Metzger, Shenai, investigated the performance of ten notable strategies during and after the financial crisis using over 9500 hedge funds from the Hedge Index [1]. They found that although no strategy delivered a large alpha during the financial crisis, every strategy is able to outperform the benchmark S&P 500.

Stafylas, Andrikopoulosb, examined the basic raw returns of 11 strategies and 6 portfolios containing all HFs [2]. They found that hedge funds produce alpha exclusively in 'good' periods, some deliver negative alpha in 'poor' periods. Hedge funds reduce their systematic risk during 'poor' times.

DOI: 10.54254/2754-1169/5/20220066

With data from Eurekahedge, Rayner found that, for a 12-month track record, only 29% of the randomly generated track records have an average monthly return within 0.5% of true values [3,4]. Even with a more-than-10-year track record, investors still could not predict returns confidently.

Nonetheless, Avramov et al, pointed out that hedge fund returns could be predicted using macro indicators [5]. By evaluating their strategies' in-sample and out-of-sample performance, they concluded that these strategies were able to generate superior alpha, 'easily beating' the hedge fund indices.

Argyropoulos et al. took a statistical approach [6]. With increased forecast accuracy, their models could select portfolios that significantly outperform the benchmark. Stafylas et al, provided an integrated view of the implicit and statistical factor models [7]. They found out a few exposures that were valid for nearly every hedge fund strategy, such as macroeconomic risk.

Swartz and Emami-Langroodiuse 4 models in 11 risk strategies and to introduce 3 new variables [8]. A combination of two or more models may lead to a better result compared to using only one of them.

3. Data

The data on each hedge fund strategy's returns is collected from Eurekahedge.com, from January 2000 to August 2022. We included 14 strategies in total. Monthly data for the excess return on the market, risk free rate, SMB (Small Minus Big), and HML (High Minus Low) is gathered from Kenneth R. French's data library [9].

The data is then separated into two groups: (A) January 2000 to August 2017 and (B) September 2017 to August 2022. Dataset A is used to train the model and derive the parameters. Dataset B is used to test the model generated.

4. Methodology

For each strategy, we divided monthly hedge fund returns into 3 groups: the bottom quartile (worst 25%), the top quartile (best 25%), and the middle (remaining 50%). We calculated the average excess return, the standard deviation of excess return, and Sharpe Ratio for the months after three quantiles and then repeated yearly data.

We then build a linear model. Regressions were estimated as follows:

$$R_i - R_f = \alpha_i + \beta_{iM}(R_M - R_f) + \beta_{is}SMB + \beta_{ih}HML$$

Where:

- $(R_i - R_f)$ is the strategy's return less risk-free rate for that period.
- $(R_M - R_f)$ is the excess return on the market
- SMB is the size premium and HML is the value premium.

We derived our model from the Fama-French 3 factors model. Since their model explains the returns of portfolios built on stocks and bonds, they should be able to predict the portfolio of hedge funds. Using the 3 factors from the previous period, we hoped to predict the targeted period's $(R_i - R_f)$. We repeated the process for both in sample and out-of-sample testing.

DOI: 10.54254/2754-1169/5/20220066

5. Data Analysis

5.1. Descriptive Statistics of Dataset A

Table 1 provides monthly descriptive statistics of hedge fund returns from January 2000 to August 2017 (annualized). Each row is a unique strategy listed on Eureka hedge. It describes returns for months after the months in the 3 quantiles.

Table 1: The mean, standard deviation (St Dev) and Sharpe ratio (SR) of each month in three periods: bottom, middle, and top periods (01/2000-08/2017) (1 month after 1 month).

Period	Bottom			Middle			Top		
Strategy	Mean	St Dev	SR	Mean	St Dev	SR	Mean	St Dev	SR
Arbitrage	2.40%	4.12%	0.58	7.20%	2.74%	2.63	12.12%	2.63%	4.60
CTA/Managed Futures	9.48%	5.75%	1.65	7.56%	5.75%	1.31	12.72%	8.52%	1.49
Distressed Debt	-3.60%	7.72%	-0.47	11.28%	4.40%	2.56	22.68%	5.54%	4.09
Event Driven	-0.48%	8.49%	-0.06	10.92%	5.13%	2.13	16.44%	5.85%	2.81
Fixed Income	1.56%	4.92%	0.32	8.52%	2.53%	3.37	12.48%	3.33%	3.75
Long Short Equity	1.44%	8.56%	0.17	9.72%	6.75%	1.44	14.76%	6.13%	2.41
Marco	6.48%	3.43%	1.89	8.52%	3.78%	2.26	10.20%	4.50%	2.27
Multi-Strategy	4.44%	5.40%	0.82	9.12%	3.74%	2.44	14.40%	4.75%	3.03
Relative Value	4.44%	4.78%	0.93	8.28%	3.60%	2.30	14.76%	3.78%	3.91
Equity Market Neutral	5.64%	3.33%	1.70	5.40%	1.94%	2.78	5.04%	1.91%	2.65
Equity Long Bias	-1.92%	14.10%	-0.14	10.68%	9.73%	1.10	18.84%	9.15%	2.06
Trend Following	10.92%	7.97%	1.37	8.28%	8.49%	0.98	13.56%	12.12%	1.12
FX	7.20%	2.74%	2.63	6.84%	4.12%	1.66	9.96%	5.72%	1.74
Commodity	2.88%	9.21%	0.31	12.72%	8.90%	1.43	13.44%	10.60%	1.27

Note: Cells deeper in green are more favorable; cells deeper in red less favorable.

All but only 3 negative Sharpe ratios appear in the bottom period. The growth rate of funds exceeds the risk-free rate for all strategies in the month after the middle and top periods. Most strategies also have good returns in the month after a bottom period.

For 13 of 14 strategies, the mean return after a ‘good month’ is higher than that of a ‘middle month’. For 10 of 14 strategies, the mean return after a ‘middle month’ is higher than the mean return of a ‘bad month’. It follows the actual trend that bad returns are often after bad months, while good returns are often after good months.

In addition, for yearly returns (table 4 in the appendix), all 14 strategies have a higher Sharpe ratio after a good year as compared to a middle year. For 4 strategies, investing after a bad year yields the best Sharpe ratio.

DOI: 10.54254/2754-1169/5/20220066

5.2. Prediction

We examined 6 statistics to analyze the accuracy of our predictions: MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), Training R-squared, Predicted R-squared, and number of predictions.

5.3. Results of Prediction on Dataset B

Table 2 shows our model's performance on monthly returns for the middle quantile.

Table 2: The regression prediction for the middle quartile of each month, based on the previous month's factor. (1 month after middle 1 month) In the order of decreasing Predicted R-squared.

Strategy	Training R-squared	Predicted R-squared	MSE	RMSE	MAE	Number of Predictions
Arbitrage	0.015	0.229	0.00007	0.811%	0.516%	30
Relative Value	0.006	0.219	0.00024	1.545%	1.071%	30
CTA/Managed Futures	0.126	0.144	0.00010	1.023%	0.788%	30
Trend Following	0.059	0.138	0.00048	2.197%	1.676%	30
Commodity	0.040	0.135	0.00044	2.107%	1.640%	29
	0.040	0.133	0.00012	1.074%	0.890%	29
Multi-Strategy	0.016	0.084	0.00020	1.426%	1.069%	30
Distressed Debt	0.106	0.066	0.00012	1.091%	0.810%	30
Fixed Income	0.019	0.032	0.00005	0.729%	0.539%	30
Event Driven	0.097	0.010	0.00031	1.774%	1.323%	30
Equity Market Neutral	0.088	0.007	0.00005	0.671%	0.570%	29
FX	0.049	0.007	0.00013	1.119%	0.932%	30
Long Short Equities	0.012	0.005	0.00060	2.439%	1.847%	30
Equity Long Bias	0.006	<0.001	0.00117	3.424%	2.615%	30

Note:
*MSE is the mean of all (predicted **-** actual) ^2 and RMSE is the root of MSE.
*MAE is the mean of all **(**predicted – actual**)**
*Cells deeper in green are more favorable; cells deeper in red are less favorable.

For monthly data, some strategies show very small MSE and RMSE. Smaller MSE, RMSE, and MAE suggest predicted values are closer to true values. They generally follow a similar trend while MSE is more sensitive towards outliers.

Despite small errors, all predicted R-squared values are small, including those of bottom and top quartiles. This suggests our model cannot explain monthly returns.

Table 3 displays the model's performance on yearly returns also for the middle quantile. This data is calculated using the same linear regression. Instead of using monthly data, this table compounds previous 12 months' factors and returns.

We observed that yearly predictions showed significant improvement in accuracy compared to monthly predictions, symbolized by much higher R-squared values. This increase in R squared may be due to yearly returns cancel out some monthly fluctuations and allow less randomness. 5 strategies have both R-squared greater than 0.6, suggesting our model can largely explain their variations from 2000 to 2022.

Comparatively small MSE, RMSE, and MAE occur on Arbitrage and Fixed Income. Their predicted R-squared are 0.6-0.7, but training R-squared are 0.4-0.5. This significant improvement in R-squared should not be attributed to the accuracy of our model. Instead, it is probably due to the lack of outliers.

Table 3: The regression prediction for the middle quartile of each year, based on the previous year's factor. (1 year after middle 1 year) In the order of decreasing Predicted R-squared.

Strategy	Training R-squared	Predicted R-squared	MSE	RMSE	MAE	Number of Predictions
Distressed Debt	0.629	0.862	0.00242	4.920%	3.844%	32
Relative Value	0.580	0.828	0.00231	4.804%	4.247%	30
Event Driven	0.769	0.807	0.00240	4.901%	3.713%	28
Equity Long Bias	0.744	0.796	0.00767	8.760%	7.180%	28
Multi-Strategy	0.610	0.701	0.00303	5.501%	4.910%	30
Fixed Income	0.474	0.683	0.00162	4.021%	3.284%	27
Long Short Equities	0.716	0.643	0.00548	7.405%	6.084%	30
Arbitrage	0.448	0.629	0.00124	3.525%	3.003%	32
Macro	0.344	0.535	0.00232	4.815%	3.322%	31
Commodity	0.064	0.441	0.00677	8.230%	6.799%	26
FX	0.118	0.367	0.00251	5.011%	4.379%	32
Equity Market Neutral	0.159	0.233	0.00269	5.191%	4.593%	31
Trend Following	0.431	0.057	0.01589	12.607%	10.743%	20
CTA/Managed Futures	0.495	<0.001	0.00591	7.685%	6.497%	19

Out of the 14 strategies, Distressed Debt, Relative Value, and Event Driven are highly predictable by our models. Both of the R-squareds are high, suggesting our model well explains actual data not only from 2000 to 2017 but also from 2017-2022. These strategies present comparatively small MSE, RMSE, and MAE. Nonetheless, MSE, RMSE, and MAE are still high in absolute value. In fact, it is possible to coexist with large R-squared. This situation implies that actual yearly variations are big, so that the sum of squared total (SST) increases and so does R-squared.

In our model, Trend Following and CTA/Managed Futures show a significant decrease in R-squared from Dataset A to B. It may be due to their trend-following nature. Trend-following strategies are able to generate returns regardless of direction of the trend. As our model only considers factors from the previous period, it does not capture the time series aspect of the factors. Our model's low explanatory power to them allies with the paper by Akindynos-Nikolaos Balta and Robert Kosowski [10].

DOI: 10.54254/2754-1169/5/20220066

5.4. Graphs

We plotted the prediction vs. actual value graphs for typical strategies in Figure 2. The line in the middle is where predicted values equal actual values. The closer the points are to the middle, the better the prediction.

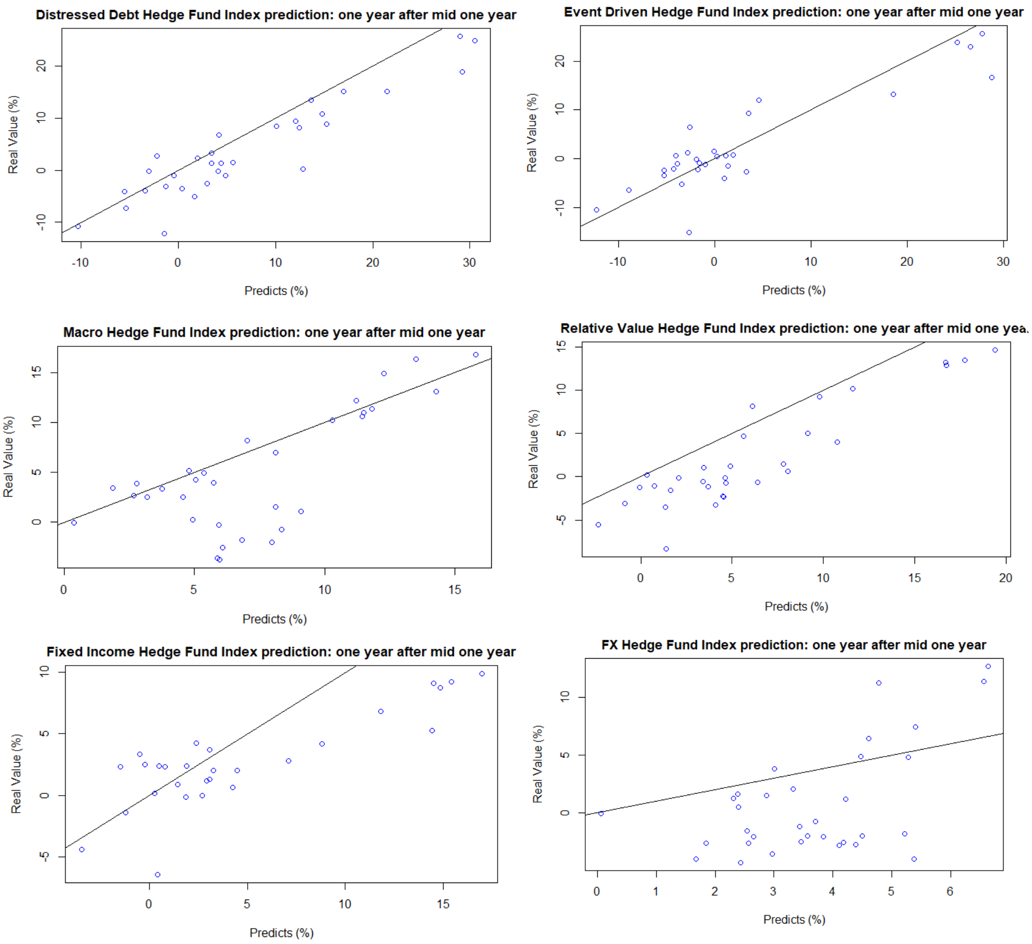

Figure2: Actual vs Predicted Plot for strategies' yearly return after a middle year.

For Distressed Debt and Relative Value, although most points are close to the line, there are much more points to the right of the line, suggesting that our model consistently overestimates their returns. Strategies’ returns may decrease over time due to capacity constraints, resulting in systematic overestimation.

Regarding Event driven, Fixed Income, and Macro, our model fails to predict extreme statistics. It consistently overestimates the returns of Event Driven when return is greater than 10% or less than –10%. Same overstatements happen when Fixed Income returns are more than 5% and Macro returns are less than 0%. Adding more factors to our model helps to capture those cases and make better predictions.

Some strategies may follow a logarithmic relationship instead of a linear one as our model. In strategies such as CTA/Managed Futures, Equity Market Neutral, and Commodities, there does not exist a line that could fit all the points. It means the trend of their returns has more complexity than what a linear model can capture

Another case can be represented by FX, as its actual returns fall into a smaller range compared to others. FX shows a higher standard deviation than most of others. Thus, there is more random noise in its returns, increasing difficulty to predict.

DOI: 10.54254/2754-1169/5/20220066

Overall, graphical illustration allied with our previous implications from Table 3. Comparatively, Distressed Debt, Relative Value, and Event Driven are predicted by our model in a more accurate way. Other strategies, with worse statistics in Table 3 indicating non-reliability of our predictions, reveal points farther to the line in graphs.

For our predictions of top and bottom years, fewer strategies show an ideal combination of the R-squareds and errors. This may be attributed to various reasons.

Firstly, the middle quantile has a sample size twice as large as the other two. Increased training data could increase accuracy. Secondly, financial markets are more volatile during best and worst performing periods, resulting in increased standard deviation of returns and more outliers. Thirdly, hedge fund performance may be attributed to luck and randomness. Thus, their performance is more unpredictable after a particularly good or bad time period. In addition, getting a top or bottom return may change managers' behaviors. The direction of their behaviors can be binary and then the next period returns are hard to predict. Manager may get complacent or discouraged by extreme results. By contrast, managers are less affected by a period of middle returns, and more likely to follow their original plan, making returns more stable and predictable.

6. Conclusion

Across all hedge funds, almost all strategies have their average net value growth rate exceeding the risk-free rate. Most of them yield a greater average return after a good period and a lower average return after a bad period. However, we could not accurately predict a hedge fund's monthly returns using our model based on Fama/French 3 factors. Meanwhile, for Distressed Debt, Relative Value, and Event Driven, we can predict their yearly returns for the middle quantile to an extent of about 70%. In regards to yearly hedge fund returns as a whole, our model shows greater accuracy in predicting the middle quantile compared to top and low quartiles, possibly due to the fact that the training dataset is doubled in size. As a result, more future data is required to test whether our model for top and low years is useful.

Acknowledgement

Mingxuan Zhou and Xintian Han contributed equally to this work and should be considered co-third authors.

References

[1] *Metzger, Nicola, and Vijay Shenai. "Hedge Fund Performance during and after the Crisis: A Comparative Analysis of Strategies 2007-2017." Int. J. Financial Stud 7, no. 1: 15 (2019), https://doi.org/10.3390/ijfs7010015*

[2] *Stafylas, Dimitrios, and Athanasios Andrikopoulosb. "Determinants of hedge fund performance during 'good' and 'bad' economic periods." Research in International Business and Finance 52, (2020), https://doi.org/10.1016/j.ribaf.2019.101130*

[3] *Eurekahedge. "Eurekahedge Hedge Fund Index." Accessed September 14, 2022. https://www.eurekahedge.com/Indices/IndexView/Eurekahedge/474/Eurekahedge-Hedge-Fund-Index*

[4] *Rayner, Ian. "Hedge Fund Return Predictability." Rayner Gorman. Last modified June 13, 2018, https://www.raynergobran.com/2018/06/hedge-fund-return-predictability/*

[5] *Avramov, Doron, Laurent Barras, and Robert Kosowski. "Hedge Fund Return Predictability under the Magnifying Glass." The Journal of Financial and Quantitative Analysis 48, no. 4 (2013): 1057-83, https://www.jstor.org/stable/43303833*

[6] *Argyropoulos, Christos, Ekaterini Panopoulou, Nikolaos Voukelatos, and Teng Zheng. "Hedge fund return predictability in the presence of model risk*." The European Journal of Finance, (2022), https://doi.org/10.1080/1351847X.2021.2020146*

[7] *Stafylas, Dimitrios, Keith Anderson, and Moshfique Uddin. "Recent advances in explaining hedge fund returns: Implicit factors and exposures." Global Finance Journal 33, (2017): 69-87, https://doi.org/10.1016/j.gfj.2016.08.001*

DOI: 10.54254/2754-1169/5/20220066

[8] *Swartz, L. Mick, and Farrokh Emami-Langroodi. "Relative Value Hedge Funds: A Behavioral Modeling of Hedge Fund Risk and Return Factors." Journal of Behavioral Finance, (2018): 19:4, 462-482. https://doi.org/10.1080/15427560.2018.1434654*
[9] *The Tuck School of Business at Dartmouth College. Kenneth R. French. " Fama/French 3 Factors." Accessed September 9, 2022. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html*
[10] *Baltas, Nick, and Robert Kosowski. "Momentum Strategies in Futures Markets and Trend-following Funds." SSRN, (2013), http://dx.doi.org/10.2139/ssrn.1968996*

Appendix

Table 4: The mean, standard deviation (St Dev) and Sharpe ratio (SR) of each year in three periods: bottom, middle, and top periods (01/2000-08/2017) (1 year after 1 year).

Period	Low			Median			High		
Strategy	Mean	St Dev	SR	Mean	St Dev	SR	Mean	St Dev	SR
Arbitrage	9.55%	5.48%	1.74	5.32%	4.32%	1.23	6.93%	3.30%	2.10
CTA/Managed Futures	6.65%	7.15%	0.93	10.19%	7.98%	1.28	9.64%	5.64%	1.71
Distressed Debt	11.92%	17.52%	0.68	9.77%	11.81%	0.83	13.61%	7.35%	1.85
Event Driven	13.97%	13.40%	1.04	7.80%	9.52%	0.82	9.14%	8.30%	1.10
Fixed Income	8.00%	10.29%	0.78	7.27%	4.94%	1.47	8.91%	2.88%	3.10
Long Short Equity	13.76%	8.74%	1.57	6.36%	9.01%	0.71	8.34%	9.06%	0.92
Marco	7.26%	3.95%	1.84	8.83%	5.74%	1.54	9.03%	4.43%	2.04
Multi-Strategy	9.19%	5.75%	1.60	8.87%	7.64%	1.16	10.67%	7.03%	1.52
Relative Value	9.67%	7.20%	1.34	8.60%	6.54%	1.32	8.46%	5.10%	1.66
Equity Market Neutral	5.44%	1.95%	2.78	4.57%	2.30%	1.99	5.37%	2.49%	2.16
Equity Long Bias	17.22%	15.11%	1.14	6.29%	14.69%	0.43	9.49%	13.11%	0.72
Trend Following	5.25%	9.06%	0.58	10.54%	11.24%	0.94	14.29%	9.12%	1.57
FX	7.04%	3.81%	1.85	6.30%	4.18%	1.51	10.33%	5.39%	1.92
Commodity	8.13%	13.16%	0.62	11.35%	14.88%	0.76	14.06%	14.98%	0.94