Can we select outperforming hedge funds? A set-identification approach based on efficient pairwise comparisons

Auteur : Muller, Guillaume
Promoteur(s) : Hambuckers, Julien
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en ingénieur de gestion, à finalité spécialisée en Financial Engineering
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/15301

CAN WE SELECT OUTPERFORMING HEDGE FUNDS? A SET-IDENTIFICATION APPROACH BASED ON PAIRWISE COMPARISONS

Jury:
Supervisor:
Julien HAMBUCKERS
Reader:
Marie LAMBERT

Master thesis by
Guillaume MULLER
For a master's degree in Business Engineering with specialization in Financial Engineering
Academic year 2021/2022

Ecole de Gestion de l'Université de Liège

Acknowledgements

More than a year ago, I embarked on a more than tumultuous and exciting adventure which is this end-of-study thesis. Indeed, many challenges and difficulties have arisen ahead of me, but that only makes the pride and joy of this accomplishment greater.

The completion of this thesis would not have been possible without the support of many people. I would like to particularly thank my thesis supervisor Prof. Julien Hambuckers for his invaluable assistance and guidance throughout this journey, as well as for his accessibility. I have benefited greatly from his knowledge and pieces of advise.

In addition, I wanted to thank Prof. Marie Lambert for dedicating her precious time to reading and analysing this work. I very much look forward to receiving your feedback and opinion, and to discussing this thesis topic more in details during my oral defence

Finally, I wanted to thank my family and friends for their unfailing support, patience and assistance during this last year, without which it would have been considerably more complicated to achieve this project.

Contents

Acknowledgements . . . 2
1 Introduction . . . 6
2 Literature Review . . . 9
2.1 General performance of Hedge Funds . . . 9
2.2 Performance based on Hedge Funds features . . . 9
2.3 Choice of outperforming Hedge Funds . . . 10
3 Methodology . . . 12
3.1 Alpha and Predictive alpha computations . . . 12
3.2 HFCS Algorithm . . . 14
3.3 HFCS Interpretation . . . 16
3.4 Operational requirements . . . 16
4 Data Presentation . . . 18
5 Results of the analysis . . . 20
5.1 HFCS example . . . 20
5.2 Results of the complete analysis . . . 21
5.2.1 HFCS Performance . . . 22
5.2.2 Performance of alpha and predictive alpha sorted portfolios . . . 24
5.2.3 Performance of alternative approaches . . . 26
6 Conclusions and Discussions . . . 28

A Alpha and predictive alpha distributions: Kalman filter **29**

B HFCS Example : Predictive alphas **30**

C HFCS Example : Predictive alphas differences **31**

D Example: *'The Vilas Fund LP'* **32**

E Returns of HFCS Portfolios **33**

F Performance of single-fund portfolios **34**

G Alpha-sorted portfolios: Number of funds (Rolling window) **35**

H Alpha-sorted portfolios: Number of funds (Kalman Filter) **36**

Bibliography **39**

1 Introduction

The choice of the nature of our investments has always been a challenge, for investors obviously, but also by academics. Choosing an investment fund based on historical data alone is often misleading, which is why it is interesting and meaningful to look at other investment techniques. Hedge funds, by their mysterious and sometimes more complicated nature, make this issue even more challenging. Since 1990, hedge funds have experienced significant growth and their popularity has largely increased. As you can notice on Figure 1, hedge funds' Assets under Management (AuM) jumped from approximately USD 300 million in 2000 to more than USD 4.5 billion in Q3 2021. These assets are mainly concentrated in the United States, representing more than 80 percent of the industry's AuM in 2019 (See Figure 2). The number of hedge funds has also undergone a strong increase, accounting for approximately 9,500 in Q3 2021[1]. This large development has enabled this industry to become the second largest alternative asset class after the Private Equity industry. In order to have a better understanding and a better scope of what a hedge fund is and represents, let us start with an overview of the definition and particularities of the latter.

We can define a hedge fund as an "actively managed, pooled investment vehicle that is open to a limited group of investors and whose performance is measured in absolute return units" (Connor and Woo, 2004). Hedge funds managers seek, by their use of various sophisticated investment tools such as derivatives or leverage, to generate absolute returns. Indeed, thanks to the light regulation they are subject to, these funds have the possibility to invest in such instruments without having to face strict legal requirements. This last point represents the first main difference between hedge funds and traditional mutual funds. As Mitra (2009) stated, hedge funds also target mainly institutional or high net worth individuals, while mutual funds are accessible to the general public. Moreover, hedge and mutual funds' returns differ considerably and have a low correlation, which is appealing to investors.

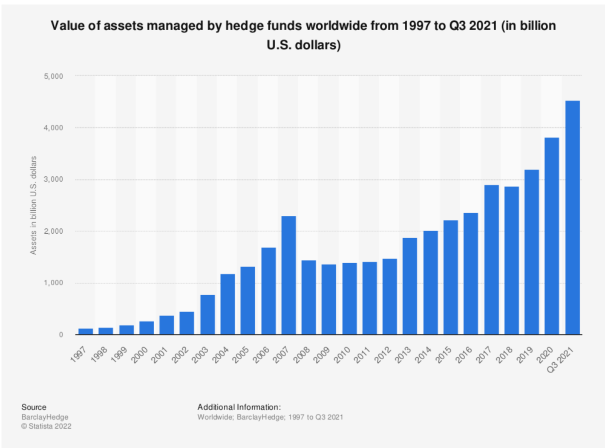

Figure 1: HF AuM growth from 1997 to 2021 (Q3)

Source: *Statista* & *BarclayHedge*

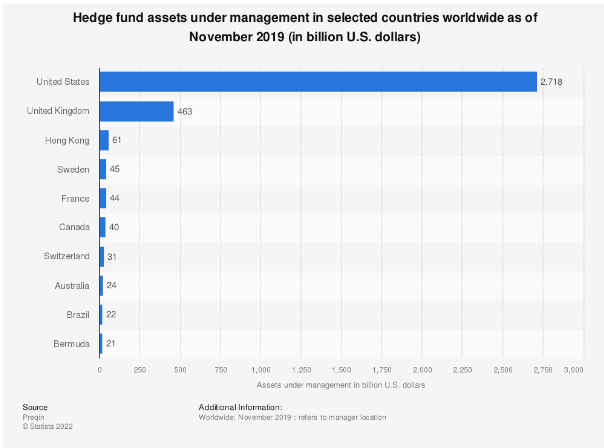

Figure 2: HF AuM by country in 2019

Source: *Statista* & *Preqin*

Regulations represent an important factor concerning hedge funds' performance as they are subject to less legal requirements on their investment methods. Despite their impressive growth in the last decades, hedge funds face few regulatory requirements, and it was especially the case prior to the 2008 crisis. Indeed, as explained by Helleiner and Pagliari (2009), the legal framework of hedge funds before 2008 consisted of indirect regulation through the control of the credits granted to them and consisted of self-regulations mechanisms led by private equity entities. However, the situation changed from 2009 onwards as G20 leaders decided to implement stricter regulations on hedge funds, by requiring the disclosure of information such as leverage. However, these newer requirements are not as stringent as the legal framework of traditional mutual funds, as less transparency and restrictions regarding the funds' investment policies are required.

[1]Data according to the quarterly "Private Funds Statistics Q3 2021" of the SEC

Hedge funds have many characteristics that differentiate them from traditional funds. Firstly, they have a particular fee structure. Indeed, hedge fund managers are remunerated by two types of fees. The first is a management fee, which is usually a percentage of the fund's assets, and the second is called the performance-based fee, which is a percentage of the profit generated by the fund. This fee structure is often called the "2/20" structure as the management fee accounts usually for 2 % while the performance-based fee is generally 20%. Hedge funds managers therefore have a considerable incentive to generate high absolute returns, no matter the current market conditions. It has been proven that higher performance fees have a positive relationship with the fund's performance, as it is indicated further in this report. Moreover, performance fees are often charged only if some conditions are met. For example, a high hurdle rate or water mark could be set up in order to further boost profits and performance.

Hedge funds follow certain main investment strategies and each fund usually focuses on a single strategy. The main hedge fund styles are the followings, as indicated by Connor and Lasarte (2003): Long/Short (or Equity Hedge), Event Driven, Tactical Trading and Relative Value strategies. L/S hedge funds mainly invest, as their name suggest, in equity positions and their derivatives. These funds have the ability to reduce risk by taking short positions on different assets (i.e. short selling). Event Driven funds benefit from events such as mergers or bankruptcies by exploiting pricing inefficiencies. This strategy focuses on distressed securities or merger arbitrage. Tactical Trading strategies, which encompass Macro and Commodity trading strategies, generate returns by speculating on macroeconomic variables swings, such as interest rate, inflation or political stability. Finally, Relative value funds use Arbitrage to produce returns, which is the bet on securities wrongly priced. Naturally, latter strategies encompass multiple sub-categories which won't be discussed here. The net returns and proportion of these strategies can be found in Figure 3 and 4 respectively. As you can notice, there has been a significant net returns decrease from 2020 to 2021 impacting the whole hedge fund industry.

Figure 3: Net returns by HF Strategy (Q420-Q121)

Source: *Statista* & *Preqin*

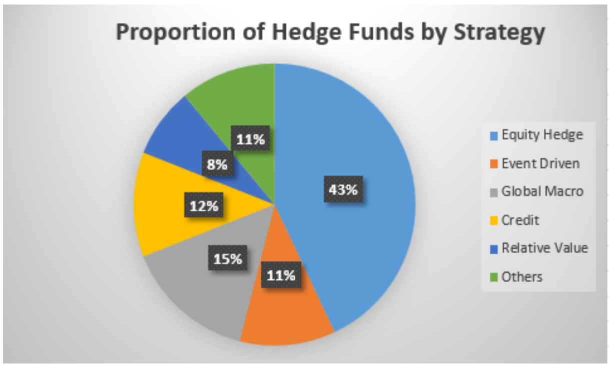

Figure 4: Proportion of hedge funds by Strategy

Source: *Preqin*

Although they generate high absolute returns, hedge funds also face multiple types of risks due to their intense use of sophisticated investment tools such as derivatives or leverage. Naturally, the level of risk incurred by the hedge fund depends heavily on the strategy it focuses on. Hedge funds combine risks from both the liquid financial world and the alternative (and illiquid) fund industry, depending on the strategy used. First, market risk naturally affects multiple strategies, such as the Tactical Trading and Equity strategies. Then, some hedge funds have limited market liquidity. This is especially the case for funds targeting illiquid assets such as mispriced corporations or distressed assets. In order to deal with such an issue, hedge funds can set up restrictions for their investors and therefore improve their liquidity levels. For instance, lock-up periods are time periods during which investors do not have the possibility to redeem their shares. The latter are set up following the investor's contribution. Gates can also be implemented in order to limit the amount of shares redeemed. Finally, redemption periods are usually not daily but rather monthly of quarterly in order to cope with these liquidity issues.

As a last topic, hedge funds are known to encounter data biases within their returns data. These biases can strongly disrupt analyses based on hedge fund returns from specific databases. Fung and Hsieh (2000) presented some of these biases. Survivorship Bias is the fact that the databases only contain funds that are still active and operating and therefore do not include any bankrupt of liquidated funds. Instant History Bias represents the fact that funds' historical returns are back filled when they are added to a database. Funds' managers usually decide to join a database only when their results become historically positive. Finally, Selection Bias consists of the hedge funds managers' choice to disclose their results in a database or not. We can easily assume that only performing hedge funds will make that decision, therefore distorting the database accuracy. These different biases impact naturally any studies based on hedge funds returns. In our analysis, Instant History and Selection Bias can hardly be avoided. However, our study will be based on both live and obsolete funds, which fix the Survivorship Bias.

This introduction enables us to have a better understanding of the hedge fund industry, its various and sophisticated features and trends. As we will see in the next section, almost all of these aspects have an effect on the funds' performances. As it is complex for investors to analyse and compare these funds individually, it is essential to find methods and studies to determine which hedge funds outperform the competitors in order to achieve high long-term absolute returns. In this article, we will present a method based on multiple pairwise comparisons that will allow investors to choose the best performing hedge funds, whose results are not based on luck but rather on skill.

2 Literature Review

In this section, we will discuss several studies concerning the performance of hedge funds. This part will first address the short and long term persistence of hedge funds performance. Then, latter performance will be assessed by associating it to hedge funds' main characteristics. Afterwards, a brief review of the proposed hedge funds' factor models will be presented. Finally, we will review several quantitative studies carried out in order to find and pick outperforming hedge and mutual funds in the market.

2.1 General performance of Hedge Funds

First of all, hedge funds' performance has been proven to have both short-term and long-term persistence, as summarized by Stafylas et al.(2016). Indeed, Agarwal and Naik (2000a, 2000b) indicated that hedge funds present substantial performance persistence on a quarterly horizon for all strategies. Harri and Brorsen (2004) reached approximately the same result, discovering a short-term performance persistence for all styles besides short sales. Bares et al. (2003), as well as Eling (2009) found that there was indeed a performance persistence up to two and six months respectively. The similarity of these findings is that this performance seems to weaken as the time horizon gets longer. However, Ammann et al. (2013) showed that hedge funds, by using the probit regression method and by using a two-way sorted portfolio, presented alpha persistence up to three years. Moreover, Jagannathan et al. (2010) proved, by implementing a method measuring hedge funds' performance and taking into account a constructed peer group, that the best performing funds present performance persistence up to three years. More recently, Gonzalez et al. (2016) examined the persistence of hedge funds performance and indicated that top quantile hedge funds based on their alphas and sharpe ratios outperformed the competition. However, this performance persists up to only 12 months and is partly explained by luck. Nevertheless, we can conclude from these studies that hedge funds experience significant performance which persist through time. Furthermore, the results of Ackermann et al. (1999) on a sample of hedge funds data existing from 1988 to 1995 indicated that hedge funds consistently had higher performance than mutual funds. This result is confirmed by Liang (1999) who analysed approximately the same time-frame. This outcome was mainly explained by the hedge funds' use of more flexible strategies and by their strong managerial incentives. Finally, Bali et al. (2013) showed that a couple of hedge fund strategies (i.e. Long/Short equity hedge and emerging markets) outperformed the US equity market and that some others (i.e. Long/Short equity hedge, multi strategy, managed futures and global macro) had a better performance than the US treasury market.

2.2 Performance based on Hedge Funds features

Now that we have discussed the performance persistence of hedge funds in general, let us decompose this performance with regards to their main features. The first characteristic which might have a relationship with the performance of a hedge fund is its Asset under Management, henceforth called its size. Opinions differ among the various authors who tackled this matter. Agarwal et al.(2004) showed that larger funds, which present larger money inflows, have lower future performance. Harri and Brorsen (2004) and Joenvaara et al. (2012) also found a negative relationship between the performance of hedge funds and their size. Finally, Teo (2009) showed that small hedge funds dominated largers ones by 3.65$\%$ per year taking risk adjustments into account. However, other studies indicate that larger hedge funds outperform smaller entities. Amenc and Martellini (2003) discovered that small funds' mean alphas were consistently outperformed by larger funds. Getmansky (2004) indicated that there is a positive correlation between performance and previous size. Finally, Ammann and Moerth (2005) showed a negative relationship between size and performance, yet they also indicated that very small funds could not compete with medium and large ones.

The hedge funds' age is also a factor that influences the performance and there is a clear tendency favoring younger hedge funds. Meredith (2007) showed results indicating that young and small funds bring maximized returns to investors. Frumkin and Vandegrift (2009) indicated that age and performance have a negative relationship as well.

Finally, the performance fees charged by the fund, as well as the lock up period and the fund's domicile have been proven to have a effect on the fund's performance. Most studies show that higher performance fees indicate higher performance. Indeed, Amenc and Martellini (2003), as well as Joenvaara et al. (2012) proved that high incentives funds' alphas surpassed that of low incentives funds. Liang (1999) indicated that hedge funds implementing high watermarks performed better. A watermark is simply a provision ensuring that hedge funds' managers do not have the possibility to receive incentive fees before having returns above the hurdle rate. Furthermore, lockup periods also have an impact on funds' performance. Aragon (2007) proved that hedge funds implementing lockup periods presented outperforming performance. Agarwal and al. (2009) indicated that hedge funds with greater manager incentives (e.g. higher fees, managerial ownership and watermarks) outperformed. Finally, onshore funds seem to outperform offshore funds, as shown by Joenvaara et al. (2012).

These multiple findings show that the hedge fund industry has undergone short and long-term performance and will therefore remain a considerable investment opportunity in the future. As a consequence, it is important and insightful to develop methods in order to pick the best performing hedge funds in the industry. Moreover, most studies agree on the fact that size and age have a negative relationship with hedge funds' performances, while high performance fees and the implementation of high lockup periods and watermarks influence positively the funds' performances. These results show that hedge funds' performances are strongly impacted by many features characterizing them. Picking funds with regards of these features is a complex and time-demanding task. Procedures enabling for the choice of outperforming hedge funds is therefore a suitable solution.

2.3 Choice of outperforming Hedge Funds

Having explained the features that can influence hedge funds' performance, it is now interesting to look at the different analytical methods to find the best performing hedge and mutual funds, without being based solely on latter variables. First of all, performances of hedge funds, as it is for traditional funds, are usually assessed thanks to factor models in order to retrieve the fund's alpha and to figure out the funds' risk exposures to certain variables. This means that hedge funds' returns are generated through risk factors and not solely thanks to market inefficiencies. The most popular and frequently used model is certainly the seven-factor model of Fung and Hsieh (1997,2004). Liang (1999) used eight different factors to consider hedge funds' returns as an extension of the Fung and Hsieh model. Brealey and Kaplanis (2001) even considered more than thirty factors for their analysis. However, multiple studies still rely on traditional single or multiple factor models such as the CAPM or Fama-French models in order to explain hedge funds' returns. Determining the hedge funds' alphas is however not sufficient in order to determine which fund outperforms its competitors. That is why we will introduce some analytical and quantitative methods used in order to find outperforming mutual and hedge funds.

Kosowski et al. (2006), by using a bootstrap statistical technique, tried to identify US mutual funds which can outperform based on a risk-adjusted basis. Their conclusions indicated that the best and worst funds' performances are not solely explained by luck. Moreover, Barras et al. (2010) assessed the ability of mutual funds to genuinely outperform their benchmarks. For this matter, they ranked funds in three categories based on their ability to choose correctly their holdings, namely "Unskilled", "Zero-alpha" and "Skilled" funds. This is carried out by applying threshold on the cross-sectional alpha t-statistic distribution. The implemented procedure accounts for luck by identifying false discoveries within the t-test distribution. Their conclusion indicates that the proportion of skilled funds has decreased significantly from 1990 to 2010. The methodology of Barras et al. (2010) and Kosowski et al. (2006) does not enable investors to identify so-called skilled and outperforming funds respectively. Finally, Gronborg et al. (2021) proposed a method for the selection of the best mutual funds by performing pairwise comparisons between all funds. The particularity of this method is the fact that it enables investors to identify the name and number of funds outperforming the market through time. Moreover, the authors included the funds' holdings and returns data for the funds' alphas computations. Gronborg et al. (2021) based their pairwise comparison methodology on the work of Hansen et al. (2011), HLN henceforth, whose goal was to compare and pick outperforming models.

Methodologies naturally also exist in order to find outperforming hedge funds. Kosowski et al. (2007) proposed a analysis similar to their previous study in 2006 on mutual funds. Indeed, they also used a bootstrap analysis, as well as Bayesian methods to cope with the small sample issues, in order to prove that hedge funds' abnormal performance is not explained exclusively by luck. Their findings also indicate that hedge funds' superior performance persists. Titman and Tiu (2010) suggested that hedge funds managers which hedge their exposure to risk, whose R-squared is therefore low, present higher alphas, Information and Sharpe ratios, and hence a better performance. They ended up with latter results by identifying risk factors thanks to stepwise regressions. Then, hedge funds are ranked based on their R-squared generated by latter regression. Future performance of higher R-squared funds are proven to have a weaker performance than low R-squared funds. Chen et al. (2017) presented a method classifying hedge funds into three distinct groups, as Barras et al. (2010) did for mutual funds. They then used a modified expectation-maximization algorithm in order to find out the proportion of funds within each group, as well as the individual probability that a specific fund belongs to each group. Finally an individual fund's performance was performed by combining the fund's estimated alpha and the cross-sectional distribution of fund's skill. Their findings suggest that nearly 50% of the hedge funds analysed (i.e. 48%) are skilled, as opposed to the very small proportion of superior mutual funds indicated by the study of Barras et al. (2010).

In this paper, we will base our analysis on the methodology proposed by Gronborg et al. (2021), GLTW henceforth, applied to hedge funds instead of mutual funds. The purpose of this analysis is to be able to detect hedge funds significantly outperforming the competition. As stated by GLTW, the method employed in this paper allows us to know the proportion of funds that are outperforming, as well as to identify the funds with this superior performance. This last point differs from the methods presented here above. Following common practise, the eight-factor model of Fung and Hsieh will be our base to determine funds' risk-adjusted return (namely their alpha). Within this paper, we are pursuing two main objectives, which are first the correct adaptation of the GLTW's approach to hedge funds, and secondly the evaluation of latter method.

3 Methodology

As stated previously, the methodology used within this paper will be mainly driven by the paper of Gronborg et al. (GLTW), itself derived by the paper of Hansen et al. (HLN). This part of the report will precisely describe the different stages followed during this analysis in order to retrieve the outperforming set of hedge funds from our initial sample, thanks to efficient pairwise comparisons. This section will be divided in two main parts, namely the alpha computations details and the HFCS algorithm's explanation and interpretation.

3.1 Alpha and Predictive alpha computations

The first step of the analysis is to figure out a suitable measurement tool to assess hedge funds' performances. The calculation of the funds' alphas appears to be the most common practise for this purpose. As opposed to GLTW, which used as benchmark a four-factor model composed of the market, size, value and momentum factors in order to assess mutual funds' performance (respectively R_{mt}, SMB_t, HML_t, MOM_t), we will use the eight-factor model of Fung and Hsieh (1997,2004) which takes the following form:

$$R_{i,t} = \alpha_i + \beta_i \mathbf{z_t} + \epsilon_{i,t} \tag{3.1}$$

where

$$\mathbf{z_t} = (PTFS_{B_t}, PTFS_{Cur_t}, PTFS_{Com_t}, EQ_t, ES_t, BM_t, BS_t, EM_t)' \tag{3.2}$$

R_{it} represents the monthly return of fund *i*. $PTFS_{B_t}$, $PTFS_{Cur_t}$ and $PTFS_{Com_t}$ are respectively the Bond, Currency and Commodity Trend-Following factors introduced by Fung and Hsieh (2001)[2]. EQ_t is the Equity Market factor based on the S&P 500, while ES_t is the size spread factor. ES_t is calculated as the difference between the Russel index monthly total return and the S&P 500 monthly total return. BM_t and BS_t are the Bond Market and Bond Size Spread factors respectively. BM_t uses the monthly change in the 10-year treasury constant maturity yield[3], while BS_t represents the difference between the monthly change in the Moody's Baa yield[4] and above-mentioned BM_t. Finally, EM_t is the latest introduced Emerging Market factor, corresponding to the MSCI Emerging Market index. The regression of each hedge fund's returns with the above-mentioned factors enables us to find every fund's abnormal return, denoted α_i, as well as its risk exposures to said factors, denoted β_i.

As this pairwise comparison analysis will be performed monthly, it is necessary to have time-varying parameters α_i and β_i for each hedge fund. The traditional OLS regression on the whole time-series of returns does not allow us to benefit from such time-varying parameters. In order to implement a conditional factor model, multiple choices are available. Two main solutions are presented here below, namely the Rolling Window and the Kalman filter.

Firstly, estimating the above regression parameters on a rolling window is certainly the most straightforward method for this purpose. In this context, applying a rolling window regression is simply the act of running regressions repeatedly, with sub-samples of the original set of returns. The rolling window size usually depends on the type of analysis and data conducted. Secondly, the dynamic factor model can be based on a Kalman filter, as GLTW did in their analysis. The Kalman filter is a recursive mathematical tool estimating a state taking into account the previous estimate in its calculation. The conditional model on which the Kalman filter will operate is denoted as follows, with Equation (3.3) being the measurement equation and Equation (3.4) the process equation.

$$y_t = H_t x_t + v_t \tag{3.3}$$

[2] Data available at the following URL: http://faculty.fuqua.duke.edu/ dah7/DataLibrary/TF-Fac.xls

[3] Data available at the following URL: https://fred.stlouisfed.org/series/DGS10

[4] Data available at the following URL: https://fred.stlouisfed.org/series/DBAA

where y_t is the observable values which depend on the non observable x_t variable. H_t is the measurement matrix and v_t the measurement noise.

$$x_t = F_t x_{t-1} + w_t \tag{3.4}$$

where x_t is the state vector, F_t the transition matrix from time *t-1* to *t* and w_t being the process noise. We deliberately omit the control input in the process equation. These two state space equations can be applied to our eight-factor model in the following way:

$$R_{i,t} = \alpha_i + \beta_i \mathbf{z_t} + \epsilon_{i,t} \tag{3.5}$$

$$\alpha_t = \alpha_{t-1} + \epsilon_t^{\alpha} \tag{3.6}$$

$$\beta_t = \beta_{t-1} + \epsilon_t^{\beta} \tag{3.7}$$

The Kalman filter enables us to have conditional time-varying values of the fund's α, representing the intercept of Equation 3.5, and of the risk exposures with regard of the eight factors of Fung and Hsieh. As you can notice in Equations 3.6 and 3.7, both the $alpha$ and the different $beta$ estimates follow a simple random walk and the state space model put in place is a pure recursive model. Moreover, we make the assumption that the noise terms (i.e. $\epsilon_{i,t}$, ϵ_t^{α} and ϵ_t^{β}) are normally distributed, with a mean of zero and a constant variance.

We have decided to implement a Kalman filter to solve our state-space model because this tool has the aptitude to easily adapt to continuously changing data. Indeed, the Kalman filter consists of a prediction and an update step and therefore incorporates a term of imprecision within its own model. It therefore suits our hedge funds sample, whose returns can be close to unpredictable. However, the Kalman filter we are using differs from the work of GLTW. Indeed, their analysis is based on the work of Mamaysky et al. (2007,2008) which assumed that funds' managers receive unobserved information that has a certain correlation with the funds' returns and performance. This unobserved signal's parameters can be estimated through a Kalman filter. On the contrary, our use of the Kalman filter will estimate observed risk and performance parameters of hedge funds, namely the α and β of latter funds directly.

Many studies have used or assessed the Kalman filtering approach in order to estimate the abnormal returns and risk exposures of hedge funds. For instance, van Vuuren and Yacumakis (2015) indicated that dynamic αs and βs are estimated more accurately through a Kalman filter than with traditional OLS methods. Moreover, Amenc et al. (2010) suggested that the Kalman filter outperforms other dynamic approaches such as MRS (Markov Regime Switching model) in terms of explanatory power. The Kalman filter has also been used by many authors, such as Brunnermeier and Nagel (2004), Racicot and Théoret (2013) or Can and Liang (2012).

Once we have obtained the necessary time-varying alphas for all of our hedge funds, we can compute, as proposed by GLTW within their methodology, the following predictive alpha for all funds at each time *t*:

$$P_{i,t} = (R_{i,t} - \hat{\beta}'_{i,t}) \times \textit{sign}(\hat{\alpha}_{i,t|t-1}) \tag{3.8}$$

In this equation, *sign*$(\hat{\alpha}_{i,t|t-1})$ represents the sign of the fund's alpha forecast at time *t* based on information available in *t-1*. The latter can only be assigned a value of 0, -1 or +1 depending on the situation. The purpose of this equation is to monitor whether the alpha forecasts estimated thanks to information from *t-1* are reliable in predicting the performance of the different funds at time *t*. Indeed, the predictive alpha will control whether the information provided by the alpha forecasts is confirmed in the following time period. On the one hand, hedge funds whose alpha estimate does not accurately predict the fund's risk adjusted return will be penalized. This scenario will occur if a positive (negative) alpha estimate predicts a negative (positive) fund's risk adjusted return. On the other hand, funds whose positive (negative) risk adjusted return is anticipated by a positive (negative) alpha forecast will be rewarded.

As a follow up to the above, we can compute the average predictive alpha as shown here below:

$$\bar{P}_{i,t} = \frac{1}{t - t_{i0}} \times \sum_{\tau = t_{i0}+1}^{t} P_{i,\tau} \tag{3.9}$$

Following GLTW, the average predictive alpha described above is calculated on at least 12 monthly return observations but may include up to 60 observations if latter data is available.

3.2 HFCS Algorithm

Now that we have computed the time-varying alphas and predictive alphas of the whole initial set of hedge funds, the next step is the implementation of the HFCS methodology and algorithm. HFCS states for "Hedge Fund Confidence Set". Latter name is based from the work of GLTW whose superior final set of funds is called the Fund Confidence Set (FCS). The name FCS is itself inspired by the Model Confidence Set (MCS) implemented by HLN. As explained previously, the methodology of the HFCS will also be mainly based on the one of GLTW and HLN. Indeed, the initial set of hedge funds will be pairwise compared until a final superior set of funds is selected. HLN implemented this procedure with the aim of selecting superior models, while GLTW adopted this idea and applied it to pick outperforming mutual funds. As they are accounting for luck, these effective pairwise comparisons will enable us to find outperforming hedge funds which are selected solely based on their skill. The below-explained methodology relies of two main stages, which are (i) an equivalence test $\delta_{\mathcal{HF}}$ and (ii) an elimination rule $\epsilon_{\mathcal{HF}}$.

Let us start with an overview of the algorithm, as well as with some definitions. Let us denote the initial set of hedge funds, consisting of all hedge funds respecting the requirements explained in Part 3.4 of the report, $\mathcal{HF}_t^0 = \{HF_{1t}, ..., HF_{nt}\}$. The funds performance will be assessed through their predictive alphas, as described in Equation 3.8. The performance difference between fund *i* and fund *j* at time *t* is simply defined as below:

$$d_{ij,t} = P_{i,t} - P_{j,t}, \quad i, j \in \mathcal{HF}_t^0 \tag{3.10}$$

$\mu_{ij} = E[d_{ij,t}]$ is defined as the expected performance difference between fund *i* and *j* at time *t*. One will prefer fund *i* to fund *j* if $\mu_{ij} > 0$. Through multiple pairwise comparisons, our goal is to determine $\mathcal{HF}_t^*$, which represents the superior set of hedge funds. Funds which are significantly inferior to others are eliminated until only funds of significantly equal superior performance remain within this outperforming set of funds. Latter set is expressed as follows:

$$\mathcal{HF}_t^* = \left\{ i \in \mathcal{HF}_t^0 : \mu_{ij} \geq 0,\ for\ all\ j \in \mathcal{HF}_t^0 \right\} \tag{3.11}$$

As stated previously, the HFCS algorithm is constructed based on an equivalence test $\delta_{\mathcal{HF}}$. $\delta_{\mathcal{HF}}$ will consist in the testing of the null hypothesis stating that all funds within the remaining sample have an equal performance. The null hypothesis is written as follows:

$$H_{0,\mathcal{F}_t} : \mu_{ij} = 0, \quad for\ all\ i, j \in \mathcal{HF}_t \tag{3.12}$$

On the other hand, the alternative hypothesis states that the expected performance of at least one pair of funds differs and is described here below:

$$H_{A,\mathcal{F}_t} : \mu_{ij} \neq 0, \quad for\ at\ least\ one\ pair\ i, j \in \mathcal{HF}_t \tag{3.13}$$

Now that we have gone through these definitions, we can finally describe precisely how the HFCS algorithm will function. The first step of this algorithm is to take the whole set of hedge funds within the initial sample into consideration by setting $\mathcal{HF}_t = \mathcal{HF}_t^0$ at time *t*. The second step uses $\delta_{\mathcal{HF}}$ in order to test the null hypothesis described in Equation (3.12) at a significance level lambda (λ). The choice of λ will be discussed further in this report. Let us detail the specific constructed t-statistic of the equivalence test. Firstly, in order to assess the relative performance difference between fund *i* and fund *j* at time *t*, we denote Equation 3.14, taking into consideration Equations (3.8) and (3.10). In order to simplify the analysis and to to better adapt to the programming tools at our disposal, the time-frame on which both funds will be pairwise compared is set to 24 months and remains constant for all pairwise comparisons. This differs from the work of GLTW which implemented a variable time period depending on the two funds being compared.

$$\bar{d}_{ij} = \sum_{\tau=1}^{t} d_{ij,\tau} \tag{3.14}$$

Thereafter, we construct the t-statistic here below by dividing $\bar{d}_{ij}$ by its standard error, where *vâr* $(\bar{d}_{ij})$ is an estimate of *var* $(\bar{d}_{ij})$.

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\hat{var}\left(\bar{d}_{ij}\right)}} \tag{3.15}$$

As it is performed by HLN and GLTW, the test of the null hypothesis $H_{0,\mathcal{F}_t}$ will be based on the statistic described in Equation (3.16). The latter consists of the largest t-statistic out of the multiple pairwise t-tests performed as in Equation (3.15).

$$T_{\mathcal{HF}_t} = \max_{i,j \in \mathcal{HF}_t} |t_{ij}| \tag{3.16}$$

As HLN stated in their work, the distribution of the test-statistic $T_{\mathcal{HF}_t}$ is nonstandard, as it is impacted by nuisance parameters. Therefore, a bootstrap procedure can be implemented for the distribution of the t-statistic taken into consideration. The bootstrap approach used for this algorithm is mainly based on the work of White (2000). One of the characteristics of the bootstrapped implementation of HLN is the fact that the bootstrapped critical values of the distribution are adapted depending on the elimination round carried out. The latter enables the algorithm to be more robust statistically. The set of superior hedge funds is established if the p-value of the bootstrap procedure is higher than our λ, representing a quantile of the bootstrap distribution. Equation 3.17 displays the p-value formula, where $Tb_{\mathcal{HF}_t}$ is the bootstrapped statistics vector and B the number of bootstrapped samples. Using a bootstrapped procedure has the advantage of avoiding using a large covariance matrix. Moreover, the multiple bootstrap refinements allow for the consideration of initial data discrepancies.

$$Pvalue = \frac{\sum \left(Tb_{\mathcal{HF}_t} \geq T_{\mathcal{HF}_t}\right)}{B} \tag{3.17}$$

The last step of the HFCS algorithm depends on the equivalence test's results performed in the second step and described here above. Indeed, an accepted null hypothesis indicates that no funds is dominated by any others within the tested sample. This third step therefore sets $\mathcal{HF}_t^* = \mathcal{HF}_t$ and the procedure ends. On the other side, a rejected $H_{0,\mathcal{F}_t}$ implies that at least one fund is dominated by another. The elimination rule will therefore eject the weakest fund out of the sample. Latter elimination rule identifies the weakest fund between fund *i* and fund *j* being pairwise compared and is described as follows:

$$\arg \min_{i \in \mathcal{HF}_t} \sup_{j \in \mathcal{HF}_t} t_{ij} \tag{3.18}$$

This elimination rule defines the worst performing fund from all t-statistics conducted in Equation (3.15). Taking into account Equation (3.16), latter rule will only eject the worst performing fund out of the pair showing the largest t-statistic. Above Equation differs from the Model Confidence Set's elimination rule of HLN, which was defined as $\arg\max_{i\in\mathcal{HF}_t}\sup_{j\in\mathcal{HF}_t} t_{ij}$. Indeed, the data of HLN being pairwise compared are losses encountered by the analysed models. Therefore, the model contributing the most to the t-statistic t_{ij} is considered worse, hence the elimination rule amendment. Using Equations (3.16) and (3.18), we are able to figure out the weakest fund in the sample tested. The procedure outlined above will be repeated as long as the null hypothesis gets rejected, with the weakest fund being eliminated at each round. Once the null is no longer rejected, the process ends and the set of superior hedge funds $\mathcal{HF}_t^*$ is established.

3.3 HFCS Interpretation

In order to have a better understanding of the HFCS algorithm methodology outlined in the previous subchapter, let us have a discussion regarding its interpretation. Moreover, this part will address the choice of the significance level λ of the equivalence test. As previously explained, the HFCS procedure is based on multiple pairwise t-tests (detailed in Equation (3.15)) performed in order to eliminate weaker funds out of the initial sample. The elimination rule will eject the fund that is part of the pair showing the largest t-statistic. This t-stat will be increased if either the predictive alpha of fund *i* outperforms significantly the one of fund *j* for the period analysed (i.e. 24 months), or if the returns' correlation between fund *i* and fund *j* is high. Indeed, the standard error of the two funds' performance difference, denoted $\hat{var}(\bar{d}_{ij})$, has a tendency to be lower when correlation of the two funds returns is higher. The latter therefore results in a higher t-test value. These two aspects make perfect sense for our analysis. Indeed, two funds generating highly different predictive alphas are naturally good candidates for being considered as inferior. Moreover, considering two funds having correlated risk exposures as superior is of little interest for portfolio diversification purposes.

As briefly mentioned previously, the null hypothesis $H_{0,\mathcal{F}_t}$ test is naturally performed at a specific significance level or critical level , here denoted λ. This λ represents the probability of committing a type I error (wrongly rejecting the null hypothesis). In this context, this would indicate the probability of wrongly eliminating some funds not dominated by the set of superior funds. On the contrary, type II errors are interpreted as wrongly including an inferior fund within $\mathcal{HF}_t^*$. Following GLTW, we will compare the outcomes of resulting from different values of λ. Setting a high value of λ enables us to decrease the type II error. In other words, this will reduce the probability of incorrectly including poorer funds. However, the analysis will also be performed with a significance level of 0.1 in order to have a reference point for comparison purposes. A lower value of λ will increase the likelihood of including truly inferior hedge funds.

As a last point of interpretation, and as GLTW states, this HFCS procedure seems to ignore the positive diversification effects that less performing funds may add in the overall superior portfolios. Nevertheless, the HFCS can be interpreted as a tool that removes funds which are not expected to improve the information ratio of the portfolio. In other words, the hedge funds not included in the final HFCS have a low probability to improve the superior fund's performance through diversification effects. However, this does not prevent the improvement of the suggested HFCS by applying optimal weights to the selected funds. GLTW adopted the mean-variance efficient portfolios as proposed by Treynor and Black (1973). As the implementation of optimal weights does not improve significantly the performance of GLTW's outperforming portfolios, and in order to simplify our analysis, we decided to assign equal weights to the hedge funds selected by the HFCS procedure.

3.4 Operational requirements

As you may have noticed, the HFCS algorithm requires a large number of pairwise comparisons. Indeed, if 150 funds are being compared for a certain month, (150*149) pairwise comparisons are necessary in order to eliminate only the first fund. Then (149*148) comparisons are needed for the second elimination, and so on. This quantity requires a large amount of time and computing power. Following GLTW, we therefore set some constraints on the hedge funds selected for this procedure. First, at each time *t*, we will remove hedge funds whose alpha forecasts are negative in the subsequent period *t+1*. Indeed, our goal is to pick funds that will outperform in the future. In addition, we will remove hedge funds which have a negative average predictive alpha, as calculated as in equation 3.9. Latter constraint enables us to make sure that the hedge funds' alphas are actually predictable. These requirements allow us to reduce significantly the monthly number of pairwise comparisons needed. Furthermore, as stated previously, we require the hedge funds to be pairwise compared on a 24-month time period, which also narrows the number of funds available at the date under analysis. In order to implement this algorithm, we will primarily use the software R, and more specifically the R package "MCS" of Catania and Bernardi (2017). This package was designed to conduct the Model Confidence Set of HLN and we slightly modified it in order to cope with the elimination rule update we discussed earlier. For more clarity, we have summarised the HFCS steps here below, mainly derived from the GLTW, which will be repeated at each month from April 2001 to September 2021.

1. Calculation of the alpha forecasts $\hat{\alpha}_{i,t+1|t}$ for the subsequent month (using a Rolling window or a Kalman filter) and elimination of funds with a negative alpha estimation.
2. Computation of the predictive alpha $P_{i,t}$ (as in Equation 3.8) and discard of hedge funds with a negative average predictive alpha $\bar{P}_{i,t}$ (as in Equation 3.9).
3. Execution of the HFCS algorithm on the remaining funds with pairwise comparisons on a 24-month time period.
4. Formation of a portfolio consisting of outperforming hedge funds selected by step 3.
5. Collection of portfolio returns on the analysed time-frame
6. Model's assessment through the risk-adjusted return α of the portfolio (using the eight-factor model of Fung and Hsieh from Equations 3.1 and 3.2)

4 Data Presentation

The data used for this empirical analysis is a sample of monthly hedge funds returns domiciled in the United States. The United States are indeed undoubtedly the country with the highest hedge funds Asset Under Management worldwide, accounting for 2.7 billion in 2019[5]. Funds' returns are provided by the database EUREKAHEDGE and the time period analysed goes from April 1994 to September 2021. We require at least 84 months of returns observations and six months of continuous data available for each hedge fund. Moreover, the analysis is based on hedge funds whose strategy is driven by Long/Short equities. Indeed, this strategy is the most widespread within the hedge fund industry (as shown on Figure 4) and is the one most similar to the initial mutual funds sample of GLTW. 743 hedge funds fulfil these data and strategy requirements. Following GLTW, we also have restricted our initial sample in accordance with the individual funds' goodness of fit with the chosen regression model. In their paper, a R^2 threshold of 0.7 was set for individual funds to be included in the analysis. However, only 27 hedge funds meet this requirement when being applied to the eight-factor model of Fung and Hsieh. Therefore, we have amended latter limitation to 0.5. 173 hedge funds actually comply with this modified limit. As this R^2 level is far from substantial, these results therefore attest that Fung and Hsieh 8-factor model does not explain adequately most of the analysed hedge funds' data variations. Finally, both obsolete and live funds will naturally be taken into consideration for this study.

One of the biggest differences with the analysis of GLTW is the absence of holdings data in our study. Indeed, latter paper included mutual funds' holdings in the calculation of the characteristic selectivity (CS) of Daniel et al. (1997), which represents, in a nutshell, the abnormal return of an individual stock compared to a relevant benchmark. We have decided not to include hedge funds holdings for several reasons. First of all, only 230 out of the 743 L/S Equity hedge funds had disclosed holdings information within the SEC 13F Form. Furthermore, only few of these 230 hedge funds reported comprehensive and reliable holdings information. Indeed, unlike mutual funds, hedge funds do not have the obligation to disclose their holdings information. Despite the fact that larger funds (calculated through their AuM) have stricter obligations regarding their holdings information than smaller ones, the information disclosed by the latter remains very opaque and incomplete. We can therefore question the accuracy and reliability of the information provided by the few hedge funds disclosing their holdings details. Then, as Shi (2017) indicated, hedge funds disclosing their holdings information are subject to a decrease of performance and to higher returns correlation with other competitors subsequently. As a consequence, the analysis of these particular hedge funds may seem less meaningful with respect of the procedure we implement. Finally, the non incorporation of holdings returns simplify significantly the model allowing for time-varying alphas, as the right hand side of the state-space's measurement equation becomes linear instead of quadratic.

Let us first take a closer look at the behaviour of our data, as well as at some statistics referring to it. The AuM average of the 173 considered hedge funds is worth approximately US 1,100 million and its median is about US 200 million. Both figures are slightly higher than the average of L/S equity hedge funds as a whole. Moreover, the average lifetime of the funds analysed is 160 months, corresponding to slightly more than 13 years. Again, the tested sample has a lifespan slightly above average.

With respect to the performance of the hedge funds analysed, let us analyse the funds' alphas. The simple OLS regression constant alpha of all 743 L/S equities hedge funds[6] has a negative mean of (0.35%) and a median of 0%. Our analysed set of hedge funds shows higher abnormal returns, with an average of 0.75% and a median of 0.68%. Both distributions are presented in Figure 5 and 6. We can notice that the analysed sample's distribution is more negatively skewed. Cross-sectional time-varying alphas present different results and related statistics are displayed in Table 1. An individual fund's average time-varying alpha is simply the mean of its dynamic alpha on the time-frame it existed. The average time-varying alpha's mean of the analysed sample is 2.55% and the cross-sectional distribution is presented in Figure 7. These results are estimated through the rolling window methodology and are significantly different from the outcomes of GLTW. Indeed, the mutual funds' cross-sectional time-varying alphas had a negative mean of (0.072%). This result is mainly

[5]According to Statista

[6]Respecting the 72-month data requirement

explained by the extremely high time-varying alphas' values of certain top hedge funds. Indeed, the top 10 percent of hedge funds, ranked by time-varying alpha performances, has an alpha estimate of 14.08% and the top 5 percent funds achieve an abnormal time-varying return of nearly 20%. In contrast, the bottom 10 and 5 percent of hedge funds have a negative time-varying alpha value of (4.63%) and (6.61%) respectively. the results concerning predictive alphas are even more extreme as merely none of the hedge funds analysed has a negative average predictive alpha within our analysed sample. The related distribution can be found in Figure 8. The annualized Sharpe ratio of the tested sample[7] has a mean of 0.59. This performance overview remains approximately similar when estimating the time-varying alphas through the Kalman filter approach. The average time-varying alphas and predictive alphas' distributions regarding the Kalman filter approach are displayed in Appendix A.

	Rolling window	
	Alpha	**Sharpe Ratio**
Mean	2.55%	0.59
Bottom 5%	-6.61%	-0.11
Bottom 10%	-4.63%	-0.01
Top 10%	14.08%	1.18
Top 5%	19.97%	1.27

Table 1: Time-varying alphas and Sharpe Ratio statistics

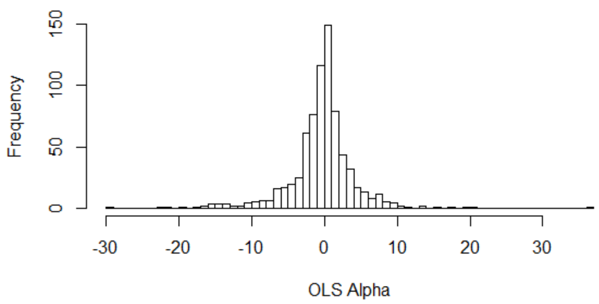

Figure 5: L/S equities HF: Constant OLS alpha distribution

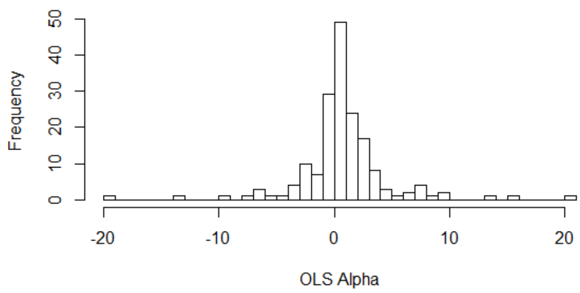

Figure 6: Analysed sample: Constant OLS alpha distribution

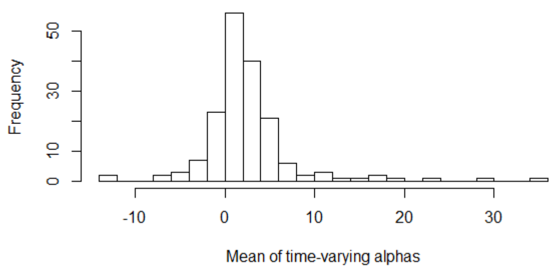

Figure 7: Mean of time-varying alphas

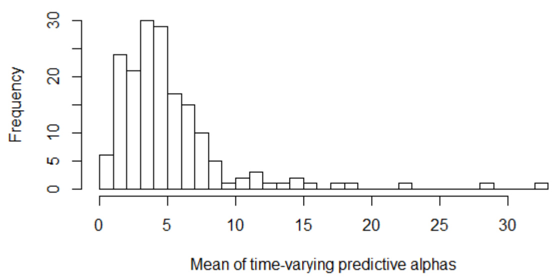

Figure 8: Mean of time-varying predictive alphas

[7] Calculated by dividing the fund's average return by its standard error, and then multiplying it by $\sqrt{12}$

5 Results of the analysis

Now that the methodology of our analysis has been detailed and now that our data has been introduced, it is time to present our results. As a reminder, the procedure employed is based on multiple pairwise comparisons within our initial hedge funds sample at each time t. As it takes 60 months for the first alpha predictions to be computed (i.e. size of the rolling window for the regression parameters in Equation 3.1) plus 24 months for the collection of sufficient historical predictive alphas values, our analysis will be carried out from April 2001 to September 2021. Although the Kalman filter approach does not require 60 months of historical returns, we will also remove this period of time out of our analysis in order to reduce estimation errors that might occur for first alpha estimates. Our main goal is to find the subset $\mathcal{HF}_t^*$ which contains superior and outperforming hedge funds. On a first part, we will describe and analyse the procedure on three hedge funds at a certain date in order to have a better understanding of the methodology applied. In a second stage, we will describe the final results of the HFCS procedure applied to the entire fund sample and on the whole time-frame. Finally, we will evaluate the performance of our models by assessing the performance of the portfolios formed of funds selected by our approach. For your information, this methodology could also be applied in order to select an inferior set of hedge funds, as GLTW did in their analysis. Nevertheless, latter study will not be conducted in this paper.

5.1 HFCS example

In order to have better insights with regards of the HFCS procedure and the methodology applied, let us start this discussion with a small, easy-to-understand example. For this purpose, let us analyse and compare three hedge funds on a randomly chosen date, which will be October 2010. The three hedge funds are called 'Alder Capital Partners I LP', 'Alpha Equity Global L/S Equity' and 'Archer Equity Fund LLC'. These funds meet the data requirements stated in part 3.4 of the report and are part of the 53 hedge funds pairwise compared in October 2010. You can find the predictive alphas of the three funds for the 24 months on which the latter are compared In Appendix B. Indeed, the predictive alphas are the values on which the HFCS algorithm will be based, as detailed in Equation (3.10). These predictive alphas will be pairwise compared by applying the t-stat defined in Equation 3.15. The numerator of latter equation is displayed in Appendix C. Thereafter, the mean of all performance differences is calculated (as shown in Table C.2 of Appendix C) and will be divided by its standard error. The standard error estimate is determined thanks to the the bootstrapped procedure implemented. The six resulting t-statistics are displayed in Table 2.

Fund I - Fund II	Fund I - Fund III	Fund II - Fund I	Fund II - Fund III	Fund III - Fund I	Fund III - Fund II
-0.46	-0.02	0.46	0.60	0.02	-0.60

Table 2: Pairwise T-statistics - Round 1

Fund I - Fund II	Fund II - Fund I
-0.46	0.46

Table 3: Pairwise T-statistics - Round 2

As explained in the methodology section of the report, the null hypothesis $H_{0,\mathcal{F}_t}$ will be tested through the larger t-statistics of the pairwise compared funds, as defined in Equation 3.16. Table 2 shows us that the t-test comparing Fund II ('Alpha Equity Global L/S Equity') and Fund III ('Archer Equity Fund LLC') is the largest, worth 0.6. The p-value of the our procedure will therefore be calculated on this largest value. As a reminder, this p-value is computed with respect to the bootstrapped t-statistics of resampled data. For this specific case, the p-value equals 0.849. In order to check if the superior set of funds is accepted, this p-value should be higher than the λ quantile of the bootstrapped distribution. For this example, we will choose a λ value of 0.9 and the requirement is therefore not respected. This means that one of the funds within our initial sample is outperformed by at least one other at a significance level of 0.9. Using Equations 3.18 and 3.16, we can conclude that Fund III, namely 'Archer Equity Fund LLC', is inferior and is therefore eliminated from the initial sample. The procedure is repeated with the two remaining funds. The HFCS p-value is now worth 0.65 and is, here again, below our significance level λ set at 0.9. It means once again that one fund outperforms the other in the remaining sample. Table 3 shows the pairwise t-tests and Fund I ('Alder Capital Partners I LP') will be removed from the sample. Fund II ('Alpha Equity Global L/S Equity') represents therefore the only hedge fund considered as superior out of the initial sample. However, as we could see, the choice of the confidence level λ impacts strongly the result of the analysis. Indeed, a λ value lower than the first HFCS p-value worth 0.849 would have involved that all three funds are considered as superior, as the null hypothesis $H_{0,\mathcal{F}_t}$ would have been accepted. The bootstrapped distributions of rounds one and two of this HFCS example, on which the HFCS probability has been computed, can be found on Figures 9 and 10 respectively. For this example, a number of 1,000 bootstrapped samples have been used to construct the statistic test.

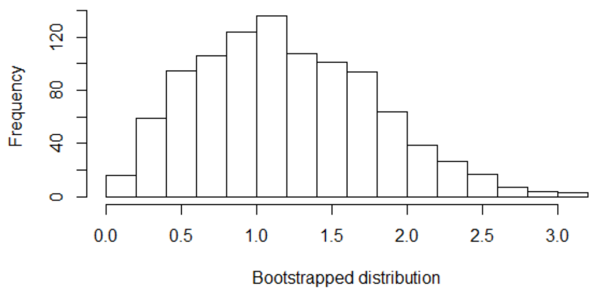

Figure 9: HFCS Example Round 1 - Bootstrapped distribution

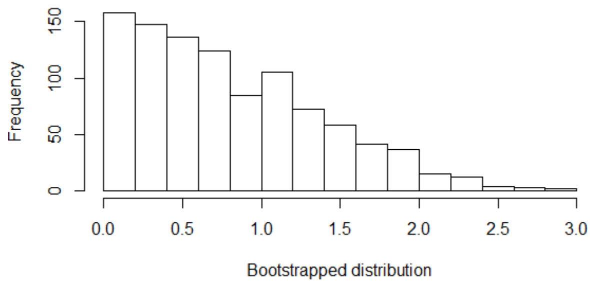

Figure 10: HFCS Example Round 2 - Bootstrapped distribution

5.2 Results of the complete analysis

The HFCS algorithm enables us to pick several hedge funds each month from April 2001 to September 2021. In this part, we will analyse how our algorithm performed for this time-frame. As previously stated, we have carried out our study with two different time-varying alpha computation approaches. First, we used a 60-month rolling window in order to calculate the alpha estimates, henceforth called the RW approach. Second, we implemented a state-space model whose parameters (namely the α and β) are estimated thanks to a Kalman filter, henceforth called the KF approach. This section will take into consideration and compare both approaches. Furthermore, as GLTW did, we carried out this analysis using different significance levels, denoted λ. As stated previously, we decided to compare a λ value of 0.9 with a value of 0.1 to check the impact of said significance level on the equivalence test. Results of both options will be discussed here below. As a reminder, choosing high values of λ enables us to reduce the probability of committing a type II error, representing the probability of wrongly including inferior hedge funds in our final HFCS.

In order to ensure that our HFCS algorithm accurately predicts outperforming hedge funds, let us use the following strategy. Imagine an investor mimicking the algorithm outcomes (i.e. investing in all funds considered as superior by the latter). At each month t, he will invest in the funds included in the HFCS and record the returns of the resulting portfolio for the following month. To simplify this step, an equal-weighting is applied to all superior funds considered within the portfolio. As time passes, a time-series of returns from April 2001 to September 2021 is formed. From these returns, we will simply apply the same eight-factor model of Fung and Hsieh that we used for the individual funds' alphas estimations, described in Equations 3.1 and 3.2. The regression of our portfolio returns with the eight risk variables will allow us to determine the final alpha of our different portfolios and latter alpha will be the base of our models' performance assessments. Furthermore, a similar analysis will be carried out to alternative performance models, more traditional and easy-to-use than our HFCS procedure. Following GLTW, we will compare our HFCS methodology to two main models. Firstly, we will simply base the choice of our hedge funds at time *t* on the funds' alpha forecasts calculated in *t-1*. We will therefore be able to conclude whether funds with high historical alphas are likely to continue outperforming in the future. As for our HFCS procedure, we will assess this technique by investing in the higher ranked alpha-generating funds at each time *t* and get returns of latter funds at *t+1*. Similarly, the final alpha of our portfolio will be determined applying the eight-factor model of Fund and Hsieh to the collected returns. The second main model on which we will compare the HFCS with is basically equivalent to the alpha-ranking method explained here above but with the difference that it will be ranked with respect of the funds' predictive alphas, as defined in Equation 3.8.

5.2.1 HFCS Performance

The methodology presented as an example in section 5.1 will be now applied to the entirety of the hedge funds over the complete time-frame. Many portfolios will be compared for this purpose. As explained previously, the differences will mainly come from the significance levels used for our equivalence test and from the time-varying alphas computation approaches we used. Henceforth, the terms "HFCS tight" and "HFCS wide" will be used to refer to the significance levels of 90% and 10% applied to the equivalence test, respectively. As a trade-off between the statistical robustness of our procedure and the computing power necessary for the carrying out of the analysis, we have decided to use 1,000 bootstrapped samples in order to construct the t-stat on which our procedure is based.

As anticipated, the size of the superior set of funds is considerably different depending on these two variables. As you can see in Figures 11 to 14, the HFCS wide using the RW approach is the one constituted by the largest number of funds, amounting to 883 funds in total, or 3.6 funds per month on average. In comparison, the number of hedge funds picked by the HFCS tight reaches 366 in total, corresponding to approximately 1.5 funds per month on average. These results are distinctively different when it comes to the KF approach. Indeed, the algorithm is considerably more stringent regarding the choice of the funds which constitute it. As a matter of fact, the wide HFCS of the KF approach only includes 331 funds in total, while the tight KF HFCS comprises 251 funds in total, representing almost only one hedge fund per month, as can be seen in Figure 14. This can be explained by the fact that the Kalman filter is a tool which adapts quickly to the data variations encountered by our hedge funds' returns. This may therefore result in greater differences in performance between the funds compared. Furthermore, the fact that tight HFCS are made up of fewer superior hedge funds than wide HFCS is logical and reflects the same outcome as GLTW. Indeed, the equivalence testing approach discussed earlier explains this result.

One of the striking characteristics of Figures 11 to 14 is that the different monthly HFCSs are often composed of a single hedge fund. As you can notice, the latter is mainly the case for the KF approach. This outcome is primarily due to the large range of alphas generated by the hedge funds analysed. As shown on Figures 8 and A.2, some hedge funds have very high predictive alphas and it is therefore more probable to experience high performance differences between paired funds. As an example, Appendix D shows the RW time-varying alphas, predictive alphas and returns of the hedge fund named 'The Vilas Fund'. Latter fund is, as you could expect, outperforming any others for the majority of its lifetime. Nevertheless, we decided not to remove these super performing funds as their performance is pairwise compared on a substantial time period of 24 months and they are therefore considered as reliable and accurate.

As we have explained previously, we will assess our HFCS algorithm performance by forming portfolios of the hedge funds picked by the latter, and by collecting their returns on the analysed time-frame. The models assessments will be carried out by retrieving the portfolios' alphas from these time-series, applying the eight-factor model of Fung and Hsieh. It is plausible that some of the funds selected by our procedure at time *t* are no longer active at time *t+1* (for instance due to liquidation or a simple cessation of activities). It is therefore not possible to collect the returns from this fund at time *t+1* and to check whether our procedure is performing correctly. We have therefore decided to proceed as follows. If the fund concerned by this eventuality is accompanied by other superior funds in the month in question, the return of the portfolio will only depend on the remaining funds, on an equal-weighting allocation. However, this solution is not applicable if the HFCS is

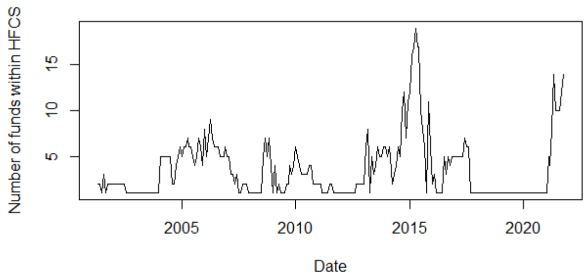

Figure 11: HFCS Wide - Number of funds (Rolling window)

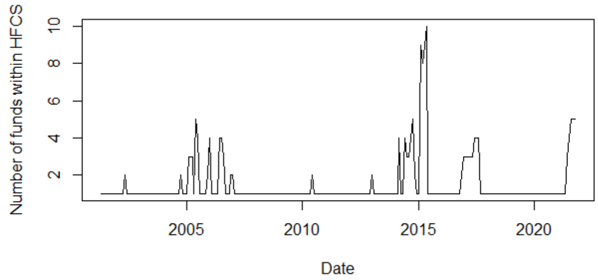

Figure 12: HFCS Tight - Number of funds (Rolling window)

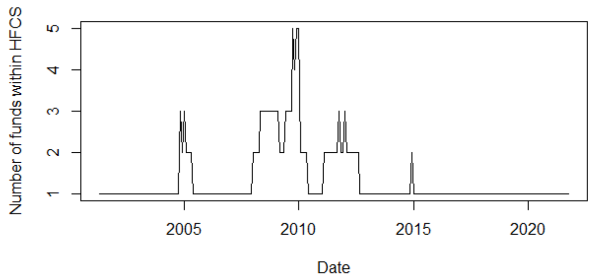

Figure 13: HFCS Wide - Number of funds (Kalman filter)

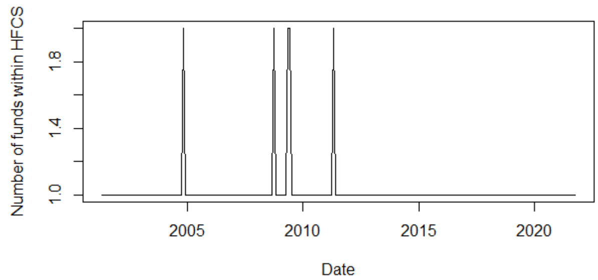

Figure 14: HFCS Tight - Number of funds (Kalman filter)

composed of a single fund, which is, as showed above, very frequent in our analysis. In this case, we replace the concerned fund by the last eliminated hedge fund during the procedure and the latter's return will be taken into consideration in our time-series. The returns of all four final portfolios are displayed in Appendix E. You will notice the high variance of these returns from 2018 to 2020, corresponding to a single top-performing hedge fund selected by the HFCS and generating highly volatile returns and alphas, namely the hedge fund 'The Vilas Fund LP' whose returns, alphas and predictive alphas are displayed in Appendix D.

HF in portfolio	**Alpha**	**T-value**	**P-value**	**SR**
Rolling window				
Wide HFCS	1.63	0.55	0.58	0.14
Tight HFCS	1.17	0.38	0.71	0.15
Kalman Filter				
Wide HFCS	0.84	0.30	0.77	0.16
Tight HFCS	0.85	0.29	0.78	0.15

Table 4: Performance of HFCS Portfolios

Table 4 represents the four portfolios created by the previously explained methodology. As you can see, the portfolios are assessed according to both their final alphas and to the significance of said alphas. A first encouraging comment is the fact that all formed portfolios have a positive abnormal performance, ranging from 0.84% for the Wide KF HFCS to 1.63% for the Wide RW HFCS. This result is not surprising as our initial set of hedge funds has a positive alpha and time-varying alpha average, as shown in Figures 6, 7 and A.1. Nonetheless, the portfolios' alphas are far from convincing, as none of them are close to being significant (T-values ranging from 0.29 to 0.55). Furthermore, we can notice that the RW approach outperforms the KF approach in terms of both risk-adjusted return and significance. Lastly, the significance level adopted for our analysis did not have the intended impact. As you can see, the HFCS wide outperforms the HFCS tight for the RW approach. This outcome differs from the results of GLTW, whose Tight portfolios performed significantly better than Wide ones. Here are an assumption that may explain this result. A high significance level of λ indeed decreases the probability of wrongly including an inferior fund but also increases the chance of omitting genuinely superior hedge funds. A value of 0.9 may therefore have caused the non-integration of hedge funds which would have positively impacted the risk-adjusted return of the whole portfolio. As many months are composed of a single fund, failing to include a truly superior fund can have an insightful impact on the portfolio's performance. In addition, portfolios containing a single fund are seemingly outperformed by portfolios consisting of multiple equally-weighted assets. Appendix F.1 shows multiple predictive alpha-ranked portfolios[8]. These portfolios maintain their good performance as long as the selection criterion gets stricter, with the exception of the "Top 1" portfolio, consisting solely on the highest predictive alpha out of the analysed sample, which performs substantially worse. This outcome could also explain the lower performance of the KF approach, whose portfolios consist mainly of single hedge funds.

5.2.2 Performance of alpha and predictive alpha sorted portfolios

As we could figure out in the previous section, the choice of outperforming hedge funds based on our HFCS procedure provides us with mixed results. Indeed, all constructed portfolios have positive abnormal returns over the period April 2001 to September 2021 but their alphas are all far from the significance desired. In order to evaluate latter approach more thoroughly, let us compare it with more traditional approaches solely based on funds' alpha or predictive alpha estimates.

For this section, we will base our analysis on three portfolios for both the sorted alphas and sorted predictive alphas strategies. Firstly, we will focus on a portfolio consisting of all hedge funds having a positive alpha (predictive alpha) forecast. Then, we will frame our analysis by choosing the top 10 and top 20 percent highest ranking funds (with regards to their alphas or predictive alphas). As a first step, the analysis will be conducted on the sample of funds included in the initial set on which the HFCS procedure will be applied. This enables to better compare both approaches. The results can be found in Tables 5 and 6 for alpha-ranked and predictive alpha-ranked portfolios respectively. As you can notice, the tables are separated with regards of the time-varying alphas computation approach, namely the RW and KF approaches. The number of funds taken into consideration for these portfolios over time are reported in Appendix G using the RW approach for time varying alphas computation and in Appendix H for the use of the KF approach. Figures G.1 and G.2 represents the number of funds part of the top 20% and 10% portfolios, while Figures G.3 and G.4 feature the number of funds in the portfolios consisting of funds with positive alphas and predictive alphas respectively. The same layout is adopted for the KF approach in Appendix H.

HF in portfolio	Alpha	T-value	P-value	SR
Rolling window				
Positive α	0.92	1.96	0.05	0.68
Top 20$\%$ α	0.91	1.15	0.25	0.57
Top 10$\%$ α	1.06	0.96	0.34	0.47
Kalman Filter				
Positive α	0.62	1.30	0.20	0.68
Top 20$\%$ α	0.58	0.74	0.46	0.47
Top 10$\%$ α	0.88	0.82	0.41	0.31

Table 5: Model performance - Alpha-ranked portfolios (HFCS sample)

[8] Alpha and Predictive alpha-ranked portfolios will be assessed in the subsequent section

HF in portfolio	Alpha	T-value	P-value	SR
Rolling window				
Positive Pred. α	1.00	2.08	0.04	0.69
Top 20$\%$ Pred. α	0.53	0.64	0.52	0.61
Top 10$\%$ Pred. α	0.51	0.45	0.65	0.57
Kalman Filter				
Positive Pred. α	0.76	1.58	0.12	0.66
Top 20$\%$ Pred. α	0.56	0.67	0.50	0.47
Top 10$\%$ Pred. α	1.32	1.24	0.22	0.36

Table 6: Model performance - Predictive alpha-ranked portfolios (HFCS sample)

As you can see, these tables show the statistics for both alpha-sorted and predictive alpha-sorted portfolios, and for both the RW and the KF dynamic approaches. The first row consists of all funds having a positive alpha (predictive alpha in Table 6) forecast. These tables consist then of the top 20 and 10 percent of funds with respect to their alphas (predictive alphas in Table 6) in the second and third row respectively. The first interesting comment is that, as was the case for our HFCS procedure, all portfolios generate a positive risk-adjusted return by the reporting period. The second insightful remark is that predictive alpha-ranked portfolios do not necessarily outperform alpha-ranked portfolios. Indeed, their α and T-values remain more or less identical. As a third comment, all portfolios but the 10% predictive alpha-ranked one, generate higher alphas and are more significant when using the RW approach. As an example, the top 20% higher ranked funds with respect to their RW time varying alphas generate a portfolio α of 0.91%, while its counterpart using the KF approach has a portfolio α of 0.58%. In addition, you can notice that the portfolios' performances has the tendency to improve, the tighter the selection of funds in the portfolio becomes. In other words, selecting the top 10% of hedge funds to compose our portfolio will usually produce greater results than picking the top 20% of hedge funds. As a last point, almost none of the portfolios generate significant values of alphas, with the exception of the portfolio consisting of funds having positive Predictive alpha estimates.

In general, HFCS portfolios' (see Table 4) have a superior performance than alpha and predictive alpha-sorted portfolios (see Tables 5 and 6). With the RW approach, Wide and Tight HFCS have a performance of 1.63% and 1.17% respectively. On the contrary, following the same approach, the most performing portfolio consists of the Top 10% funds with regards of their alpha forecasts (with a value of 1.06%). The KF approach does not demonstrate such clear findings, as the Top 10% of both alpha and predictive alpha-ranked portfolios have a higher and more significant abnormal return than the HFCS procedure. Effectively, both KF HFCS portfolios generate an α close to 0.84%, while the Top 10% of alpha and predictive alpha-ranked portfolios have a final performance of 0.88% and 1.32%. Moreover, the alpha significance of alpha/predictive alpha-ranked portfolios is constantly higher than the one of HFCS final sets of hedge funds. Finally, the annualized Sharpe Ratio of alpha and predictive alpha-sorted portfolios are substantially higher than for our HFCS portfolios.

As a reminder, the above-explained analysis consists of hedge funds included in the initial set of funds evaluated by the HFCS procedure. In other words, as stated in section 3.4 of the report, funds included in the portfolios have at least 84 months of past data (i.e. 60-month rolling window for the time-varying alpha computations and 24 months for the the pairwise comparisons of predictive alphas). Moreover, hedge funds whose alpha forecasts for the subsequent month, and whose average predictive alpha at the current month, are negative, are removed from the initial sample. Naturally, a strategy focusing solely on the funds selection based on alpha estimates takes into consideration the entirety of funds having an alpha estimate available at the current month analysed. The constraint requiring a 24-month period of continuous alphas, as well as the one removing funds having a negative alpha estimate for the period ahead, should not be incorporated here. Tables 7 and 8 display the same analysis as performed previously, but focusing on all alphas available at the analysed month, rather than on funds meeting the HFCS requirements. The subsequent analysis has not been performed by GLTW as they compared FCS portfolios with alpha-sorted portfolios following the FCS requirements.

HF in portfolio	Alpha	T-value	P-value	SR
Rolling window				
Positive α	0.76	1.59	0.11	0.68
Top 20$\%$ α	0.84	1.30	0.19	0.68
Top 10$\%$ α	1.12	1.27	0.20	0.58
Kalman Filter				
Positive α	0.95	2.05	0.04	0.71
Top 20$\%$ α	1.26	2.00	0.05	0.67
Top 10$\%$ α	1.41	1.73	0.09	0.62

Table 7: Model performance: Alpha-ranked portfolios

HF in portfolio	Alpha	T-value	P-value	SR
Rolling window				
Positive Pred. α	1.01	2.08	0.04	0.64
Top 20$\%$ Pred. α	1.50	2.17	0.03	0.50
Top 10$\%$ Pred. α	1.92	2.20	0.03	0.49
Kalman Filter				
Positive Pred. α	1.20	2.67	0.008	0.69
Top 20$\%$ Pred. α	2.14	3.19	0.002	0.68
Top 10$\%$ Pred. α	2.18	2.64	0.009	0.67

Table 8: Model performance: Predictive alpha-ranked portfolios

The first comment standing out from these tables are the significantly better performances of the portfolios analysed compared with the analysis of Tables 5 and 6. This is especially the case for the KF approach and for predictive alpha-sorted portfolios. Indeed, predictive alpha-ranked portfolios all have an α higher than 1.00 and the latter is always significant, which was not necessarily the case previously. Furthermore, the KF approach now outperforms considerably the RW approach for all formed portfolios and predictive alpha-ranked portfolios perform considerably better than alpha-sorted portfolios. This means that the use of the proposed predictive alpha approach defined in Equation 3.8 might help generating better performances.

5.2.3 Performance of alternative approaches

The portfolios performances obtained by taking all available alphas into consideration radically changes the comparison between alpha-ranked and HFCS portfolios. Regarding the RW approach, the HFCS Wide and Tight portfolios α still outperform the alpha-ranked portfolios but the latter are strongly competed by predictive alpha-ranked portfolios performances. Furthermore, KF alpha-sorted portfolios perform approximately similarly to HFCS ones, while predictive alpha-ranked performances significantly outperform our HFCS procedure. Finally, most alpha and predictive alpha-sorted portfolios have significant values of their risk-adjusted returns, which is not the case regarding HFCS alphas. The HFCS procedure therefore outperforms alpha and predictive alpha-ranked portfolios when the latter are based on the same analysed sample of hedge funds but is outperformed by the former when the HFCS restrictions are not taken into account.

The interpretation of these surprising results is presented as follows. The lower performance of alpha and predictive alpha-ranked portfolios when considering the HFCS hedge funds sample suggests that the funds selection criteria implemented for the HFCS procedure (see section 3.4) reduce the subsequent performance of the formed portfolios. Those criteria, namely (i) the discard of funds with a negative alpha forecast for the following period, (ii) the rejection of funds having a negative average predictive alpha and (iii) the period of time needed to collect predictive alpha values in order to pairwise compared them (i.e. 24 months)[9] must therefore result in the rejection of some hedge funds whose performance improves the final portfolio created

[9]This third criteria differs from the approach of GLTW, which implemented a varying period of time depending on the two funds compared.

by the set of superior funds. If criteria (i) is the one impacting this lower performance, this would suggest that the factor model used (namely the eight-factor model of Fung and Hsieh) does not accurately estimate our hedge funds' alphas. As a consequence, funds generating a negative alpha forecast for the subsequent period will be removed following the HFCS procedure, although latter fund might have been performing well and may have improved the portfolios' performances. If criteria (iii) is the one having this negative impact, this would suggest that the time-period implemented to collect the data necessary for our pairwise comparisons (i.e. 24 months) made us miss out on outperforming funds. Indeed, the latter did not have enough historical data available. Criteria (ii) is not likely to have impacted the study as most of the hedge funds analysed have a positive average predictive alpha throughout their lifetimes (as shown on Figure 8).

In order to further investigate latter hypothesis, let us slightly modify the initial sample on which our best performing portfolio, namely the RW Wide HFCS, was constructed. For this purpose, we will first consider a sample removing criteria (i), meaning that funds with a negative alpha forecast for the subsequent period will not be removed from the initial HFCS sample. Then, we will consider a shorter time-period on which funds' predictive alphas are being pairwise compared (i.e. the time period requirement will go from 24 to 12 continuous months). Table 9 presents the results.

HF in portfolio	**Alpha**	**T-value**	**P-value**	**SR**
Rolling window				
Wide HFCS: initial	1.63	0.55	0.58	0.14
Wide HFCS: α criteria	1.81	0.62	0.53	0.14
Wide HFCS: time criteria	-0.94	-0.34	0.74	0.21

Table 9: Performance of the RW Wide HFCS - Amended initial criteria

As you can notice, the removal of the alpha criteria slightly improves the overall portfolio performance, with a portfolio alpha increasing from 1.63% to 1.81%. The significance of latter alpha also increased but remains disappointing. On the other side, the hedge fund sample pairwise compared on a 12-month time period show a substantially lower performance, with a negative alpha value. This rapid assessment suggests that the removal of funds with negative alpha forecasts for the period ahead slightly decreased the performance of our HFCS portfolios and it therefore also suggests that the eight-factor model of Fung and Hsieh fails at accurately predicting the analysed hedge funds' abnormal returns. Nevertheless, this analysis should ideally be carried out on all remaining portfolios to ensure the validity of this outcome.

6 Conclusions and Discussions

The choice of outperforming hedge funds has been a challenging subject to academics and institutional investors for a long time. Although the hedge fund industry has increased in popularity for the last decades, no accurate forecasting model of future outperforming funds proved its efficiency as of today. Our set-identification approach, derived from the models of Gronborg et al. (2021) and Hansen et al. (2011), carries out multiple pairwise comparisons between U.S. L/S Equities hedge funds. The final funds included within the set of superior hedge funds (HFCS) are theoretically outperforming for a significant amount of time and are therefore expected to continue generating a high performance in the future. Although this methodology produces highly satisfactory results on mutual funds, the conclusion differs with respect to the analysed hedge funds.

Indeed, the abnormal returns of the four constructed HFCS portfolios (namely the Wide and Tight HFCS of both the RW and KF approaches) range from 0.84% to 1.63% on the period April 2001 to September 2021, which outperform the mutual funds' risk-adjusted returns values retrieved by GLTW on their FCS implementation. Nonetheless, the alphas generated by our portfolios are not statistically significant and none of them surpasses the value of 0.55. Without considering this significance, HFCS portfolios perform substantially better than alpha and predictive alpha-sorted portfolios when the latter consists of the same hedge fund sample as our procedure. However, these results differ when we consider a fund sample that is not subject to any of the HFCS criteria. Indeed, the performance of our HFCS is then almost always outperformed by alpha-ranked strategies. This outcome questions the validity of the selection criteria established by the HFCS procedure. As our approach performs slightly better than alpha and predictive alpha-sorted portfolios on an identical set of initial funds, it is likely that it will approximate the good performance of alpha and predictive alpha-ranked portfolios considering all available alphas at the date analysed. This assumption is supported by the better performance of our Wide RW portfolio, considering funds having negative future alpha forecasts within the initial HFCS sample. Latter result also suggests that the eight-factor model of Fung and Hsieh (1997,2001,2004) fails at estimating alpha estimates as several hedge funds with negative alpha forecasts ended up improving the overall portfolio performance. Such alternative analysis should however naturally be applied on all four considered portfolios in order to confirm our hypothesis.

Several proposals may improve the HFCS procedure in order to obtain better and more significant results. First of all, the factor model choice has an impact on the time-varying alphas, and therefore naturally on the predictive alphas further used in our HFCS procedure. Furthermore, this model is also used to determine the final performance of our portfolios and models. As previously explained, less than 5% of U.S L/S equities hedge funds have a R squared higher than 0.7 through the regression with the eight-factor model of Fund and Hsieh, and slightly more than 20% of them have one over 0.5. Therefore, basing our performance assessment on a portfolio α retrieved thanks to this factor model may not be reliable and accurate. The choice of a factor model estimating time-varying alphas with a better goodness of fit would provide us with more accurate measures of alphas. Second, the approach used to determine the time-varying αs also highly impacts our study. We have decided to use two methods, namely the rolling window and the Kalman filter. These two methods differ from the approach used by GLTW who used both an unobserved signal perceived by hedge funds managers and funds' holdings in order to determine the funds' time-varying parameters. The replication of the GLTW method may then yield more insightful results. Finally, our methodology simply assigned equal weights to all hedge funds included within the HFCS. Alternative possibilities exist, such as the method of Treynor and Black (1973) as performed by GLTW. Naturally, with more time and computing power, the HFCS procedure can be replicated on a larger number of funds, following other strategies or domiciled in different geographical locations. Furthermore, a more comprehensive analysis, comparing multiple alternative significance levels λ and time-varying computation approaches could only improve the accuracy of our models' assessments.

As a last comment, we noticed throughout this paper that hedge funds behave very differently from mutual funds. We could particularly notice the latter with the very different distribution of their returns and alphas, which were considerably higher than the ones of mutual funds. As a result, large performance differences arise and a small amount of top performing hedge funds may have biased our analysis. We strongly believe that the implementation of this set-identification approach may produce performing results. It is however necessary to further assess the whole set of variables impacting the performance assessment of hedge funds.

Appendix A

Alpha and predictive alpha distributions: Kalman filter

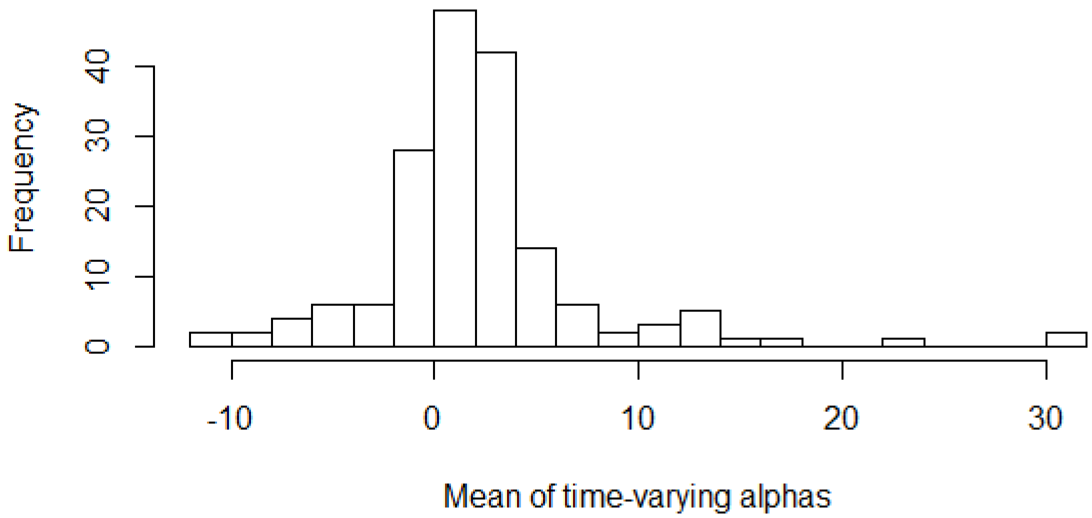

Figure A.1: Mean of time-varying alphas (Kalman filter)

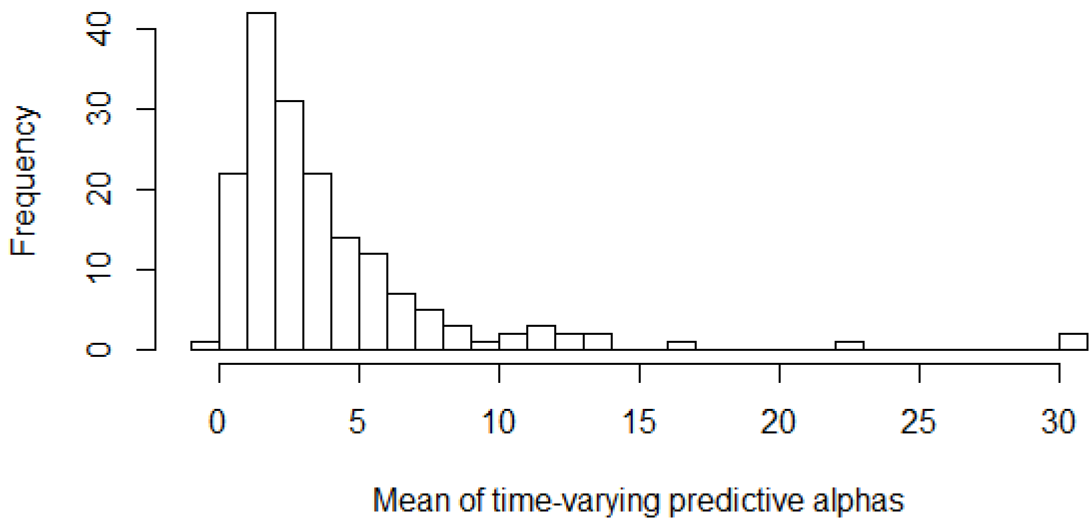

Figure A.2: Mean of time-varying predictive alphas (Kalman filter)

Appendix B

HFCS Example : Predictive alphas

	Alder Capital Partners I LP	Alpha Equity Global Long/Short Equity	Archer Equity Fund LLC
1	-6.22	5.71	5.39
2	-0.41	3.22	-3.80
3	3.34	6.34	3.38
4	0.54	4.37	-3.05
5	-0.39	4.71	-0.14
6	-9.80	6.36	-4.24
7	0.03	6.99	5.71
8	-4.98	9.71	13.77
9	4.47	6.25	8.70
10	18.71	9.85	3.07
11	18.59	5.47	-2.41
12	-1.00	7.24	-1.90
13	14.28	4.05	-1.99
14	17.45	5.08	12.05
15	13.58	8.38	1.67
16	23.41	6.67	14.49
17	20.32	6.81	19.33
18	17.97	5.46	22.47
19	4.34	0.88	13.20
20	-10.26	-0.48	4.73
21	3.54	6.97	-0.61
22	15.85	-2.60	-0.90
23	21.05	8.04	10.04
24	8.29	8.42	3.10

Table B.1: Predictive alphas from October 2008 to October 2010

Appendix C

HFCS Example : Predictive alphas differences

	Fund I - Fund II	Fund I - Fund III	Fund II - Fund I	Fund II - Fund III	Fund III - Fund I	Fund III - Fund II
1	-11.92	-11.61	11.92	0.32	11.61	-0.32
2	-2.81	-3.40	2.81	-0.58	3.40	0.58
3	-9.68	-6.72	9.68	2.96	6.72	-2.96
4	-3.83	3.59	3.83	7.42	-3.59	-7.42
5	-4.33	0.25	4.33	4.57	-0.25	-4.57
6	3.45	14.05	-3.45	10.60	-14.05	-10.60
7	-6.95	-5.67	6.95	1.28	5.67	-1.28
8	-14.69	-18.75	14.69	-4.06	18.75	4.06
9	-10.72	-13.17	10.72	-2.44	13.17	2.44
10	8.86	15.65	-8.86	6.78	-15.65	-6.78
11	13.12	21.00	-13.12	7.88	-21.00	-7.88
12	-8.24	0.90	8.24	9.14	-0.90	-9.14
13	10.22	16.27	-10.22	6.05	-16.27	-6.05
14	12.37	29.51	-12.37	17.13	-29.51	-17.13
15	5.20	11.91	-5.20	6.71	-11.91	-6.71
16	16.74	8.92	-16.74	-7.82	-8.92	7.82
17	13.51	0.99	-13.51	-12.52	-0.99	12.52
18	12.51	-4.49	-12.51	-17.01	4.49	17.01
19	3.46	-8.86	-3.46	-12.32	8.86	12.32
20	-9.77	-14.99	9.77	-5.22	14.99	5.22
21	-3.43	4.15	3.43	7.58	-4.15	-7.58
22	-13.25	-14.95	13.25	-1.70	14.95	1.70
23	-29.09	-31.08	29.09	-2.00	31.08	2.00
24	-0.12	5.20	0.12	5.32	-5.20	-5.32

Table C.1: Pairwise performance differences from October 2008 to October 2010

	Fund I - Fund II	Fund I - Fund III	Fund II - Fund I	Fund II - Fund III	Fund III - Fund I	Fund III - Fund II
1	-1.22	-0.06	1.22	1.17	0.06	-1.17

Table C.2: Pairwise mean of performance differences

Appendix D

Example: *'The Vilas Fund LP'*

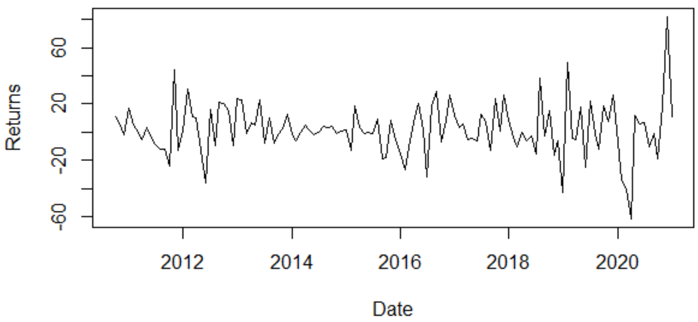

Figure D.1: 'The Vilas Fund LP' - Returns

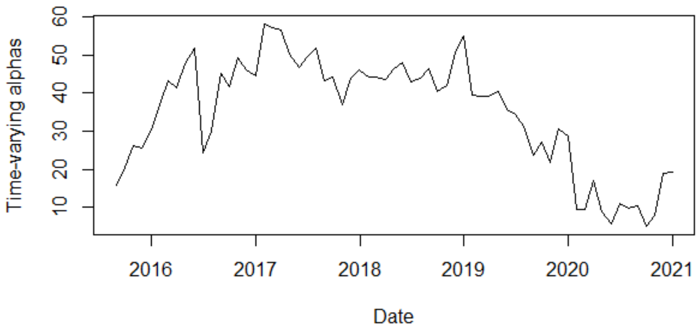

Figure D.2: 'The Vilas Fund LP' - Time-varying alphas

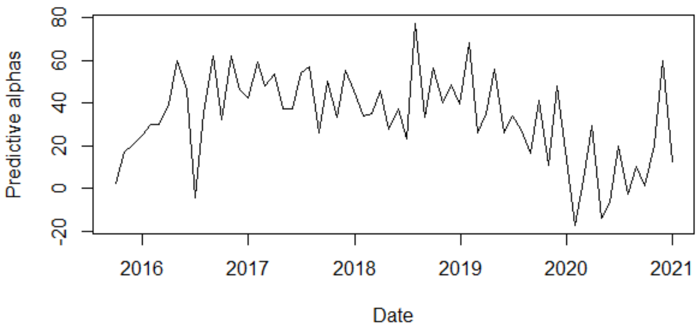

Figure D.3: 'The Vilas Fund LP' - Predictive alphas

Appendix E

Returns of HFCS Portfolios

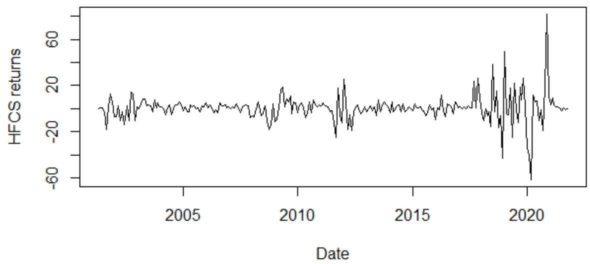

Figure E.1: HFCS Wide (RW approach) - Returns distribution

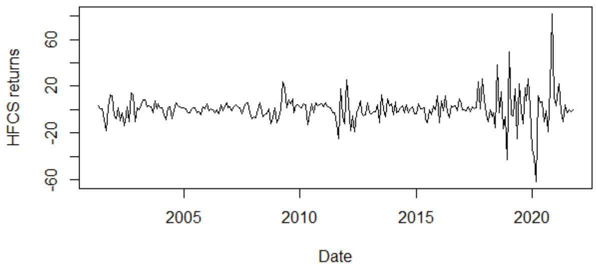

Figure E.2: HFCS Tight (RW approach) - Returns distribution

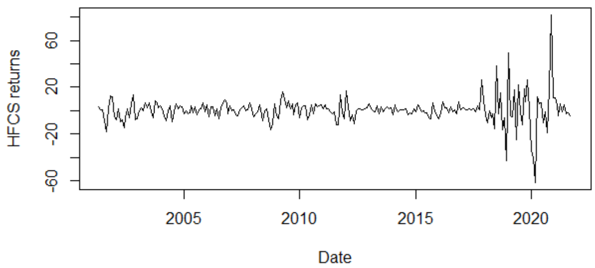

Figure E.3: HFCS Wide (KF approach) - Returns distribution

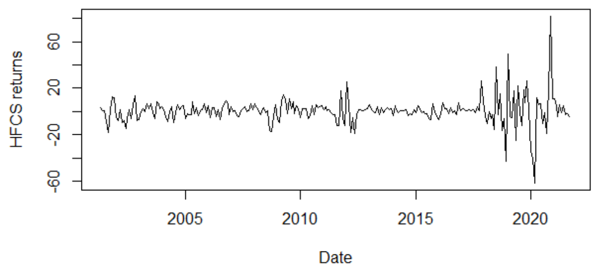

Figure E.4: HFCS Tight (KF approach) - Returns distribution

Appendix F

Performance of single-fund portfolios

HF in portfolio	Alpha	T-value
Rolling window		
Top 5% Pred. α	2.69	2.20
Top 3% Pred. α	3.90	2.28
Top 3 Pred. α	3.34	2.09
Top 2 Pred. α	3.68	1.77
Top 1 Pred. α	0.03	0.01

Table F.1: Top Predictive alpha-ranked portfolios

Appendix G

Alpha-sorted portfolios: Number of funds (Rolling window)

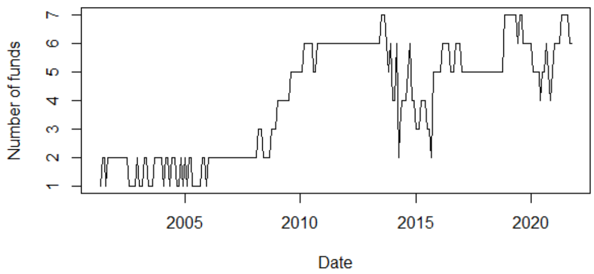

Figure G.1: Number of portfolio funds - Top 10%

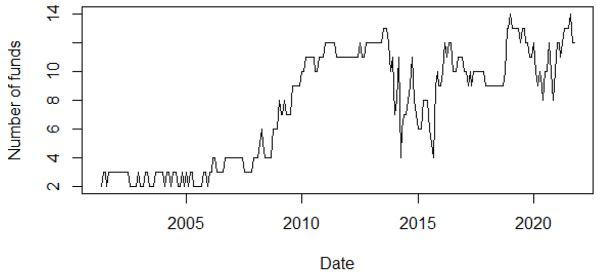

Figure G.2: Number of portfolio funds - Top 20%

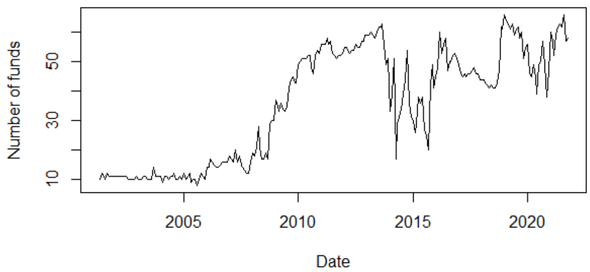

Figure G.3: Number of portfolio funds - Positive alphas

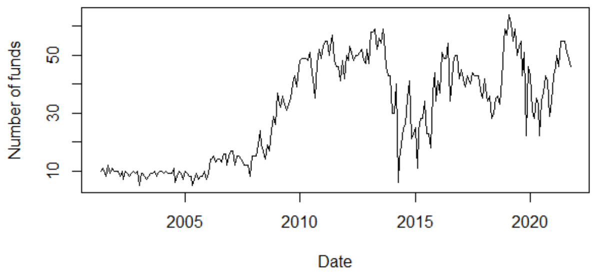

Figure G.4: Number of portfolio funds - Positive Predictive alphas

Appendix H

Alpha-sorted portfolios: Number of funds (Kalman Filter)

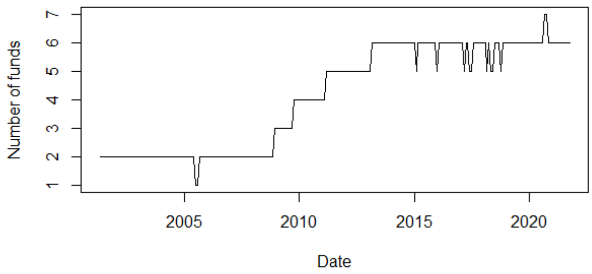

Figure H.1: Number of portfolio funds - Top 10%

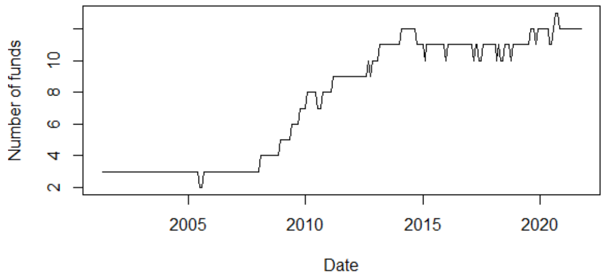

Figure H.2: Number of portfolio funds - Top 20%

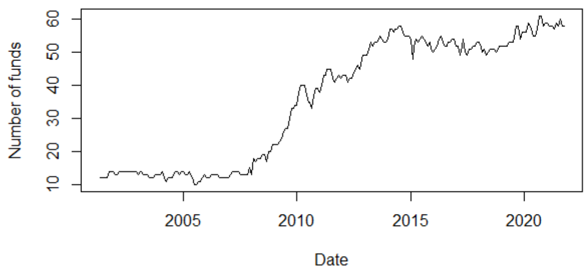

Figure H.3: Number of portfolio funds - Positive alphas

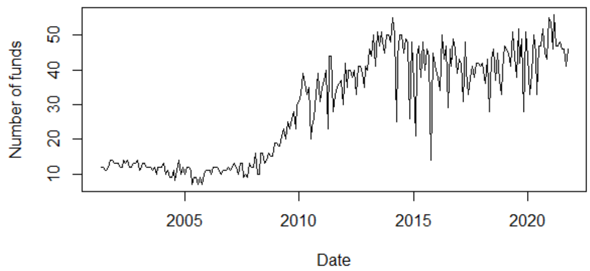

Figure H.4: Number of portfolio funds - Positive Predictive alphas

Bibliography

Ackermann, C., McEnally, R., and Ravenscraft, D. (1999). The performance of hedge funds: Risk, return, and incentives. *The Journal of Finance*, LIV(3):833–874.

Agarwal, V., Daniel, N. D., and Naik, N. Y. (2004). Flows, performance, and managerial incentives in hedge funds. https://ssrn.com/abstract=424369. Accessed: 2022-07-10.

Agarwal, V., Daniel, N. D., and Naik, N. Y. (2009). Role of managerial incentives and discretion in hedge fund performance. *The Journal of Finance*, 64(5):2221–2256.

Agarwal, V. and Naik, N. Y. (2000a). Multi-period performance persistence analysis of hedge funds. *The Journal of Financial and Quantitative Analysis*, 35(3):327–342.

Agarwal, V. and Naik, N. Y. (2000b). On taking the 'alternative' route: Risks, rewards style and performance persistence of hedge funds. *The Journal of Alternative Investments*, 2(4):6–23.

Amenc, N., Curtis, S., and Martellini, L. (2003). The alpha and omega of hedge fund performance measurement. http://performance-measurement.org/AmencCurtisMartellini2003.pdf. Accessed: July 2022.

Amenc, N., Martellini, L., and Meyfredi, J.-C. (2010). Passive hedge fund replication – beyond the linear case. *European Financial Management*, 16(2):191–210.

Ammann, M., Huber, O., and Schmid, M. (2013). Hedge fund characteristics and performance persistence. *European Financial Management*, 19(2):209–250.

Ammann, M. and Moerth, P. (2005). Impact of fund size on hedge fund performance. *Journal of Asset Management*, 6:219–238.

Aragon, G. O. (2007). Share restrictions and asset pricing: Evidence from the hedge fund industry. *Journal of Financial Economics*, 83(1):33–58.

Bali, T. G., Brown, S. J., and Demirtas, K. O. (2013). Do hedge funds outperform stocks and bonds? *Management Science*, 59(8):1887–1903.

Barras, L., Scaillet, O., and Wermers, R. (2010). False discoveries in mutual fund performance: Measuring luck in estimated alphas. *The Journal of Finance*, 65(1):179–216.

Barès, P.-A., Gibson, R., and Gyger, S. (2003). Performance in the hedge funds industry: An analysis of short- and long-term persistence. *The Journal of Alternative Investments*, 6(3):25–41.

Brealey, R. and Kaplanis, E. (2001). Changes in the factor exposures of hedge funds. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.197.792&rep=rep1&type=pdf. Working paper. Accessed: 2022-07-10.

Brunnermeier, M. K. and Nagel, S. (2004). Hedge funds and the technology bubble. *The Journal of Finance*, 59(5):2013–2040.

Cai, L. and Liang, B. (2012). On the dynamics of hedge fund strategies. *The Journal of Alternative Investments*, 14(4):51–68.

Catania, L. and Bernardi, M. (2014). Mcs: Model confidence set procedure. `https://cran.r-project.org/web/packages/MCS/MCS.pdf`. R package version 0.1.3.

Chen, Y., Cliff, M., and Zhao, H. (2017). Hedge funds: The good, the bad, and the lucky. *Journal of Financial and Quantitative Analysis*, 52(3):1081–1109.

Connor, G. and Lasarte, T. (2003). An overview of hedge fund strategies.

Connor, G. and Woo, M. (2004). An introduction to hedge funds. `http://eprints.lse.ac.uk/24675/1/dp477.pdf`. Accessed: 2022-07-10.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance*, 52(3):1035–1058.

Eling, M. (2009). Does hedge fund performance persist? overview and new empirical evidence. *European Financial Management*, 15(2):362–401.

Frumkin, D. and Vandegrift, D. (2009). The effect of size, age, beta and disclosure requirements on hedge fund performance. *Journal of Derivatives & Hedge Funds*, 15(3):241–251.

Fung, W. and Hsieh, D. A. (1997). Empirical characteristics of dynamic trading strategies: The case of hedge funds. *The Review of Financial Studies*, 10(2):275–302.

Fung, W. and Hsieh, D. A. (2000). Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases. *The Journal of Financial and Quantitative Analysis*, 35(3):291–307.

Fung, W. and Hsieh, D. A. (2001). The risk in hedge fund strategies: Theory and evidence from trend followers. *The Review of Financial Studies*, 14(2):313–341.

Fung, W. and Hsieh, D. A. (2004). Hedge fund benchmarks: A risk-based approach. *Financial Analysts Journal*, 60(5):65–80.

González, M. D. O., Papageorgiou, N. A., and Skinner, F. S. (2016). Persistent doubt: An examination of hedge fund performance. *European Financial Management*, 22(4):613–639.

Gronborg, N. S., Lunde, A., Timmermann, A., and Wermers, R. (2021). Picking funds with confidence. *Journal of Financial Economics*, 139(1):1–28.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Harri, A. and Brorsen, B. W. (2004). Performance persistence and the source of returns for hedge funds. *Applied Financial Economics*, 14(2):131–141.

Helleiner, E. and Pagliari, S. (2009). The end of self-regulation? hedge funds and derivatives in global financial governance. In Helleiner, E., Pagliari, S., and Zimmerman, H., editors, *Global Finance in Crisis: The Politics of International Regulatory Change*. Taylor & Francis.

Jagannathan, R., Malakhov, A., and Novikov, D. (2010). Do hot hands exist among hedge fund managers? an empirical evaluation. *The Journal of Finance*, LXV(1):217–255.

Joenvääräa, J., Kosowskib, R., and Tolonen, P. (2012). New 'stylized facts' about hedge funds and database selection bias. `https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.295.1960&rep=rep1&type=pdf`. Accessed: 2022-07-10.

Jones, M. (2013). Examination of fund age and size and its impact on hedge fund performance. *Derivatives Use, Trading & Regulation*, 12(4):342–350.

Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006). Can mutual fund "stars" really pick stocks? new evidence from a bootstrap analysis. *The Journal of Finance*, 61(6):2551–2595.

Kosowski, R., Y.Naik, N., and Teo, M. (2007). Do hedge funds deliver alpha? a bayesian and bootstrap analysis. *Journal of Financial Economics*, 84(1):229–264.

Liang, B. (1999). On the performance of hedge funds. *Financial Analysts Journal*, 55(4):72–85.

Mamaysky, H., Spiegel, M., and Zhang, H. (2007). Improved forecasting of mutual fund alphas and betas. *Review of Finance*, 11(3):359–400.

Mamaysky, H., Spiegel, M., and Zhang, H. (2008). Estimating the dynamics of mutual fund alphas and betas. *The Review of Financial Studies*, 21(1):233–264.

Mitra, S. (2009). An introduction to hedge funds. `https://arxiv.org/pdf/0904.2731.pdf`. White paper. Accessed: 2022-07-10.

Preqin (n.d.). Hedge fund strategies. `https://www.preqin.com/academy/lesson-3-hedge-funds/hedge-fund-strategies`. Accessed: 2022-07-10.

Racicot, F.-E. and Théoret, R. (2013). The procyclicality of hedge fund alpha and beta. *Journal of Derivatives & Hedge Funds*, 19:109–128.

Securities and Commission, E. (2022). Private funds statistics : Third calendar quarter 2021. `https://www.sec.gov/divisions/investment/private-funds-statistics/private-funds-statistics-2021-q3.pdf`. Accessed: 2022-07-10.

Shi, Z. (2017). The impact of portfolio disclosure on hedge fund performance. *Journal of Financial Economics*, 126(1):36–53.

Stafylas, D., Anderson, K., and Uddin, M. B. (2016). Recent advances in hedge funds' performance attribution: Performance persistence and fundamental factors. *International Review of Financial Analysis*, 43:48–61.

Statista (n.d.). Hedge funds - statistics facts. `https://www.statista.com/topics/5064/hedge-funds/#dossierContents__outerWrapper`. Accessed: 2022-07-10.

Teo, M. (2009). Does size matter in the hedge fund industry? `https://ink.library.smu.edu.sg/lkcsb_research/5174`. Accessed: 2022-07-10.

Titman, S. and Tiu, C. (2010). Do the best hedge funds hedge? *The Review of Financial Studies*, 24(1):123–168.

Treynor, J. L. and Black, F. (1973). How to use security analysis to improve portfolio selection. *The Journal of Business*, 46(1):66–86.

van Vuuren, G. and Yacumakis, R. (2015). Hedge fund performance evaluation using the kalman filter. *Studies in Economics and Econometrics*, 39(3):1–24.

White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5):1097–1126.

Executive Summary

Despite decades of increasing popularity and various academic studies, the forecast of hedge funds future performance remains an unresolved challenge. This end-of-master project in Financial Engineering focuses on the selection of successful and superior hedge funds. Inspired by the work of Groborg et al (2021) tackling the performance of mutual funds, we present an approach that enables to determine the number and the identity of outperforming hedge funds throughout the time-period analysed, by conducting multiple efficient pairwise comparisons.

Our study compares U.S. hedge funds that pursue a Long/Short equity strategy. The latter are assessed by their risk-adjusted returns, referred to as the funds' *alphas*, using the eight-factor model of Fung and Hsieh (1997, 2001, 2004). Time-varying alphas are estimated through two main approaches, the first of which is a 60-month rolling window and the second is a Kalman filter estimating the parameters of a pure recursive state space model. In order to ensure that alpha estimates accurately predict the hedge funds' performances, the *Predictive alphas* of the funds are computed as well.

Our goal is to construct a set of superior hedge funds by performing multiple pairwise tests that iteratively eliminate inferior funds from our initial sample. The procedure stops when the performance of the remaining funds is considered equal, with a given level of confidence. The final set of expected superior hedge funds is called the Hedge Fund Confidence Set, denoted *HFCS*.

The first objective of this project is the application of the procedure implemented by Gronborg et al. (2021) to hedge funds, considering the modifications that it entails. Our second goal is the assessment of latter model. This evaluation will be conducted by collecting returns generated by hedge funds contained in the HFCS. Consequently, our model is assessed through the alpha of the constructed portfolio, which can be estimated using the same factor model as mentioned here above.

Our findings suggest that the implemented procedure generates positive abnormal returns regardless of the confidence level and the calculation used to determine the dynamic alphas. However, none of these show statistically significant results. In addition, the HFCS procedure outperforms more conventional alpha-sorting methods. Nevertheless, this result applies only if the sample of funds considered and analysed is identical. Indeed, we note that the hedge fund selection criteria of our HFCS procedure tend to diminish the potential performance of the superior portfolios. The improvement of our best performing portfolio by the removal of one of the criteria established by the HFCS procedure strengthens this statement.

We believe that the implementation of this pairwise comparison approach can be accurately applied to hedge funds and that it is able to generate performing results. Nonetheless, this high performance is only possible by conducting a comprehensive study of several variables, such as the choice of the asset pricing model, the significance level chosen for the pairwise t-tests, the approach for the estimation of time-varying alphas and the criteria on which the initial set of hedge funds is based.

Number of words: 12,000