Leibniz University Hannover
School of Economics and Management
Institute of Banking and Finance

Leibniz Universität Hannover

Hedge Funds: Trading Strategies and Performance Evaluation

(Belegnummer: 374040)

Thema:

Hedge Funds: Impact of Size and Cycle Phase on Performance

Prüfer: Prof. Dr. Maik Dierkes
Betreuer: Dr. Sebastian Schrön

vorgelegt von:

Dipl.-Kfm. Hans-Christian Schneider
Am Borberg 21, 31787 Hameln
Gasthörer

Hannover, den 22.02.2022

Table of Contents

List of Tables

Abbreviations and Symbols

ALPHA	Excess Return of Financial Instruments (> Excess Return of Market Portfolio)
BETA	Volatility of Financial Instruments as Compared to Market Volatility
df	Degrees of Freedom
E	Expected Value
EXC	Excess Return on Value at Risk
HF	Hedge Fund
IBF	Institute of Banking and Finance at Leibniz University Hannover
IR	Information Ratio
LPMPM	Lower Partial-Moments Performance Measure
MAR	Mimimum Acceptable Rate of Return
MWU	Mann-Whitney U Test
n	Number of
NBER	National Bureau of Economic Research, USA
n obs.	Number of Observations
RET	Gross Return
RETRF	Excess Return (> RF)
RF	Risk-Free Rate
SOR	Sortino Ratio
SR	Sharpe Ratio
USLSHF	US Long/Short Hedge Fund
VaR	Value at Risk

α	Alpha: Excess Return (see above)
β	Beta: Measure of Market Risk (see above)
ϵ	Epsilon: Error Term (e.g. in CAPM)
σ	Sigma: Standard Deviation
Σ	Capital Sigma: The Sum of

1. Introduction

Capital market research is manifold concerning focus, perspectives, objectives and and and … This certainly applies to hedge funds as well. As a consequence every interesting study raises new topics relating to focus, perspectives, objectives and and and …

At 2021/22-lecture on hedge funds at Leibniz University Hannover a thesis topic was proposed concerning fund age and size as determinants of hedge fund performance. Specified literature refers to Jones (2007), who studied the impact of fund size and age on performance. Jones (2007) analyzed portfolios of funds. Findings of Jones (2007) seemed interesting to us. We wondered, however, if her *portfolio* focus can be supplemented by a *single fund* perspective.

The proposed thesis topic refers to fund age and size. Regretfully available data did not permit a comprehensive and differentiated analytical pursuit of size impact. Besides age we thus considered another second factor for our impact analysis. We looked for a factor that was (1) not already proposed for other topics of our lecture, for which (2) data are freely available and that (3) would plausibly have impact on hedge fund performance. Motivated by Stafylas & Andrikopoulos (2020) we chose cycle phase as a second factor for our impact analysis.

Fund data are most prevailingly available for US hedge funds in our data set. Concerning investment style the largest number is represented by US long/short equity hedge funds. We thus focused on US long/short equity funds. Our focus on this fund class has a particular relation to cycle phase, our second impact factor: Funds with an explicit *short* emphasis should do better in recessionary phases as other funds.

The main part of our survey is divided into two chapters.

We start Chapter 2 with terms and concepts that are important for the understanding of our study. Afterwards hypotheses are developed by shortly introducing related research. Chapter 2 will be finished by presenting and describing our data.

Chapter 3 covers performance measures of hedge funds and implements a selection for our survey. We take up several questions outlined in the task description of this thesis: How do hedge funds perform when different measures are applied? Do hedge fund rankings under different measures generate similar or different results? Is there an impact of age and cycle phase on hedge fund performance?

Chapter 3 is subdivided als follows:

- Sections 1 and 2 outline definitions of performance measures and analyze rankings of hedge funds under different .
- Sections 3 to 5 addresses performance impact of fund age and cycle phase.
- Section 6 expands on hedge fund performance in a multivariate context.

We finally sum up our findings and give an outlook.

2. Methological Basis of Our Survey

2.1. Terms and Concepts

Hedge funds (HFs)[1] are generally considered as alternative investments. They are pooled investment devices that make extensive use of complex trading and portfolio construction techniques such as short selling and derivatives. HFs are distinct from private-equity funds and other similar closed-end funds: Hedge funds generally invest in relatively liquid assets.

Hedge funds employ a wide variety of financial instruments and risk management techniques. Nonetheless can they be very different from each other with respect to their strategies, expected returns and risks. Regulators by and large restrict hedge fund marketing to institutionals, high net worth individuals and other sufficiently sophisticated investors.

At the top level, Pedersen differentiates between *equity strategies, macro strategies* and *arbitrage strategies*.[2] The former ones - equity strategies - are subdivided into *discrectionary long/short equity, dedicated short bias* and *quant equity*.[3]

The goal of *long/short equity hedge funds* is to invest in stocks with superior return characteristics and to disinvest or shorten stocks with inferior return profiles. Put simply, there are three potential outcomes. The most favourable outcome is when long positions rise in value and short positions decline - known as *double alpha*. The other outcomes - *single alpha* and *double splat* - are self-explanatory.

2.2. Development of Hypotheses

Hedge fund (HF) literature is rather broad. It i.a. covers the relationship between HF performance and various factors, e.g. investment strategies.

Investment Strategies

Numerous studies found out a significant impact of investment strategy on fund performance. We suppose that we will get the same result for our sample:

[1] For an introduction to HFs see e.g. Bodie (2014), p. 926 - 949.
[2] Pedersen (2015), p.8.
[3] Ibid., p. 9.

H1: US long/short HFs (USLSHFs) feature a significantly different performance profile compared to other funds.

Performance Measures

Our first hypothesis raises the question of *HF performance measurement*: Does HF performance depend on the way, it is measured?

Eling & Schuhmacher (2007) conducted an empirical study and compared Sharpe ratio to 12 other performance measures. They found out that the other performance measures generated an almost identical rank ordering across hedge funds. We will replicate this study with a smaller selection of performance measures and more current data. Supposedly we will get the same result:

H2: Different performance measures generate almost identical orders compared to Sharpe ratio.

Single Determinants of HF Performance

Our survey takes up the question which single fund characteristics are important to explain differences in HF performance.

Jones (2007) analyzed whether younger HFs offer stronger performance than older HFs. Her results indicate that younger funds perform better. Concerning fund age Stafylas & Andrikopoulos (2020) cite several studies.[4] Those studies for the most part also found out that young funds outperform old ones. Stafylas & Andrikopoulos (2020) quote one single study with digressive findings[5]. Their own findings support the majority result. On this basis we suppose that USLSHFs conform to the majority result:

H3: Young USLSHFs deliver better performance as older ones.

The notion that HFs always deliver positive alpha is named “conventional wisdom” by Stafylas & Andrikopoulos (2020).[6] They test this notion and find out that some strategies - e.g. *long only* - deliver significantly negative alpha in “bad” times. According to them other strategies provide high excess returns even then[7].

We suppose that USLSHFs belong to the latter category. Not without reason do they explicitly admit *short* strategies. Thus our hypothesis is:

H4: USLSHFs perform significantly better in recessionary phases as other HFs.

[4] Stafylas & Andrikopoulos (2020), p. 6.
[5] Ibid.
[6] Ibid., p. 8.
[7] Ibid., p. 28.

Contributions of Factors to HF Performance

Impact of single factors raises the question of their contribution to performance in a multivariate context. We take up this question. Our hypothesis is:

H5: As far as factors show significant impact on HF performance on a univariate basis, they will also decisively contribute to HF performance on a multivariate basis.

2.3. Origin of Data

Data sets of our HF analysis have been mainly provided by Institute of Banking and Finance (IBF) at Leibniz University Hannover. IBF HF data cover a total of 979 HFs over the period from February 1990 to August 2020. They contain base data of HFs (e.g. security id, Morningstar category) as well as price data on a monthly basis.

Market data of IBF dataset have been drawn from library of Fama and French[8]. Fama-French data contain a performance stock-market index comprising all firms listed at NYSE, AMEX or NASDAQ. This stock-market index is furtheron referenced as *the market*. Fama-French data represent risk free returns by one-month Treasury bill rate.

US business cycles are officially announced by National Bureau of Economic Research (NBER). The period from 1990 to 2020 has seen 4 complete business cycles as denoted by NBER. There are four growth periods (01/1990–07/1990, 04/1991–03/2001, 12/2001–12/2007 and 07/2009–02/2020) and four recession periods (08/1990–03/1991, 04/2001–11/2001, 01/2008–06/2009 and 03/2020-04/2020).

2.4. Data Set of Our Survey

Fund surveys require sufficient track records of funds for well-founded results. Leaned on our lecture-practice we demand 60 monthly returns for inclusion of HFs into our study. This reduces relevant HFs from a number of 979 to 727. Those 727 funds are composed of 252 USLSHFs and 475 other funds (AOHFs).

Our survey is conducted on the basis of excess monthly returns[9]. A first impression of location and dispersion can be drawn from the following tables:[10]

Table 2.1.: USLSHFs - Monthly Excess Returns

Mean	SD	25%-Quantile	Median	75%-Quantile
0.0064	0.0518	-0.0136	0.0056	0.0264

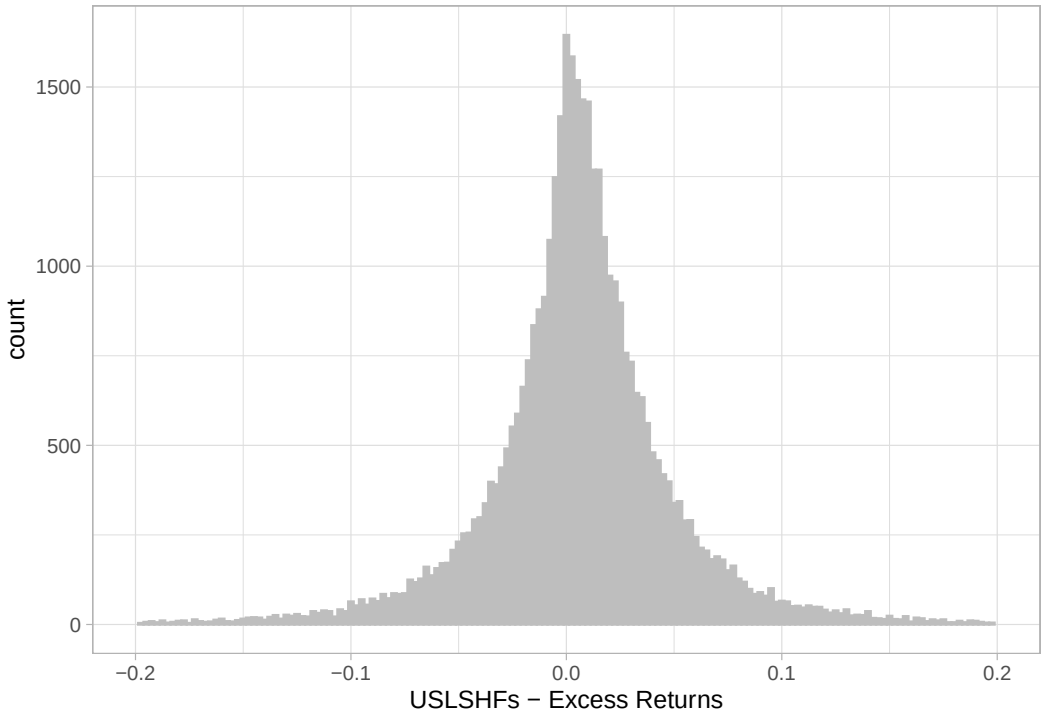
[8] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[9] For a definition see Section 3.1.
[10] Tables do always refer to the whole period 1990 - 2020, if not explicitly mentioned otherwise.

Table 2.2.: AOHFs - Monthly Excess Returns

Mean	SD	25%-Quantile	Median	75%-Quantile
0.0052	0.0559	-0.0119	0.0051	0.0219

USLSHFs display lower volatility. Nonetheless do they on average generate higher excess returns as compared to AOHFs. Furthermore returns of USLSHFs are more straddled - indicated by quartile/quartile distance.

Returns of USLSHFs were plotted as a histogram.



Distribution is rather symmetrical and sharply peaked. It obviously deviates from bell shape of normal distribution. Deviation from normality generally also occurs at single fund level.[11] This diagnosis is of importance for the further course of our study, e.g. testing. Because of non-normality we reverted to non-parametrical tests, namely Mann-Whitney U tests[12].

At this stage we regard HFs at at highly aggregated level, only differentiating between USLSHFs and AOHFs. With huge numbers of return observations over 3 decades we can easily state: USLSHFS are a clearly distinct HF class with special return characteristics - regarded as a collective in a long-term perspective.

[11] See Appendix A.1.
[12] See Hedderich & Sachs (2012), p. 481 - 488.

3. Hedge Fund Performance and Determinants

3.1. Definitions of Performance Measures

Starting point of conventional performance analyses are returns of financial instruments (RETs). Excess returns (RETRFs) designate surpluses of nominal returns over risk-free rates (RFs)[1]. Expected value of RETRF is defined accordingly:

$$E(RETRF) = E(RET - RF) \tag{3.1}$$

The systematic trade-off between returns and risk has early entailed respective performance measures. William Sharpe proposed that we measure the attraction of a portfolio by the ratio of risk premiums to standard deviation of excess returns[2].

Sharpe ratio (SR) is defined as follows:

$$SR = \frac{E(RETRF)}{\sigma(RETRF)} \tag{3.2}$$

SR has generated wide acceptance - and also wide criticism.[3] If returns are normally distributed, performance measures that rely on expected value and standard deviation are adequate measures.[4] This applies to Sharpe ratio. However, HFs often generate non-normal returns. As a consequence alternative performance measures have been proposed.

Standard deviation of excess returns - if applied as the denominator of performance measures - implies another point of criticism: It has early been argued that we should look at deviations from *minimum acceptable return rates (MAR)*, e.g. deviations from risk-free returns, rather than from the sample average.[5]

This point of criticism is adressed by *lower partial-moments performance measures (LPMPM)*. The first LPMPM to be introduced was **Sortino ratio (SOR)**. The denominator of

[1] Pedersen (2015), p. 29.
[2] See Bodie (2014), p. 134.
[3] For a summary see e.g. Eling & Schuhmacher (2007), p. 2633 or Bodie (2014), p. 135. With all other preconditions given, Sharpe ratio is regarded as an adequate performance measure for diversified portfolios, but not for individual assets: See Bodie (2014), p. 135, Eling & Schuhmacher (2007), p. 2633.
[4] Bodie (2014), p. 134.
[5] Ibid., p. 140.

SOR uses only negative deviations from MAR, squares them analogously to variance and takes the square root to obtain a "left-tail standard deviation"[6]. The numerator of SOR is formed by the expected value of excess returns over MAR.

$$SOR = \frac{E(RETRF - MAR)}{\sqrt{\sum_{t=1}^{n} min(RETRF_t - MAR,\, 0)^2/n}} \tag{3.3}$$

As another alternative reference for excess returns over MAR, *value at risk (VaR)* has been proposed. This measure specifies the possible loss of an investment which is not exceeded with a given probability of 1 - α.[7]

Excess return on value at risk (EXC) takes VaR to measure E(RETRF - MAR):

$$EXC = \frac{E(RETRF - MAR)}{VaR} \tag{3.4}$$

With the advent of CAPM[8] respective new measures were introduced to performance analysis. CAPM implies security-specific measures of α and ϵ. They are computed by running a regression of HFs´ RETRFs at time t on MKTRF:

$$RETRF_t = \alpha + \beta * MKTRF_t + \epsilon_t \tag{3.5}$$

In this equation **ALPHA** denotes excess returns of a specific financial instrument over excess returns of market portfolio. ALPHA represents the intersect of $\beta * MKTRF$ with y-axis:

$$ALPHA \;=\; \alpha \;\; = \overline{RETRF} - \beta * \overline{MKTRF} \tag{3.6}$$

In equation (3.5) ϵ describes deviations from the straight defined by $\beta * MKTRF$.

The higher $\sigma(\epsilon)$, the *idiosyncratic risk* of a specific HF, the more do risk-adjusted returns deviate from values that are specified by $\beta * MKTRF$.

Information ratio (IR)[9] condenses ϵ and α to one single performance measure. IR represents risk-adjusted α.

$$IR = \frac{\alpha}{\sigma(\epsilon)} \tag{3.7}$$

Performance Measures as defined above are now to be calculated for our data set. Before that we have, however, to question distribution of returns.

[6] See Bodie (2014), p. 140.
[7] Eling & Schuhmacher (2007), p. 2636.
[8] For an introduction see e.g. Brealey (2011), p. 185 - 212; Bodie (2014), p. 291 - 323
[9] Bodie (2014), p. 274 - 276.

3.2. Ranking of Sample Hedge Funds under Different Performance Measures

In succession of SR later performance measures have i.a. been developed because normal distribution of returns is often questionable. Our calculation of performance measures will thus be preceded by testing RETRF distribution for normality.

As presented in Section 2.4. RETRFs of HFs at a collective level show a symmetrical, sharply-peaked distribution that deviates from the typical bell shape of normal distribution. Deviating shapes were often retrieved by us when plotting RETRF histograms for individual HFs as well. Hereby often heavy tails became apparent. RETRF distributions of individual USLSHFs were thus tested for normality using Shapiro-Wilk test.[10]

Normality hypothesis had to rejected for 207 out of 252 funds ($p < 0.05$)[11].

This demonstrates: As far as HF performance measures are concerned, alternatives to SR are of legitimate interest. This statement in turn raises the question how alternative performance measures compare to SR.

Our computation of performance measures was conducted according to the a.m. formulas.[12] Afterwards the performance of HFs was ranked under each of selected performance measures. Rank correlation was calculcated on the basis of Spearman´s rank correlation coefficient[13]:

Table 3.1.: Rank Correlation Values of Performance Measures (Spearman)

	SR	SOR	EXC	IR	ALPHA	avRETRF
SR	1	0.9385	0.9300	0.8909	0.7143	0.6746
SOR	0.9385	1	0.9801	0.8257	0.7862	0.8249
EXC	0.9300	0.9801	1	0.8200	0.7756	0.8035
IR	0.8909	0.8257	0.8200	1	0.8471	0.5699
ALPHA	0.7143	0.7862	0.7756	0.8471	1	0.7838

avRETRF $\hat{=} \overline{RETRF}$

[10] Hedderich & Sachs (2012), p. 426.

[11] Appendix A.1. documents those 45 HFs, for which normality hypothesis did not have to be rejected.

[12] Minimum acceptable rate of SOR and EXC in equations 3.3 and 3.4 has been defined as the risk-free rate of return. Following Eling & Schuhmacher (2007) we have taken one single value for the whole period under investigation: in the form of the mean of monthly risk-free rates (Eling & Schuhmacher (2007) took the monthly equivalent of a long-term rate). Mean of monthly MKTRFs of our study amounts to 0.22%. SOR in equation 3.3 has been computed on the basis of our own coding. We have reviewed our findings by the respective function of PerformanceAnalytics-package and obtained identical results. VaR was determined according to our lecture practice: as the possible loss which is not exceeded with a probability of 95 %. VaR was also computed by means of our own coding. We didn´t make use of functions included in the PerformanceAnalytics-package, because these could regretfully not be reviewed respectively verified by us.

[13] Hedderich & Sachs (2012), p. 110 - 112.

Distinctly strong rank correlation (> 0.8) is consistently displayed by measures with a comparable structure: SR, SOR, EXC and IR all embody return indicators in their numerator and risk measures in their denominator.

In contrast ALPHA and $\overline{RETRF}$ are both refined to a single dimension of representation (returns). Thus rank correlations to the other performance measures seem to be lower not by chance. ALPHA and $\overline{RETRF}$ corrrelate most distinctly to IR resp. SOR. This seems to be no coincidence either as the latter measures include the former in their numerators.

Comparison to Eling & Schuhmacher (2007)

Eling & Schuhmacher (2007) tested 13 performance measures for correlation, including SR, SOR and EXC[14]. All 13 performance measures exhibit pairwise rank correlation values (Spearman) >= 0.92. Eling & Schuhmacher (2007) state, that "despite significant deviations of hedge fund returns from a normal distribtution, our comparison of Sharpe ratio to the other performance measures results in virtually identical rank ordering across hedge funds."[15]

We compared our own findings to the results of Eling & Schuhmacher (2007).

Table 3.2.: Rank Correlation of Performance Measures - Own Findings in Comparison to Eling and Schuhmacher (in brackets)

	SR	SOR	EXC
SR	1	0.94 (0.99)	0.93 (1.00)
SOR	0.94 (0.99)	1	0.98 (0.98)
EXC	0.93 (1.00)	0.98 (0.98)	1

This comparison shows that our findings are in line with Eling & Schuhmacher (2007).

We considered three other measures not included by Eling & Schuhmacher (2007): IR, ALPHA and $\overline{RETRF}$. Rank correlations of these other measures range below the a.m. values of >= 0.92.[16] These other measures including ALPHA thus obviously comprise information content that differs from SR, SOR and EXC.

3.3. Measures and Data Set of Performance Impact Analysis

With obviously additional information content of ALPHA, it would have been desirable to include this measure and composites of it in our performance impact analysis. This however would have presupposed that ALPHA calculation happens within a conclusive linear model.

[14] IR, ALPHA and $\overline{RETRF}$ are not included in their study.
[15] Eling & Schuhmacher (2007), p. 2632.
[16] See the respective table above.

Valid and significant linear models require several prerequisites. Statistical prerequisites include:[17]

1. **Overall Model Significance**
   Overall model significance can be presumed if not all parameters of the linear model must be assumed to be equal to zero. This requires that respective null hypothesis can be rejected (p-value of F-statistic < 0.05).

2. **Single Coefficient Significance**
   Single coefficient significance can be presumed if both ALPHA and BETA can be assumed as not being equal to zero. This requires that respective null hypothesis can be rejected in both cases (p-value of F-statistic < 0.05).

Besides that distribution requirement of regression residuals has to be fulfilled:

3. **Normal Distribution of Residuals**
   Regression residuals are to be supposed to be distributed normally. This requires that null hypothesis of respective Shapiro-Wilk test does not have to be rejected (p-value of W-statistic > 0.05).

We checked fulfillment of conditions for available 252 USLSHFs. The following table shows all USLSHFs that fullfil a.m. 4 conditions entirely.

Table 3.3.: Linear Models - RETRF vs. MKTRF - Measures

Secid	p-Value F-Statistic	ALPHA	p-Value ALPHA	BETA	p-Value BETA	p-Value W-Statistic
F00000GV1F	0.0242	0.0101	0.0383	0.2120	0.0242	0.7657
F00000GXGO	0.0000	0.0077	0.0012	0.2725	0.0000	0.1767
F00000GXGQ	0.0000	0.0079	0.0163	0.3905	0.0000	0.2640
F00000H58O	0.0000	-0.0049	0.0002	0.6559	0.0000	0.0744
FHUSA04BHO	0.0000	0.0067	0.0010	0.2430	0.0000	0.0628
FHUSA04EVP	0.0000	0.0040	0.0440	0.7345	0.0000	0.1034
FHUSA04EVQ	0.0000	0.0041	0.0433	0.7592	0.0000	0.1627
FHUSA04F1G	0.0000	0.0057	0.0008	0.2402	0.0000	0.0740
FHUSA04F5S	0.0000	0.0123	0.0200	0.8131	0.0000	0.2354
FHUSA04H74	0.0000	0.0083	0.0251	0.5134	0.0000	0.3885
FOUSA05L5Q	0.0000	-0.0060	0.0059	0.6676	0.0000	0.2402
FOUSA05L9N	0.0000	0.0055	0.0010	0.2586	0.0000	0.1417
FOUSA06BOR	0.0000	0.0039	0.0433	0.1958	0.0000	0.2700
FOUSA06C2K	0.0000	0.0043	0.0329	0.3191	0.0000	0.7559
FOUSA06H2K	0.0000	-0.0042	0.0441	0.6899	0.0000	0.1648
FOUSA06M1H	0.0066	0.0050	0.0452	0.1290	0.0066	0.7850
FOUSA06QO9	0.0000	0.0070	0.0426	0.4564	0.0000	0.9978
FOUSA06U6U	0.0000	0.0037	0.0140	0.9655	0.0000	0.2575
FOUSA06UT1	0.0000	0.0081	0.0164	0.4151	0.0000	0.7867
FOUSA08ZW5	0.0144	0.0066	0.0239	0.1824	0.0144	0.7368
FOUSA0909O	0.0000	0.0053	0.0049	0.4673	0.0000	0.3055

[17] See Hedderich & Sachs (2012), p. 693 - 697.

Fulfillment of all 4 conditions is given for 21 USLSHFs (8 %).
231 USLSHFs (92 %) don´t fulfill conditions at least partly.

Fulfillment percentages of *single* conditions are:
- Overall model significance given: 88 % of USLSHFs
- Significance of ALPHA given (is not zero): 38 % of USLSHFs
- Significance of BETA given (is not zero): 88 % of USLSHFs
- Normal distribution of residuals can be assumed: 23 % of USLSHFs

Synopsis of fulfillment percentages indicates that overall model significance is essentially brought about by BETA. ALPHAs can be assumed to be significantly different from zero only in a minority of cases. Distribution of residuals can be taken to follow normality with even fewer USLSHFs.

ALPHA seems - as a consequence - not to be a favorable performance measure for our impact survey. We thus looked for alternatives. To keep our performance impact analysis within fewest possible assumptions and requirements, we decided to take the most simple measure as a performance indicator: excess returns (RETRFs). This decisison has been contributed to by recommended literature: Jones (2007) also makes use of RETRFs for her performance impact study.

Data Set of Our Performance Impact Study

The goal of our performance impact study requires review of our data set. To extend our study to individual funds we need sufficient observation numbers in all cases at issue.

We started our study with 252 USLSHFs. 91 of these have at least 25 RETRF observations in all 4 settings: young age, old age, expansion and recession. Increasing the required number of observations would have reduced the number of HFs to be included drastically - especially resulting from relatively smaller scope of recessionary phases and subsequently fewer observations for these periods. We thus determined at least 25 RETRF observations in all 4 settings as a precondition for USLSHF inclusion into our performance impact study.

AOHFs were selected accordingly with at least 25 observations in all 4 settings. This leaves 169 of those 475 AOHFs that we started with.

3.4. Fund Age and Performance

As pointed out in Chapter 2 previous research largely revealed significant impacts of HF age on performance. For her study Jones (2007) classified HFs into age classes.[18] RETRFs of a specific HF were assigned to the relevant age class if and as long the single HF belonged to this class.

[18] See Jones (2007), p. 346. Jones´ study is oriented on 3 age classes.

RETRFs of our sample were classified into 2 age classes. This results in 2 portfolios:
- *young* comprising observations of funds <= 3 years after inception and
- *old* (> 3 years).

After April 2004 no data of young HFs are available. From 1993-07 to 2004-03 at least 3 observations in every group are existent. Comparison was thus confined to this period.

Table 3.4.: RETRFs of Fund Types by Age Class - 1993 - 2004

FundTtype	Age Class	Mean	SD	25%-Quantile	Median	75%-Quantile
AOHF	old	0.0084	0.0524	-0.0108	0.0061	0.0256
AOHF	young	0.0126	0.1089	-0.0119	0.0067	0.0305
USLSHF	old	0.0094	0.0606	-0.0156	0.0071	0.0319
USLSHF	young	0.0127	0.0580	-0.0128	0.0098	0.0372

The table shows expected mean RETRF differences between age classes. Difference in central tendency is highly significant for USLSHFs ($p < 0.001$, one-sided). Age groups of AOHFs don´t differ significantly ($p = 0.054$).[19]

USLSHFs exhibit performance advantages over AOHFs, remarkably only in *old* age class. SD of young USLSHFs is noticeably lower as of AOHFs.

Single Hedge Funds, Fund Age and Performance

A.m. findings *young vs. old* were reviewed at level of single USLSHFs.

RETRFs of single USLSHFs were grouped by age stages: young (<= 3 years) vs. old (> 3 years). RETRF means of age stages and difference between them were calculated. Differences of age stage RETRF distributions were covered by MWU. Alternative hypothesis corresponds to expected direction: *young* greater as *old*.

The table below shows all USLSHFs that have positive RETRF mean difference and highly significant differences of central tendency ($p < 0.01$, one-sided).

[19] See Appendix A.2.

Table 3.5.: RETRFs of USLSFs at Different Age Stages - p-Value of MWU < 0.01 (One-Sided)

Secid	Young	Old	Diff.	p-Value MWU
FHUSA04DV8	0.0465	0.0015	0.0449	0.0083
FHUSA04F3D	0.0392	0.0026	0.0366	0.0013
FHUSA04GWM	0.0409	0.0077	0.0332	0.0007
FHUSA04ASO	0.0301	0.0010	0.0291	0.0023
FHUSA04HW2	0.0350	0.0071	0.0279	0.0080
FHUSA04COQ	0.0328	0.0050	0.0278	0.0024
FHUSA04B1T	0.0298	0.0065	0.0233	0.0097
FHUSA04D0P	0.0243	0.0041	0.0202	0.0007
FHUSA04HO6	0.0207	0.0017	0.0190	0.0056
FOUSA06LNX	0.0264	0.0084	0.0181	0.0001
FHUSA04AS4	0.0223	0.0061	0.0162	0.0000
FHUSA04GHI	0.0186	0.0029	0.0157	0.0004
FHUSA04GPO	0.0176	0.0056	0.0120	0.0075
FHUSA04DDQ	0.0083	0.0006	0.0077	0.0029
FHUSA04CZ3	0.0106	0.0043	0.0063	0.0000
FHUSA04ASX	0.0060	0.0010	0.0050	0.0028

We can recognize: 16 USLSHFs show highly significant distribution difference: *young* greater as *old* ($p < 0.01$). Allover about 4/5 of funds follow our direction hypothesis.

Our analysis of USLSHFs at their young vs. old stages can be summarized as follows:

Table 3.6.: Single USLSHFs at Young vs. Old Age Stage - RETRF Mean Difference and p-Value of MWU (One-Sided)

Young/Old Difference / p-Value of MWU	> 0 / **	> 0 / *	> 0	< 0	Total
n USLSHFs	16	12	47	16	91

Conclusion: Fund Age and Performance

It can be summarized that young USLSHFs do significantly perform better as old ones at portfolio level. Advantages of young funds seem not to be restricted to USLSHFs: Regarded collectively AOHFs show the same pattern. Young USLSHFs, however, feature lower SD when compared to young AOHFs on a long-term portfolio basis.

Concerning age stage impact, *single* USLSHFs to a large extent tend in the same direction: about 4/5 do perform better in their younger as in their older years, more than 1/3 of them significantly.

Jones (2007) comes to the conclusion that investors who wish to maximize return should start their search by looking for younger funds. Our findings support her recommendation.

3.5. Cycle Phases and Performance

As pointed out in Chapter 2, previous research revealed significant impact of cycle phase on HF performance. Our study reviews this finding - at portfolio as well as single funds level.

The period from January 1990 to August 2020 has seen 4 complete business cycles as denoted by NBER. There are four growth periods (01/1990–07/1990, 04/1991–03/2001, 12/2001–12/2007 and 07/2009–02/2020) and four recession periods (08/1990–03/1991, 04/2001–11/2001, 01/2008–06/2009 and 03/2020-04/2020).

Performance of HFs - when differentiating between expansionary and recessionary phases - can be summarized as follows:

Table 3.7.: Fundtype by Cycle Phase - Monthly Excess Returns 1993 - 2004

Fund Type	Cycle Phase	Mean	SD	25%-Quantile	Median	75%-Quantile
AOHF	expansion	0.0077	0.0606	-0.0102	0.0059	0.0229
AOHF	recession	-0.0062	0.0652	-0.0286	-0.0012	0.0186
USLSHF	expansion	0.0082	0.0508	-0.0131	0.0063	0.0282
USLSHF	recession	-0.0048	0.0718	-0.0313	-0.0008	0.0260

The table shows expected cycle phase differences of USLSHF vs. AOHF RETRFs. Difference in central tendency is highly significant ($p < 0.001$, one-sided). This relates to USLSHFs and AOHFs as well.[20]

Regarded collectively USLSHFs perform better as AOHFs during expansionary phases as well as during recessions. In latter periods they do however not even generate mean RETRFs that are positive. This finding is striking: It questions legitimacy of an explicit strategic *short* orientation of USLSHFs. And it raises the question how individual funds do in recessionary phases.

Single Hedge Funds, Cycle Phases and Performance

A.m. findings were reviewed at level of single funds. RETRFs of single USLSHFs were grouped by cycle phase: expansion vs. recession. RETRF means during both cycle phases were calculated. Cycle phase RETRF distributions were compared by MWU (one-sided). The following table is arranged in order of descending RETRFs > 0 during recession.

[20] See Appendix A.2.

Table 3.8.: RETRFs of USLSFs vs. Cycle Phase

Secid	Expansion	Recession	Diff.	p-Value MWU
FOUSA08XA3	0.0062	0.0326	-0.0263	0.0233
FOUSA05L2U	0.0050	0.0191	-0.0141	0.0378
FHUSA04HW2	0.0140	0.0150	-0.0010	0.3370
FHUSA04HGQ	0.0075	0.0132	-0.0056	0.7337
F00000H5GL	0.0061	0.0130	-0.0070	0.8482
F00000IRR1	0.0067	0.0126	-0.0059	0.5183
FHUSA04B0U	0.0096	0.0124	-0.0028	0.4935
FHUSA04CB2	0.0075	0.0110	-0.0035	0.6131
FHUSA04DV8	0.0138	0.0101	0.0036	0.6488
FOUSA085WN	0.0065	0.0099	-0.0034	0.7651
FHUSA04F1G	0.0061	0.0080	-0.0019	0.5598
FHUSA04F8Q	0.0051	0.0077	-0.0026	0.8611
FOUSA08YJY	0.0034	0.0074	-0.0041	0.2208
FHUSA04APF	0.0021	0.0069	-0.0048	0.4897
FHUSA04HK2	0.0059	0.0067	-0.0008	0.8405
FHUSA04COQ	0.0093	0.0041	0.0052	0.9291
FHUSA04D3O	0.0001	0.0041	-0.0040	0.6647
FOUSA06IRF	0.0134	0.0035	0.0099	0.6004
FHUSA04D37	0.0059	0.0035	0.0024	0.8694
FOUSA05KX2	0.0036	0.0029	0.0006	0.8156
FHUSA04GHI	0.0059	0.0025	0.0034	0.7506
FOUSA05L6M	0.0118	0.0024	0.0094	0.4970
FHUSA04C3N	0.0147	0.0023	0.0125	0.2222
FHUSA04BHB	0.0096	0.0018	0.0078	0.5023
FOUSA06JZC	0.0063	0.0001	0.0062	0.9057
FHUSA04BHO	0.0088	0.0000	0.0087	0.0566

26 USLSHFs managed to deliver positive RETRFs during recession. This fact is all the more remarkable in comparison to AOHFs: No single of all 169 AOHFs delivered positive mean RETRFs during recessionary phases.

Synopsis of a.m. 26 USLSHFs reveals some interesting patterns. The first 2 funds e.g. generate noticeably and significantly higher RETRFs during recession as compared to expansion. We suppose that these funds pursued distinctive and enduring *short* strategies. They had to forego possibly higher returns during expansion. On the other hand were they prepared at any time for slumps and profited outstandingly when they occurred. But this strategic abstract naturally is pure speculation.

We hereby finish our univariate analyses. They have all in all demonstrated a manifold picture of performance impact. The question is how apparent impact of single factors materializes in a multivariate context.

3.6. Hedge Fund Performance in a Multivariate Context

Coefficient of determination, denoted R^2, is the proportion of variation in the dependent variable that is explained by independent variables. Because increases in the number of independent variables necessarily increase the value of R^2, coefficient of determination alone cannot be used as a meaningful comparison of models with numerous variables.

Our study doesn´t intend any differentiated findings concerning multivariate performance impact. Our R^2 analysis is only directed to a first impression of relative importance. We include MKTRF (M), fund age (A) and cycle phase (C), respectively combinations of these factors, into our analysis.

R^2 for different constellations of independent variables looks as follows:

Table 3.9.: RETRF vs. MKTRF (M), Fund Age (A), Cycle Phase (C) 1993 - 2004: Coefficients of Determination (R^2)

M+A+C	M+A	M+C	M	A
0.1794	0.1792	0.1737	0.1735	0.0031

With all 3 variables included 17.9 % of RETRF variance can be explained. MKTRF alone explains 17.4 % (!) of this variance.

Inclusion of age and cycle phase accounts for only fractions of a single percentage point of RETRF variance. Contribution of cycle phase is most insignificant.

On a univariate basis we have seen that HF age and - to a lesser extent - cycle phase do matter for USLSHF performance. This difference is however negligible as compared to market impact when analyzed by multivariate methods.

4. Summary

Concerning excess returns, US long/short hedge funds are a clearly distinct financial instrument - when compared to all other hedge funds (taken as a whole) over long periods. This statement is - albeit rather general - all the more remarkable as volatility of long/short hedge funds ranges below all other hedge funds (regarded as a collective over long periods).

Performance measures must recognize empirical distributions of their components. Regretfully certain measures, especially ALPHA, had to be dropped for our analysis because the underlying linear model exhibited shortcomings.

Our analysis shows - in line with prior research - that several performance measures have high rank correlation. Correlation is most pronounced with measures that are similarly constructed (not surprisingly).

Performance impact analysis shows up a number of interesting facts:

- Fund age is a decisive factor concerning hedge fund performance. This relates to US long/short equity funds as well as to all other regarded as a collective. Our findings are in line with previous surveys.
- Cycle phase is a relevant factor as well. During recessions US long/short equity funds do not even manage to deliver positive excess returns (regarded as a collective). Single US long/short equity funds however achieve positive excess returns to a considerable extent during recessions.
- Both factors, fund age and cycle phase, are thus relevant for fund performance. They are, however, rather meaningless on a multivariate basis compared to the market.

As we stated in our introduction: Capital market research is manifold concerning focus, perspectives, objectives and and and … As a consequence our study has brought many interesting insights - and possible starting points of future research.

A. Supplementary Analyses

A.1. Excess Returns of Single USLSHFs - Shapiro-Wilk Test

RETRFs are often not distributed normally. Should this be the case, mean and standard deviation are not sufficient for analysis of RETRFs. As a consequence alternative performance measures were proposed.

We tested RETRFs of USLSHFs for normal distribution.

Table A.1.: RETRFs of USLSHFs - Shapiro-Wilk Test - p-Value descending (> 0.05)

HF Name	p-Value
Grunion Fund, LP	0.9782
Connective Capital I LP	0.8767
The Sosnowski Fund, LP	0.8247
Algert U.S. Large Cap Active Extension	0.7939
Perceptive Life Sciences Fund	0.7728
Act II Partners LP	0.7353
Loomis Sayles Consumer Discr LP	0.6710
Connective Capital II QP	0.6411
Elazar Capital Management LLC	0.6318
CCI Healthcare Partners LP	0.6014
Advisory Research Energy Fund	0.5872
SC Opportunity Fund, L.P.	0.5357
EMA GARP Fund, LP	0.4929
CastleRock Partners LP	0.4723
Westerly Partners L.P.	0.4277
CBRE Clarion US LP	0.3833
FJH Associates II LP	0.3807
Krema Absolute Return Fund LLC	0.3788
The Anthem Opportunity Fund, LP	0.3545
Crawford Capital Partners, LP	0.3321
Veredus Partners LP	0.3216
Point Defiance Microcap Fund, L.P.	0.3068
The Decathlon Fund	0.3003
Bluefin Investors, LP	0.2772
Keen Vision Fund L.P.	0.2406

Table A.2.: RETRFs of USLSHFs - Shapiro-Wilk Test ... (Continued)

HF Name	p-Value
New Castle Fallen Angels	0.2345
Loch Capital I LP	0.2295
Peconic Partners International Fund Ltd	0.2031
DRMM Total Return Fund I LP	0.1859
Precept Market Neutral Fund LP	0.1670
Alkeon Growth Partners LP	0.1666
Loch Capital II LP	0.1394
P.A.W. Retail & Consumer Fund, L.P.	0.1076
T2 Accredited	0.1065
Keen MicroCap Value Fund L.P.	0.1055
Elazar Global Ltd	0.1049
GLS 193 Fund Ltd	0.0977
Ruth Asset Management I , LP	0.0931
Peconic Triumph Fund LP	0.0776
Seamark	0.0642

For more than 4/5 of USLSHFs normality hypothesis had to be rejected ($p < 0.05$).

To document this, the above table includes 45 funds with p-value > 0.5 (i.e. HFs for which hypothesis of normal distribution was not rejected).

As a consequence we resorted to non-parametrical tests.

A.2. Performance Impact of Age and Cycle Phase - Significance Tests at a Collective Level

Prior research has shown that fund age and cycle phase have significant impact on performance.

We reviewed these findings in the main part of our study.

We covered our results by MWUs, which are documented here.

Young vs. Old

Table A.3.: Young vs. Old HFs 1993 - 2004 - Mann-Whitney U Test (One-Sided)

fundtype	statistic	p.value	method	alternative
AOHF	19856758	0.9460	Wilcoxon rank sum test	greater
USLSHF	6211340	0.9998	Wilcoxon rank sum test	greater

That means: H_0 does not have to be rejected. As a consequence old HFs - regarded collectively - can be assumed to generate lower performance as young ones.

Expansion vs. Recession

Table A.4.: HFs in Expansion vs. Recession 1993 - 2004 - Mann-Whitney U Test (One-Sided)

fundtype	statistic	p.value	method	alternative
AOHF	79050543	1	Wilcoxon rank sum test	less
USLSHF	22812462	1	Wilcoxon rank sum test	less

That means: H_0 does not have to be rejected. As a consequence HFs during expansion - regarded collectively - can be assumed to generate stronger performance as funds during recessions.

Literature

Bodie, Z. et al. (2014). *Investments* (Tenth edition). New York: McGraw-Hill Education.

Brealey, R. A. et al. (2011). *Principles of Corporate Finance* (Concise ed., 2nd ed). New York: McGraw-Hill Irwin.

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance*, *31*(9), 2632–2647.

Hedderich, J., & Sachs, L. (2012). *Angewandte Statistik: Methodensammlung mit R* (14., überarb. u. erg. Aufl). Berlin Heidelberg: Springer.

Jones, M. (2007). Examination of fund age and size and its impact on hedge fund performance. *Derivatives Use, Trading & Regulation*, *12*(4), 342–350.

Pedersen, L. H. (2015). *Efficiently inefficient: How smart money invests and market prices are determined* (First paperback printing). Princeton, New Jersey, Oxford: Princeton University Press.

Stafylas, D., & Andrikopoulos, A. (2020). Determinants of hedge fund performance during 'good' and 'bad' economic periods. *Research in International Business and Finance*, *52*, 1–45. Page count from an online version of the article, available under Creative Commons License: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2692005 (downloaded Jan. 21, 2022).

Statutory Declaration

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Ort, Datum	Dipl.-Kfm. Hans-Christian Schneider