Can factor investing disrupt the hedge fund industry?

Mirroring "Global Macro" Hedge Funds performance through factor-based predictive modelling and ensemble learning

> Presented by: Gabriel Ignacio Jason Zhang



Hedge Funds

- Wealthy individuals or institutional investors
- Significant diversification benefits
- "Absolute returns" (positive risk-adjusted returns)

Has anyone invested in hedge funds before?



Sources:

US Global Macro Hedge Funds



3 trillion USD CAGR 8%-10% 5-10 years



14,000 US firms

Top fund uses global macro strat
George Soros - \$1 billion overnight



Each of the top 35 US hedge funds > entire Canadian hedge funds industry (200 hedge funds, 200 billion in AUM)

Sources:

Banquier, S. (2019). Overview of Prime Brokerage. Presented at Smith School of Business, Queen's University.

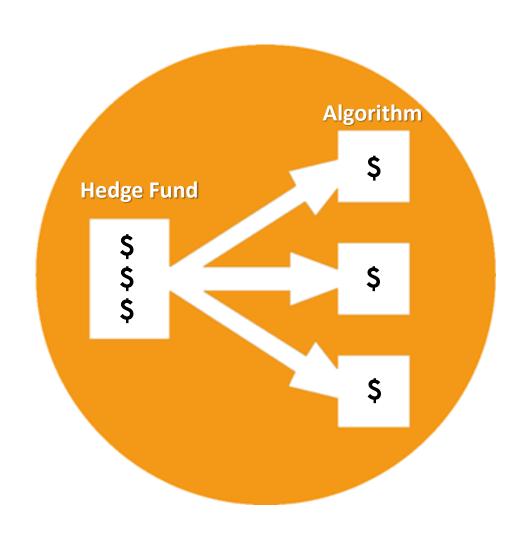
Fung, W., & Hsieh, D. A. (2011). The risk in hedge fund strategies: Theory and evidence from long/short equity hedge funds. Journal of Empirical Finance, 18(2011), 547-569.

Research Question

Are "Global Macro" Hedge Funds replicable using the factor-based frameworks for non-accredited investors (like you and me)?

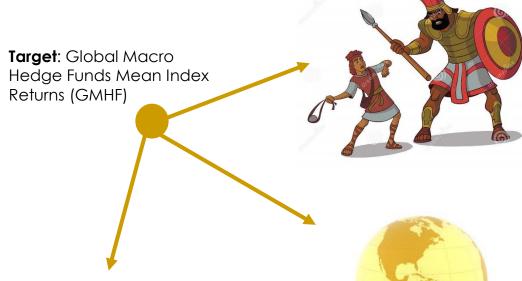
Objectives

- **Dynamic linkages**: can market variables predict GMHF performance?
- **Neutrality**: or do dynamic investment strategies (i.e. manager skill) play a factor?
- Implication: systematically replicate results without Hedge Fund fees



Data Exploration





Fama-French-Carhart 4 Factor Model
Market returns (Mkt), size factor (SMB), value
factor (HML) and momentum (MOM)



Risk-based Factor Model

Bond Trending following risk factors Equity-oriented risk factors Bond-oriented risk factors



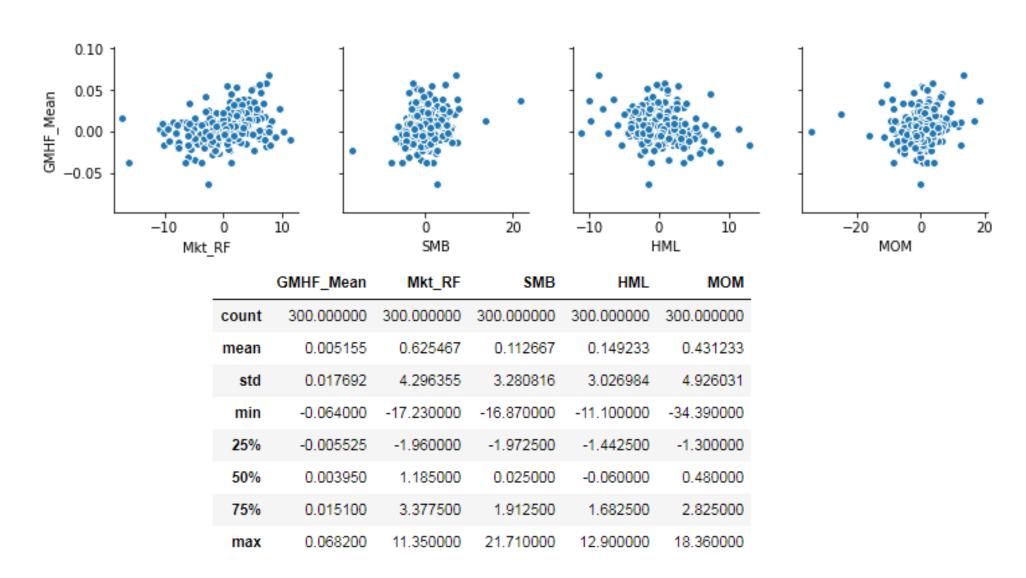
Macroeconomic Factor Model

GDP, inflation rate, interest rates, unemployment rates, etc

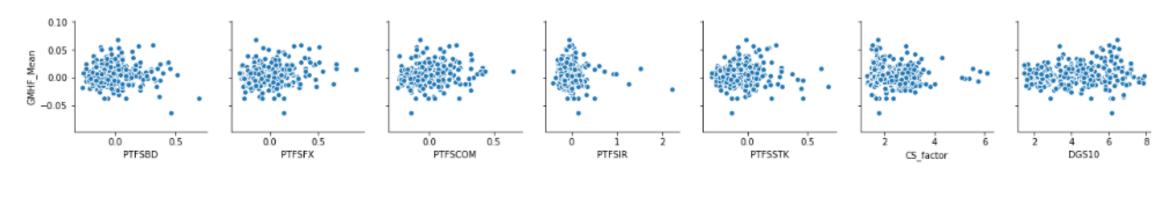
Observations: 1994-2018, 300 monthly data points

Source: Hedge Fund Research Institute, Ken French Library, David Hsieh Library, US government data

Data Exploration - 4 factor model

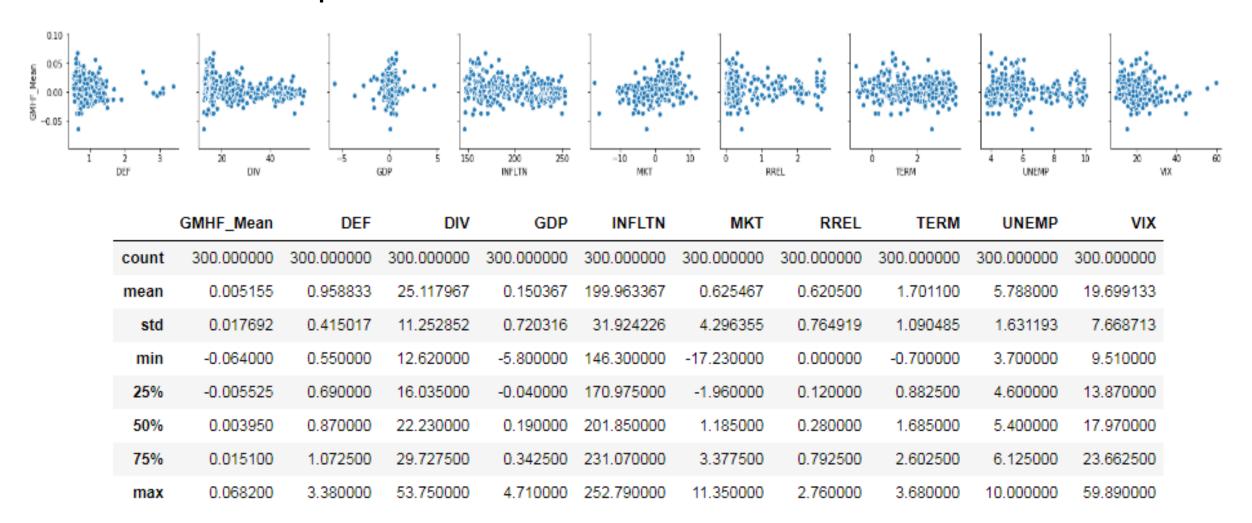


Data Exploration – Risk Based Model



	GMHF_Mean	PTFSBD	PTFSFX	PTFSCOM	PTFSIR	PTFSSTK	CS_factor	DG \$10
count	300.000000	300.00000	300.000000	300.000000	300.000000	300.000000	300.000000	300.000000
mean	0.005155	-0.01890	-0.008167	-0.005833	-0.020800	-0.048400	2.421133	4.133867
std	0.017692	0.15202	0.193697	0.140519	0.249809	0.141559	0.786240	1.636767
min	-0.064000	-0.27000	-0.320000	-0.250000	-0.420000	-0.300000	1.300000	1.460000
25%	-0.005525	-0.13000	-0.150000	-0.100000	-0.140000	-0.150000	1.770000	2.667500
50%	0.003950	-0.05000	-0.050000	-0.030000	-0.070000	-0.070000	2.335000	4.095000
75%	0.015100	0.04000	0.082500	0.060000	0.040000	0.020000	2.860000	5.422500
max	0.068200	0.69000	0.900000	0.650000	2.220000	0.670000	6.100000	7.910000

Data Exploration – Macroeconomic Model



Model Building Process

Feature Selection	Numeric Variables
Fama-French-Carhart	'Mkt_RF','SMB','HML','MOM'
Risk-based	'PTFSBD','PTFSFX','PTFSCOM','PTFSIR','PTFSSTK','DGS10','CS_factor'
Macroeconomic	'DEF','DIV','GDP','INF','MKT','RREL','TERM','UNEMP','VIX'



Algorithms Comparison	Regression	Classification
Linear Algorithms	LinearRegression	LogisticRegression
Non-Linear Algorithms	KNeighborsRegressor, SVR	KNeighborsClassifier, DecisionTreeClassifier, GaussianNB, SVC
Emsemble Algorithms	GradientBoostingRegressor, RandomForestRegressor	VotingClassifier, RandomForestClassifier





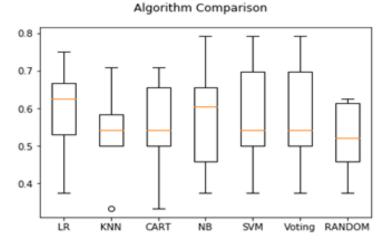
Feature Importance Use Algorithm RandomForest method best_estimator.feature_importances

Classification

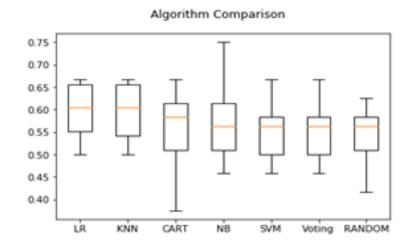
Fama-French-Carhart

Algorithm Comparison 0.7 0.6 0.5 0.4 LR KNN CART NB SVM Voting RANDOM

Risk-based



Macroeconomic



Best parameter (CV score=0.625):

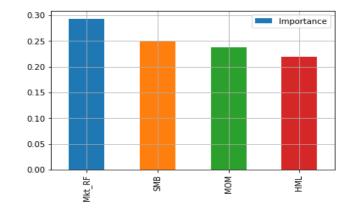
{'logistic_C': 35.564803062231285, 'logistic_penalty': '12', 'pca_n_components': 2}

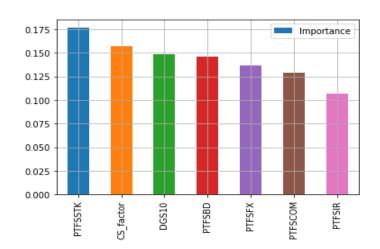
LG Model test Accuracy:: 0.68333333333333333

Best parameter (CV score=0.637):

{'logistic_C': 0.0001, 'logistic_penalty': '12', 'pca_n_components': 4}

LG Model test Accuracy:: 0.76666666666666667

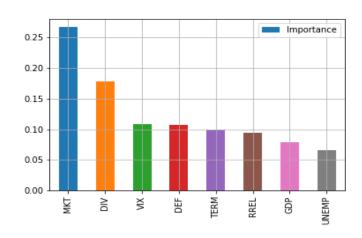




Best parameter (CV score=0.646):

{'logistic_C': 0.013257113655901081, 'logistic_penalty': '12', 'pca_n_components': 3}

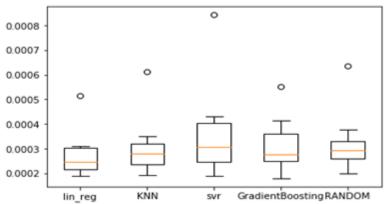
LG Model test Accuracy:: 0.63333333333333333



Regression

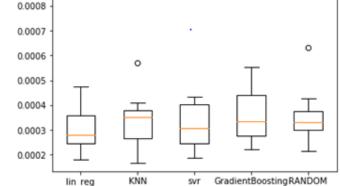
Fama-French-Carhart

Algorithm Comparison



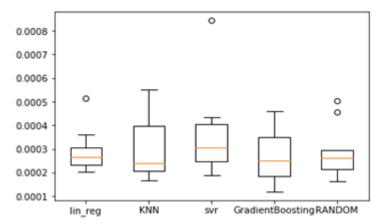
Risk-based





Macroeconomic





Best parameter (CV score=0.074):

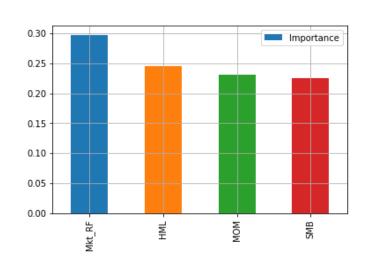
{'pca_n_components': 4}

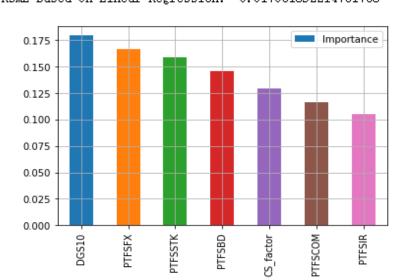
RSME Based On Linear Regression: 0.01536944488278775

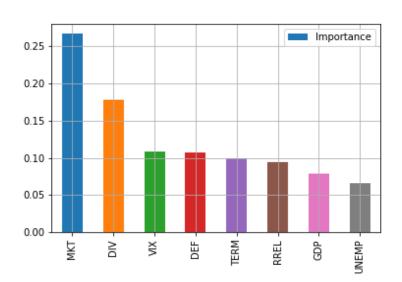
Best parameter (CV score=-0.001):

{'pca_n_components': 7}

RSME Based On Linear Regression: 0.017031892214731736







Best parameter (CV score=0.137):

['GradientBoosting_learning_rate': 0.01, 'GradientBoosting_max_depth': 4, 'GradientBoosting_n_estimators': 100, 'GradientBoosting_subsam ple': 0.5, 'pca_n_components': 6}

RSME Based On Gradient Boosting: 0.015977497904660194

Summary Results: Best Performing Model

- Evaluated each factor model using classification (accuracy) and regression (RMSE)
- The best model is chosen based on accuracy and RMSE

Models	Classification Accuracy	Regression RMSE	
Fama-French-Carhart 4-factor model	77% Logistic Regression	0.0153 Linear Regression	
Risk-based factor model	69% Logistic Regression	0.0170 Linear Regression	
Macroeconomic factor model	63% Logistic Regression	0.0159 Gradient Boosting	

Fama-French-Carhart has the best predictive power across the 3 models using logistic regression and Linear Regression. Not surprisingly, AQR also uses the same factor model

Conclusion

- Fama-French-Carhart best explain Global Macro Hedge Funds' performance among all three frameworks
- For investors, market performance (Mkt), value premium (HML) and momentum (MOM) are key considerations
- For Hedge Fund Managers, exploiting these indicators might be key to drive alpha





Appendices



Frameworks Used

Frameworks	Theoretical	Empirical	
Fama-French-Carhart Four Factor Model	Market, size, value and momentum explain stock performance	$(R_i - R_f) = \alpha_i + \beta_{\text{MKT}} (R_{MKT} - R_f) + \beta_{\text{SMB}} (\text{SMB}) + \beta_{\text{HML}}$ $(\text{HML}) + \beta_{\text{UMD}} (\text{UMD}) + e_i$	
Trend Following Factors	Trend-Following Risk Factors (3): -Bond Trend-Following Factor -Currency Trend-Following Factor -Commodity Trend-Following Factor Equity-oriented Risk Factors (2): -Equity Market Factor -The Size Spread Factor Bond-oriented Risk Factors (2): -The Bond Market Factor -The Credit Spread Factor	$(R_i - R_f) = \alpha_i + \beta_1$ (PTFSBD)+ β_2 (PTFSFX) + β_3 (PTFSCOM) + β_4 (PTFSIR) + β_5 (PTFSTK) + β_6 (DGS10) + β_7 (CS_factor) + e_i	
Macroeconomic Factors	Default spread, aggregate dividend yield, the growth rate of real Gross Domestic Product (GDP) per capita, inflation rate, equity market index, term spread, short-term interest rate changes, and unemployment rate	$(R_i - R_f) = \alpha_i + \beta_1 (DEF) + \beta_2 (DIV) + \beta_3 (GDP) + \beta_4$ $(INFLTN) + \beta_5 (MKT) + \beta_6 (RREL) + \beta_7 (TERM) + \beta_7$ $(UNEMP) + e_i$	
Volatility	Volatility index	$(R_i - R_f) = \alpha_i + \beta_1 \text{(VIX)} + e_i$	

Fama-French-Carhart 4 Factor Model (combination of market risks)

Variables	Description	# of observations	Timeframe	Corrected to	Source
Mkt-RF	Rm-Rf, the excess return on the market	1,111 monthly data points	July 1926 to Jan 2019	Jan 1994 to Dec 2018	Kenneth R. French - Data Library
SMB	Size factor: small stocks relative to large stocks	1,111 monthly data points	July 1926 to Jan 2019	Jan 1994 to Dec 2018	Kenneth R. French - Data Library
HML	Value factor: value stocks relative to growth stocks	1,111 monthly data points	July 1926 to Jan 2019	Jan 1994 to Dec 2018	Kenneth R. French - Data Library
МОМ	Tendency for the stock price to continue rising if it is going up and to continue declining if it is going down	1,105 data points Monthly	Jan 1927 to Jan 2019	Jan 1994 to Dec 2018	Kenneth R. French - Data Library

Trend-Following factors under Fung and Hsieh (directional / trend risks)

Variables		Description	# of observations	Timeframe	Corrected to	Source
PTFSBD		Return of PTFS Bond lookback straddle	301 data points	Jan 1, 1994 to Jan 1, 2019	Jan 1994 to Dec 2018	David A. Hsieh's Data Library: Hedge Fund Risk Factors
PTFSF	:X	Return of PTFS Currency Lookback Straddle	301 data points	Jan 1, 1994 to Jan 1, 2019	Jan 1994 to Dec 2018	David A. Hsieh's Data Library:
PTFSCOM		Return of PTFS Commodity Lookback Straddle	301 data points	Jan 1, 1994 to Jan 1, 2019	Jan 1994 to Dec 2018	Hedge Fund Risk Factors
PTFSIR		Return of PTFS Short Term Interest Rate Lookback Straddle	301 data points	Jan 1, 1994 to Jan 1, 2019	Jan 1994 to Dec 2018	David A. Hsieh's Data Library:
PTFSS	тк	Return of PTFS Stock Index Lookback Straddle	301 data points	Jan 1, 1994 to Jan 1, 2019	Jan 1994 to Dec 2018	Hedge Fund Risk Factors
The Credit	DBAA	Moody's Baa yield	398 data points 33 years Monthly obs	Jan 1, 1986 to Jan 1, 2019	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis
Spread Factor	DG\$10	10-year treasury constant maturity yield	686 data points	Jan 1, 1962 to Feb 1, 2019	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis
The Bond Market Factor	DG\$10	The monthly change in the 10-year treasury constant maturity yield	686 data points	Jan 1, 1962 to Feb 1, 2019	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis

Macroeconomic variables by Bali, Brown and Caglayan (2014)

Variables	Description	# of observations	Timeframe	Corrected to	Source
DEF	default spread measured as the difference between yields on BAA-rated and AAA-rated corporate bonds	1,202 data points 100 years Monthly obs	Jan 1, 1919 to Feb 1, 2019	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis
DIV	aggregate dividend yield on the Standard&Poor's (S&P)500Index	1,768 data points	Jan 1871 to Dec 2018	Jan 1994 to Dec 2018	Robert Shiller's online data library
GDP	U.S. monthly Growth rate of real GDP per capita	719 data points	Feb 1, 1959 to Dec 1, 2018	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis
INF	INF: monthly inflation rate based on the U.S. consumer price index	865 data points	Jan 1, 1947 to Jan 1, 2019	Jan 1994 to Dec 2018	Robert Shiller's online data library
MKT	MKT: excess return on the value-weighted NYSE/Amex/ Nasdaq (CRSP) equity market index	666 monthly data points	July 1963 to Dec 2018	Jan 1994 to Dec 2018	Kenneth R. French - Data Library
RREL	RREL: relative T-bill rate, defined as the difference between the three- month T-billrateandits 12- monthbackwardmoving average	1,022 monthly data points	Jan 1, 1934 to Feb 1, 2019	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis
TERM	TERM: term spread measured as the difference Between yields on ten-year and three-month Treasury securities	446 monthly data points	Jan 1, 1982 to Feb 1, 2019	Jan 1994 to Dec 2018	Federal Reserve Bank of St. Louis
UNEMP	UNEMP: the U.S. monthly unemployment rate defined as the number of unemployed as a Percentage of the labor force	854 monthly data points	Jan 1, 1948 to Feb 1, 2019	Jan 1994 to Dec 2018	US Bureau of Labor and Employment Statistics

Volatility Index (market risks)

Variables	Description	# of observations	Timeframe	Corrected to	Source
VVIX Index	CBOE simple proxy for uncertainty, calculated uncertainty betas for individual stocks using the volatility of their implied option volatilities	344 data points Monthly obs	June 30, 1986 to June 29, 2018	Jan 1994 to Dec 2018	Chicago Board Options Exchange (CBOE®) website

Global Macro Hedge Fund Index (Dependent Variable Y)

Variables	Description	# of observations	Timeframe	Corrected to	Source
HFRI GMHF Index	Index mean return on Global Macro Hedge Funds in the US	300 data points	Jan 1, 1994 to Dec 1, 2018	Jan 1994 to Dec 2018	Hedge Fund Research Index

References

Bali, T. G., Brown, S. J., & Caglayan, M. O. (2014). Macroeconomic risk and hedge fund returns. *Journal of Financial Economics*, 114(2014), 1–19.

Banquier, S. (2019). Overview of Prime Brokerage. Presented at Smith School of Business, Queen's University.

Board of Governors of the Federal Reserve System (US) (2019). Total Consumer Credit Owned and Securitized, Outstanding [TOTALSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/TOTALSL.

Carhart, M. (1997). On Persistence in Mutual Fund Performance. The Journal of Finance, 52(1997), 1-26.

Chawla, R. K., & Uppal, S. Household Debt in Canada. Retrieved from https://www150.statcan.gc.ca/n1/pub/75-001-x/2012002/article/11636-eng.htm.

Edwards, F. R., & Gaon, S. (2003). Hedge Funds: What Do We Know? Journal of Applied Corporate Finance, 15(4), 58-71.

El Kalak, I., Azevedob, A., & Hudson, R. (2016). Reviewing the hedge funds literature II: Hedge funds' returns and risk management characteristics. *International Review of Financial Analysis* 48(2016) 55-66.

Fama, E. F., & French, K. R. (2015). A Five-Factor Asset Pricing Model. Journal of Financial Economics, 116(2015), 1–22.

French, K. (2018). Fama-French 5 Factors (2x3). Retrieved from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5_factors_2x3.html.

Fung, W., & Hsieh, D. A. (1999). A primer on hedge funds. Journal of Empirical Finance, 6(1999), 309-331.

Fung, W., & Hsieh, D. A. (2011). The risk in hedge fund strategies: Theory and evidence from long/short equity hedge funds. *Journal of Empirical Finance*, 18(2011), 547-569.

Grant, T. (2019). Alternative Investing. Retrieved from http://v1.theglobeandmail.com/partners/free/srinvest/art7_hedge.html.

Hedge Fund Research Index. (2019). HFRI® Indices Performance Tables. Retrieved from https://www.hedgefundresearch.com/family-indices/hfri

References

Hsieh, D. A. (2019). David A. Hsieh's Data Library: Hedge Fund Risk Factors. Retrieved from https://faculty.fuqua.duke.edu/~dah7/HFRFData.htm

Isenberg School of Management. (2018). CISDM Equity Long/Short Index (xlsx). Retrieved from https://www.isenberg.umass.edu/resources/downloads/download-cisdm-equity-longshort-index.

Ji, X., Martin, J., & Yao, Y. (2017). Macroeconomic risk and seasonality in momentum profits. Journal of Financial Markets 36(2017)76–90

Krause, T.A. (2019). Hedge fund returns and uncertainty. North American Journal of Economics and Finance 47 (2019) 597–601.

L'Her, J. F., Masmoudib, T., & Suret, J. M. Evidence to support the four-factor pricing model from the Canadian stock market. *International Financial Markets, Institutions and Money, 14* (2004) 313–328.

Namvar, E., Phillips, B., & Pukthuanthong, K., Raghavendra, P. (2016). Do hedge funds dynamically manage systematic risk? Journal of Banking & Finance, 64 (2016) 1–15.

Panopoulou, E., & Vrontos, S. (2015). Hedge fund return predictability: To combine forecasts or combine information. Journal of Banking & Finance, 56 (2015) 103–122.

Patton, A. J. (2009). Are "Market Neutral" Hedge Funds Really Market Neutral? Review of Financial Studies, 22(7), 2495-2530.

Racicot, F.E., & Theoret, R. Multi-moment risk, hedging strategies, & the business cycle. *International Review of Economics and Finance 58*(2018) 637–675.

Slavutskaya, A. (2013). Short-term hedge fund performance. *Journal of Banking & Finance, 37* (2013), 4404–4431.

Statisa. (2017). Assets of the Hedge Funds Worldwide. Retrieved from https://www.statista.com/statistics/271771/assets-of-the-hedge-funds-worldwide/.

Stafylas, D., Anderson, K., & Uddin, M. (2017). Recent advances in explaining hedge fund returns: Implicit factors and exposures. Global Finance Journal, 33(2017) 69–87

Wigglesworth, R. (2018). Can factor investing kill off the hedge fund? Financial Times. Retrieved from https://www.ft.com/content/2b3e2eaa-6fe6-11e8-92d3-6c13e5c92914