What factors influence the likelihood of exceptionally high performance in the hedge fund industry? An extreme value approach of the hedge fund right tail

Auteur : Schillings, François
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 : 2023-2024
URI/URL : http://hdl.handle.net/2268.2/21293

WHAT FACTORS INFLUENCE THE LIKELIHOOD OF EXCEPTIONALLY HIGH PERFORMANCE IN THE HEDGE FUND INDUSTRY?

AN EXTREME VALUE APPROACH OF THE HEDGE FUND RIGHT TAIL

Jury :
Promoter :
Julien Hambuckers
Reader :
Philippe Hübner

Submitted by
François SCHILLINGS
With a view to obtaining master's degree in management engineering, specialising in Financial Engineering
Academic year 2023/2024

Executive summary

This work explores the factors that influence the likelihood of exceptionally high performance in the hedge fund industry using an extreme value approach. Hedge funds, known for their complex strategies and high risk-reward profiles, have historically attracted attention due to their ability to generate outsized returns. Despite extensive research on hedge funds, little focus has been placed on understanding the right tail of their performance distribution—the occurrence of extreme positive returns.

The study addresses this gap by examining the macroeconomic and structural variables that affect the likelihood of these extraordinary profits. By applying Extreme Value Theory (EVT) to a carefully cleaned dataset, the research seeks to establish the conditions under which hedge funds are more likely to achieve abnormal returns. The dataset includes a wide range of hedge fund characteristics as well as external financial and economic variables, ensuring a comprehensive analysis.

The findings are significant for both academic and practical purposes. Hedge fund managers can leverage insights from this research to refine their investment strategies and potentially enhance their funds' performance. Simultaneously, the study contributes to the development of models that can better predict extreme returns, offering more precise tools for future investment forecasts.

Moreover, the research highlights a gap in the existing literature: while many studies explore hedge funds' performance relative to other asset classes, few have isolated the factors that drive extreme positive outcomes. This study fills that void by focusing on the right tail of the performance distribution.

This work concludes by offering key insights and recommendations for hedge fund managers, while also addressing the limitations of the analysis. Future research is suggested to build on these findings, particularly in terms of improving predictive models and understanding the underlying dynamics of extreme returns.

1738 characters

Acknowledgements

Firstly, I would like to extend my sincere thanks to Professor Julien Hambuckers for his academic guidance throughout this work and during my last two years at HEC Liège.
Through his courses, Professor Hambuckers played a crucial role in sparking my interest in the financial sector.

I would also like to express my deep gratitude to Mr. Philippe Hübner, without whom this work would not have been possible.
His theoretical knowledge and practical expertise form the foundation of this work.

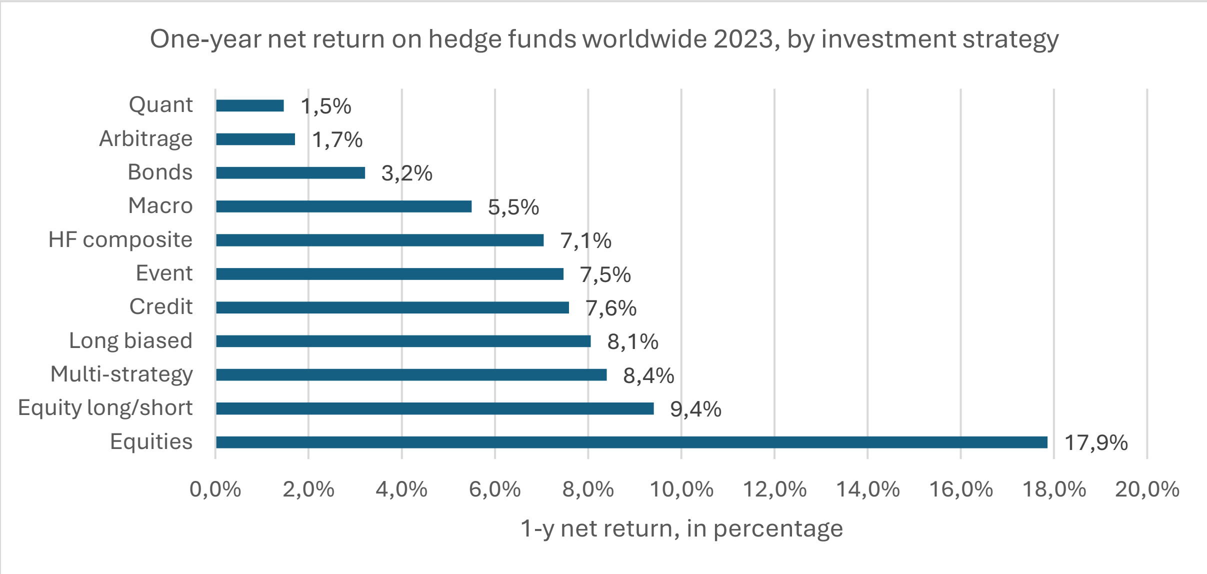
Lastly, I wish to dedicate this work to my godfather, who has always taken a special interest in my academic journey.

Contents

Introduction

In this work, we will examine the hedge fund industry and its performance. The first traces of these speculative funds date back to 1949, and since then, they have gained increasing notoriety. Their investment strategies, exotic characteristics, staggering profits, and catastrophic collapses have all contributed to the reputation of these investment funds. Despite recently celebrating their 75th anniversary, hedge funds continue to arouse the curiosity of many researchers due to their complexity and unique traits. Indeed, they have been the subject of much debate for several years, and despite hundreds of studies on the topic, no one can claim to fully understand the dynamics of these financial vehicles.

One particular issue, however, will attract our attention: that of industry performance. In 2023, some classes of hedge funds will post extremely high positive results, as the graph below shows.



Such performances inevitably raise a question:

« **What factors influence the likelihood of exceptionally high performance in the hedge fund industry?»**

The objective of this work is to identify the variables that significantly impact the right tail of the return distribution generated by hedge funds. In other words, we will attempt to assess the influence of the macroeconomic environment and fund structure on the probability of abnormally high profits.

Answering this question will have significant practical and theoretical implications. For hedge fund managers, a better understanding of the conditions that favor exceptionally high returns will enable them to adjust and optimize their investment strategies. At the same time, this study will contribute to the development of explanatory and predictive models for hedge fund returns, providing more accurate tools for anticipating future performance.

Furthermore, as we will demonstrate in our literature review, few studies exclusively address the issue of positive returns. Much research has been conducted on hedge fund performance relative to other asset classes, or on factors explaining the overall distribution of returns, but very little focuses on the right tail of hedge fund return distributions.

To answer our research question, several factors must be considered. First, it is essential to understand the theoretical context in which we are operating to grasp the intricacies of our model. Second, we need to identify a theorem that links the set of variables under consideration to the distribution of extremes. Once this framework is established, we must select a coherent data set, but determining the appropriate one poses a challenge. Lastly, after addressing the practical considerations for implementing our procedure, we should be able to obtain consistent results. We will proceed methodically to address these issues.

We will begin by setting the stage through a comprehensive literature review, defining key concepts. This will allow us to better understand the unique characteristics of hedge funds, identify variables that impact fund performance, and justify the relevance of this study. Next, we will introduce Extreme Value Theory (EVT), which will serve as the foundation of our model, enabling us to estimate the distribution of extremes. We will examine the theoretical underpinnings of EVT and assess its applicability to our research context.

Implementing the methodology will involve addressing several practical considerations, which we will discuss in detail. Once the methodology is clear, we will define the dataset on which our analysis will be based. Several restrictions, including temporal, geographical, and qualitative constraints, will be applied to the initial dataset to enhance the efficiency of our analysis. These restrictions will be elaborated on and justified in the relevant section.

Finally, we will apply our model to the refined dataset, generating results that will be carefully analyzed to draw meaningful conclusions. These results will then be interpreted in depth, with the aim of gaining key insights into the issues at hand. The study will conclude with a comprehensive summary of these findings, emphasizing the contributions they make to our understanding of hedge fund performance. We will also discuss the limitations of our analysis and suggest areas for improvement or further research. In this way, this study aims to provide a comprehensive and nuanced understanding of the impact and implications of our findings.

Literature Review

Slightly over a quarter-century ago, the hedge fund "Long-Term Capital Management" had to be rescued from an inevitable collapse that could have triggered a meltdown of the financial system (Capocci, 2013). Despite this, the hedge fund industry has seen significant growth over the past decades and is now valued at $2.944 trillion (Aurum, 2023). This remarkable growth, along with the eccentric performances exhibited by these unconventional investment vehicles, has attracted the attention of the scientific community, placing the hedge fund industry in the academic spotlight. Our work aims to contribute to this growing body of research by identifying the key factors behind the occurrence of abnormally positive returns.

To fully appreciate the scope and significance of our research, it is important to first review the progress made by the academic community in explaining hedge fund performance. The following discussion will provide a concise summary of the major findings on this topic.

I. What is a hedge fund?

Definition

The literature struggles to establish a comprehensive definition, as noted by Nabilou (2017) and Lhabitant (2006). In fact, no universally accepted definition exists due to the diversity of these investment vehicles. However, for the sake of clarity in this work, it is essential to adopt a common definition. While not perfect, we will use the definition provided by Connor & Woo (2004), which characterizes hedge funds as actively managed investment funds open only to select investors. This definition carries several implications that will be discussed in the following sections.

Investors profile

Historically, hedge funds work in junction with only private “sophisticated” wealthy investors and target “High Net Worth Individuals” (*HNWI*). Dealing only with non-public entities, allowed them to exploit legal loopholes and consequently avoid controls of public financial entities. This lack of regulation enables them to use complex and riskier strategies (Fung and Hsieh, 1999 ; Cumming et al., 2013). Nowadays, diversity among hedge funds investors has increased (Capocci, 2013) even though the need for privacy and regulatory-free environment still key (Fung and Hsieh, 1999). To avoid inappropriate investors, a substantial minimum investment amount, varying from one fund to another, is required to enter in the investment pool (De Luca, 2023). The investors’ profiles and the consequent minimum investment cap are one of the characteristics differentiating hedge funds from other investment pools.

Hedge Funds strategies

In the definition previously mentioned, the active nature of hedge funds’ management is also clearly specified. This dynamic approach translates into the implementation of a wide range of strategies. Once more, even though the subject is extensively discussed, there is no clear classification of the investment strategies.

Strategies’ classification

One might classify hedge fund strategies based on those self-reported by the funds themselves. Hedge Fund Research (HFR) defines seven classes of strategies: equity hedge, event-driven, macro, relative value, risk parity, and blockchain (Hedge Fund Research, 2024). A prominent data provider, MAR/Hedge, also acknowledges seven similar categories: event-driven, global, global/macro, market neutral, sectors, short sellers, and long-only (Jansen et al., 1998). These are just a few of the many possible classifications. Even though this approach is straightforward, it heavily relies on the honesty of fund managers, and it is difficult to determine whether a fund truly follows the reported strategy.

In their paper, Fung & Hsieh (1999) deviate from this qualitative characterization of investment strategies. Instead, they use a quantitative approach by deriving different investment strategies from return distributions. This is achieved through principal components analysis. By analyzing the correlation between the principal components and the reported strategies, they were able to identify four sub-strategies: trend-following, global or macro, long-only, and distressed securities strategies (Fung & Hsieh, 1997).

Another perspective on the issue is provided by Capocci (2013). He relies on the relationship between hedge funds and financial markets and categorizes the strategies into two groups. An investment style is considered non-directional if it is uncorrelated with the financial market, whereas styles correlated with the markets are considered directional.

As we can see, there is no consensus on hedge fund management styles and how they should be characterized. Nevertheless, the literature implicitly agrees that these styles have an impact on performance. Lhabitant (2006) further emphasizes the differences in return profiles across strategies. Given the significance of this impact, we should account for these disparities in our work.

Well known strategies

As stated, the different approaches we have just mentioned are only some of the possibilities. However, among all the strategy classification methods, five types are often emphasized: equity hedge or long-short, event-driven, macro, fixed income, and managed futures (Baker & Filbeck, 2017; Capocci, 2013; Fung & Hsieh, 1999; Hedge Fund Research, 2024; Eurekahedge, n.d.). Since these are widely used across the hedge fund industry, we will briefly review them and their characteristics.

Firstly, the equity hedge, or long-short, strategies involve constructing a portfolio that mixes long and short positions on equities. It has become one of the most widely used strategies in the non-traditional investment sector (Lhabitant, 2006). The aim is to achieve an overall position that is uncorrelated with the market, or market neutral. The performance of these strategies therefore depends on the hedge fund manager's ability to select optimal stocks for entering long and short positions (Auleta & Stefanini, 2017). This neutral exposure theoretically allows for profit generation in both bullish and bearish conditions. However, the use of hedging techniques increases the risk of losses if the manager's investment decisions turn out to be poor (Lhabitant, 2006).

When hedge fund managers seek to profit from price inefficiencies caused by corporate events, they follow an event-driven strategy (Farrelly & Lhabitant, 2017). The spectrum of corporate events here includes various situations such as mergers and acquisitions, spin-offs, financial distress, etc. (Bali et al., 2013). There are different kinds of event-driven strategies, namely distressed debt, which involves investing in the debt of financially struggling firms (Anson, 2002); arbitrage, which takes advantage of equity price discrepancies due to corporate events (Cao et al., 2016); and multi-strategies, which combine both equity and debt investing strategies (Metzger & Shenai, 2019).

The global or macro strategy relies on in-depth macroeconomic analysis to optimize asset allocations (Fung & Hsieh, 1999). More precisely, it involves analyzing macroeconomic imbalances and trends (Lhabitant, 2006) in order to enter long unhedged leveraged positions on various assets (Longo, 2022). Typically, macro managers will follow one of three approaches: a feedback-based approach, grounded in market psychology; a model-based approach, using sophisticated macroeconomic models; or an information-based approach, relying on microeconomic data analysis (Lhabitant, 2006).

Another strategy that hedge fund managers might pursue is fixed income arbitrage. As the name suggests, this strategy seeks to exploit price discrepancies between different types of fixed income assets (Anson, 2002). Such a strategy is possible due to the market structure in which the trades are made. Indeed, the absence of consensus on a standard absolute pricing model, the presence of various relative pricing relationships among different fixed income instruments, the impact of irrational but predictable supply and demand on specific assets' prices, and the intricate nature of certain fixed income securities make the market favorable for arbitrage opportunities (Lhabitant, 2006).

Finally, managed futures strategies, also known as commodity trading advisors (CTA), use complex quantitative models based on fundamental and econometric data (Capocci, 2013) to trade futures contracts on commodities, currencies, and financial markets. Their price-trend-following strategy differentiates them from macro strategies, which base their investments on macro analysis rather than prices (Pozen & Hamacher, 2015).

Obviously, regardless of the strategy adopted, hedge fund managers share a common objective: optimizing their fund's performance. Consequently, the fund's returns are naturally correlated with the strategy adopted (Lambert, 2012). From this apparent observation, a question arises: what does the literature tell us about hedge fund performance?

II. The performance of the hedge fund industry

Assessing the performance of such complex financial vehicles has always been challenging. Indeed, their intricate structure, exotic strategies, and reporting biases make performance assessment extremely difficult. Despite this apparent complexity, many scientific efforts have been deployed to characterize the performance of these speculative investment pools. One of the many questions raised is how hedge fund returns compare with those of other assets. The literature has attempted to determine whether, on a risk-adjusted basis, hedge funds significantly outperform other asset classes, such as mutual funds. The performance of the hedge fund industry has always been a subject of controversy within the scientific community, giving rise to an abundance of differing opinions. To better understand the nuances of these debates, we will review the scientific literature on the topic.

The data biases

Obviously, assessing the performance of hedge funds, or any other asset, will be based on their past performances, which are reported in various databases. Nevertheless, hedge fund databases are often characterized by inherent biases that distort performance measurements.

One type of bias is the selection bias, which stems from the requirement that hedge funds offer their services exclusively to private investors. These selection criteria limit the ability of hedge funds to attract new investors, as they are restricted from advertising. Consequently, the only way to reach new capital providers is by being listed in established databases (Fung et al., 2008). However, only the top-performing funds are likely to seek inclusion, meaning that databases capture only efficient actors, thereby overestimating market performance and creating selection bias (Fung & Hsieh, 2000). Although targeting private investors has undeniable advantages, it compromises database reliability and affects the scientific literature on the subject.

Another significant distortion is survivorship bias, which arises from the non-reporting of liquidated funds, resulting in a skewed average performance (Liang, 2000). Other existing biases include liquidity bias, caused by the omission of significant negative returns just before a fund's liquidation (Fung & Hsieh, 2006), and instant history bias, due to the overstatement of hedge fund performance prior to being included in the database (Capocci et al., 2005). These biases distort the reported performances and must be corrected or at least acknowledged before conducting any performance analysis.

Additionally, the valuation of assets under management can also bias fund performance. Hedge funds often invest in over-the-counter vehicles, which are difficult to price (Stulz, 2007). The estimation of asset prices, also known as the return smoothing process, may misrepresent a fund's risk exposure, biasing its estimated volatility (Huang et al., 2018) and inducing autocorrelation within the return time series (Bollen & Pool, 2008). Although these effects may be difficult to quantify, they should not be overlooked.

Before applying our methodology to the database, we must therefore consider all these biases to ensure the accuracy of our data.

The statistical profile of hedge funds returns

Those data biases being discussed, we will shift our attention to the funds' actual performances. However, their assessment might be blurred due to the complexity of these investment pools. Indeed, the exotic features displayed by these investment vehicles significantly transform the statistical structure of the returns. To better understand these, we will quickly review hedge funds' return statistical properties.

Distribution of the hedge funds returns

n finance, a classical assumption is that returns are normally, identically, and independently distributed. However, the reality is far from this basic assumption. Hedge funds are a prime example (Savona, 2016; Karagiorgis & Drakos, 2022; Amin & Kat, 2003). Among others, Amunderu and Murahwa (2021) have tested the normality assumption using the Jarque-Bera test, which evaluates the third and fourth moments of the distribution. The results are undisputable: among the 11 hedge fund indexes tested, none followed a Gaussian distribution, which is consistent with existing literature findings.

Tails of the distribution

Research into the non-Gaussian nature of hedge fund return distributions is complemented by studies by Kat and Brooks (2001) and Capocci (2013), which demonstrate the negative skewness and positive excess kurtosis of hedge fund returns. These characteristics translate into a fat-tailed distribution, meaning that extreme events are more likely. The authors justify these findings by arguing that hedge funds are characterized by high-order moments due to their greater risk exposure.

Since hedge funds' distributions are fat-tailed, Agarwal et al. (2017) investigate its determinants. By defining a left-tail risk measure, they show that investment in tail-sensitive stocks and options generates tail risk, while leverage and liquidity shocks impact tail risk. Additionally, Shin et al. (2018) highlighted the ability of hedge fund managers to time this tail risk exposure, allowing them to optimize their risk-adjusted returns.

As we can see, hedge fund tail risk is a well-known topic within the literature. However, we must underscore the almost nonexistent research on the hedge funds' right tail. My work will contribute to filling this gap in the literature.

Autocorrelation and correlation

Having rejected the Gaussian hypothesis, we can also reject the independence assumption, as hedge fund return time series are characterized by serial autocorrelation, as shown by the literature. Different hypotheses have been proposed regarding its sources. Some argue that it emerges from the exploitation of market inefficiencies. However, this seems improbable, as the exploitation of such inefficiencies would systematically eliminate them (Samuelson, 2015). A more likely theory, proposed by Getmansky et al. (2004), suggests that autocorrelation arises from the illiquidity of hedge funds' portfolio assets and return smoothing.

Another interesting characteristic of hedge fund returns is their low correlation with other asset classes. Many papers highlight this low correlation inherent to these funds (Fung & Hsieh, 1997). This property is key for investors, as it allows them to diversify their portfolios by including hedge funds, thereby improving the risk-return profile of their portfolios (Schneeweis & Spurgin, 1998).

Risk-adjusted performance measures

Once we have considered the data biases and the statistical features of the returns' distributions, we will assess the effectiveness of classical risk-performance metrics. Among the most commonly used are the Sharpe Ratio, Jensen's Alpha, and the Treynor Ratio (McMillan et al., 2011). However, their effectiveness can be obscured by the hedge funds' distributional characteristics (Van & Duong, 2022; Lo, 2002). Furthermore, the incentive fee structure and the application of high watermark (Panageas & Westerfield, 2009) create asymmetric payoffs (Lambert, 2012), similar to options' payoffs (Agarwal & Naik, 2003). This asymmetry presents another challenge for performance measurement.

Jensen's Alpha and the Treynor Ratio both derive from the Capital Asset Pricing Model (Maginn et al., 2007) and therefore assume normality and linear dependency between asset prices and risk premia (Amin & Kat, 2003). Since hedge fund returns are neither normally distributed (Amunderu and Murahwa, 2021) nor linearly related to risk factors (Fung & Hsieh, 1997), these metrics struggle to accurately reflect the risk-adjusted performance of the funds.

Even though the Sharpe Ratio does not have the same theoretical foundations as the metrics mentioned earlier, it still has drawbacks. The primary issue is the requirement for an equivalent distribution structure between the assets being compared, except for location and scale parameters (Schuhmacher & Eling, 2011). This assumption seems unlikely given the complexity of hedge fund statistical characteristics. Furthermore, due to the non-Gaussian distributional structure, the standard deviation fails to capture the overall risk, thereby impacting the Sharpe Ratio's effectiveness (Smith, 2017).

The multi-factors models

Return decomposition

In parallel with traditional measurements, researchers have developed multi-factor models. All these models have a common denominator: the decomposition of returns into two different components: beta and alpha.

When a hedge fund manager exposes themselves to systematic risk, the market rewards them with a premium. In this context, beta quantifies this exposure and determines the proportion of the risk premium that the manager should receive for the risk taken (Lhabitant, 2006). Throughout the literature, we may find models with either fixed or time-varying loading factors (Lambert, 2012).

On the other hand, alpha, or specific risk (Lhabitant, 2006), represents the excess return generated by the fund relative to a certain benchmark (Hossain et al., 2017). In other words, it is the portion of returns that cannot be explained by exposure to various risk factors encompassed in the model (Capocci, 2013). In the hedge fund industry, alpha is often used as a proxy for manager skill (Bui & Ganguly, 2017).

Such an approach to performance assessment is relevant to our work as it allows us to isolate abnormal returns, which is the quantity of interest in our investigation, and to highlight influencing factors that we will consider in the model. Since this approach is pertinent to our context, we will review the literature on this topic.

The models

The return decomposition principle has led to numerous scientific papers. Depending on the number and quality of the risk factors considered, models can address various aspects of hedge funds' atypical structures and returns distributions. Agarwal and Naik (2003) established a model combining the Fama and French (1993) size factors, the Carhart (1997) momentum factor, risk factors modeling the debt and equity markets, as well as an "option" factor that accounts for the non-linear payoffs exhibited by hedge funds. This latter factor is constructed through portfolios of monthly alternative long and short positions on call and put European at-the-money and out-of-the-money options on the S&P 500. These portfolios model the returns of call and put options on this index and thus include them in the return decomposition.

Other multifactor models have been developed, incorporating various risk factors. Harvey and Siddique (2000) extended the Fama and French (1998) three-factor model to account for the effect of skewness on returns. Another noteworthy model was proposed by Capocci et al. (2004). They constructed a modified version combining the models established by Fama and French (1998), Carhart (1997), and Agarwal and Naik (2002), adding an innovative factor that accounts for the fund's exposure to emerging bond markets.

Bali et al. (2014) also developed a model incorporating macroeconomic factors such as annual GDP growth, the unemployment rate, and the inflation rate, demonstrating the explanatory power of macroeconomic variables on the cross-sectional differences in hedge fund returns. Lambert and Platania (2020) further explored dynamic betas, conditioned on macroeconomic conditions, and showed their influence on asset allocation across different funds.

Performance persistence

Even though some argues that hedge funds do not outperform market indices (Ackermann et al., 1999), Agarwal et al. (2015) reviewed the scientific literature available on the topic and concluded that, in aggregate, hedge funds outperform mutual funds and other class of assets. In addition to the question on hedge funds performance, a massive strand of the literature focused on its persistence.

One of the first paper on the topic has been written by Brown and al. (1999). They investigate the performance of non-US hedge funds between 1989 and 1995 and have concluded the absence of any significant persistence in their return's distribution. Nevertheless, Agarwal and Naik (2000) challenge those findings by assessing a larger sample and indicate strong hints of short-term persistence, whatever strategy adopted. Using an extension of the Carhartt (1997) and Fama and French (1993) models, Capocci et al. (2005) have been able to demonstrate a high-performance persistence within the mid-performing funds.

According to Agarwal and Naik (2000), the persistence isn't statically significant anymore over the three months, Ammann et al. (2010), Jagannathan et al. (2010) and Ibbotson and al. (2011) went against this conclusion and demonstrate yearly alpha persistence. Those papers are based on more sophisticated econometric methods allowing them to challenge the historical assumptions of short-term performance persistence. However, further investigations must be carried to endorse those conclusions (Stafylas et al., 2016). But the most recent strand of literature seems to acknowledge the hedge funds abnormal performance as persistent over time.

Determinants of the performance

Faced with these complex investment vehicles capable of generating abnormal and persistent returns, many have tried to understand the influence of various factors on performance. As we have already seen, multifactor models and their risk premiums provide part of the answer by highlighting hedge funds' exposure and sensitivity to external factors. However, it has also been shown that the structure of the funds and their intrinsic characteristics, such as size, age, or fee structure, influence fund performance.

When considering the relationship between fund size and performance, opinions diverge. Some argue that larger funds, with more assets under management, outperform smaller ones (Amenc & Martinelli, 2003; Teo et al., 2003), while others assert that smaller structures perform better (Jones, 2007; Ammann & Moerth, 2005). Stafylas et al. (2016) settle the debate in favor of smaller funds by reviewing the literature on this specific topic, showing that there is more scientific evidence of a negative relationship between hedge fund size and returns.

We might also be interested in the interaction between returns and the age of the fund. By age, we mean the gap between the inception date and the time at which the return is observed. On this question, the literature is less uncertain and highlights a clear negative relationship between age and performance, meaning that younger funds tend to outperform older ones (Stafylas et al., 2016).

Another interesting correlation is the one between the fund's payoffs and the performance fees. Even though some studies have not been able to demonstrate any relationship (Stafylas et al., 2016), the majority agree that higher performance fees correlate with better earnings (Edwards & Caglayan, 2001; Ackermann et al., 1999). Soydemir et al. (2014) went further by asserting that funds with higher performance fees outperform those with lower incentive fees.

In some hedge funds, performance fees might be constrained by a high watermark mechanism. For an investor, the maximum share value since entering the fund is called the high-water mark. This means that performance fees are due only on the amount exceeding this mark (Goetzmann, 2003). Liang (1998) has highlighted the positive impact of this mechanism on the fund's payoffs. Hedge funds might also set a hurdle rate to attract investors. Below this predefined hurdle rate, a manager does not receive any performance fees. According to Soydemir et al. (2014), funds using these hurdles seem to be less performant.

Manager skills

The previous developments underline the abnormal performances of hedge funds over the years and attempt to explain them. Even if much of this might be attributed to external and internal factors, there remains a portion that cannot be justified by these elements. One might attribute these excess returns to pure luck; however, this does not align with researchers' opinions on the issue.

Indeed, Kosowski et al. (2007) have provided strong evidence of annual alpha persistence that cannot be attributed to luck. These conclusions are even more compelling given that they were drawn from samples extended through a bootstrap procedure. These findings are consistent with results obtained by Stulz (2007), Racicot et al. (2014), and Agarwal and Naik (200).

Therefore, since superior unexplained and persistent performance is not due to pure chance or luck, numerous papers attribute it to managerial capabilities. Although this opinion is not unanimously accepted in the literature (Malladi, 2020; Cai et al., 2018), it is still widely recognized (Ling et al., 2023; Cave et al., 2011; Agarwal et al., 2015).

This view is supported by Edwards & Caglayan (2001), who highlighted the great performance of funds with high performance fees, suggesting that it demonstrates the value added by the skills of fund managers. Some scholars go beyond merely acknowledging managerial skills and develop metrics to assess these capabilities. One such measure is defined as the probability of hedge fund returns belonging to a certain performance level, conditional on prior estimates of the fund's alpha and standard deviation (Chen et al., 2017). Berk & Van Binsbergen (2015) define the added value of a fund as the product of abnormal returns, relative to the manager's benchmark, and the inflation-adjusted assets under management, using it to measure performance. As mentioned in the previous paragraphs, the alpha from different multifactor models may serve as a proxy for managerial skills.

It is based on these studies that we will derive a skill measure. We will then integrate it into our estimation model to contribute to the academic literature on the value added by hedge fund managers.

III. Summary

This literature review has provided us with valuable insights into hedge fund performance. Firstly, it has highlighted the necessity of assessing the reliability of our database and rectifying potential biases. Ignoring these issues might have serious implications for the accuracy of our work. Furthermore, it has revealed the exotic statistical properties of hedge fund returns, features that must be considered when applying different statistical models.

Secondly, through multi-factor models and the analysis of fund characteristics, the scientific literature enables us to identify various factors impacting hedge fund performance. It underscores the need to account for external factors, related to the fund's positions, and internal factors, related to the fund's structure. Since our aim is to determine the variables explaining abnormal positive returns, these findings are valuable.

Thirdly, the academic community has emphasized the ability of hedge funds to generate significant unexplained returns over certain periods. This abnormal and persistent performance has been attributed to managerial skills, and some have defined metrics to quantify these skills. We will base our analysis on these theoretical frameworks to derive a skill measure and assess its impact on abnormal positive returns.

Although the literature provides a solid foundation for our reflection, it still has some shortcomings. Many studies focus on the left tail of hedge fund returns distribution, while less attention is given to the right extremity of the curve. Our study aims to address this gap.

Furthermore, the literature remains uncertain about the value added by managerial skills. By defining a relevant metric and incorporating it into our analysis, we will be able to determine whether managerial skills have a predominant effect on abnormal positive returns, allowing us to assess the managers' skills.

Finally, while many internal and external factors impacting returns have been acknowledged by the literature, the magnitude of their impacts remains unclear. Our model will address this issue by implementing a regression analysis to select variables with the highest explanatory power.

Methodology

The literature review has enabled us to introduce the important concepts and to understand the theoretical context of our study. We will now move from theory to practice using the methodology presented below. First, we will define the theory that links the extremes of the distribution to the variables under consideration: Extreme Value Theory. We will then focus on these variables, describing and justifying each of the restrictions placed on the set of observations used. We conclude this description of the methodology by highlighting the various key stages in the implementation of our procedure. By the end of this section, we will have clearly defined what data is considered, why it is included, and how it is processed.

I. The statistical model:

The Extreme Value Theory:

As stated in the introduction, the extreme value theory will be used to make the link between a set of factors, and the tails of a statistical distribution. Under certain assumptions, it demonstrates that the extremum of a given distribution follows a general pareto distribution, characterized by a shape and a location parameter. In the following paragraph, we'll explain the EVT in further detail and see how it applies to our context.

To build the connection between the distribution's tails and the set of external variables, the EVT requires a random variable, denoted Y, which follows a cumulative distribution function, F.

To estimate the marginal distribution of the extremum, we will assume Y belongs to the maximum domain of attraction (MDA) of the extreme value distribution, named G. In other words, when the sample size increases, the sample of extreme value of F converges in distribution to G.

Thanks to this assumption, if the sample is large enough, we can use the peak-over-threshold approach to estimate the tail distribution. According to this method, for a fixed threshold, u, the distribution of the exceedances tends to follow a generalized Pareto distribution when this limit tends to the most extreme observation, denoted y_F. A chart representing graphically the POT approach is available in the first appendix.

$$P(y \geq Y - u \mid Y > u) \xrightarrow[u \to y_F]{} \begin{cases} (1 + \frac{\gamma y}{\sigma})^{-\frac{1}{\gamma}}, & \gamma \neq 0 \\ \exp\left(-\frac{y}{\sigma}\right), & \gamma = 0 \end{cases}$$

As previously mentioned, the GPD is characterized by a scale and a shape a parameter, denoted respectively σ and γ. By estimating those two, we will be able to characterize the tail of the distribution, since both a positive shape and a high scale parameter are significative of a heavy tail. Indeed, the scale parameter will have an impact on the spread of the distribution and therefore a high value of this parameter involves a higher probability of extreme events. On the other hand, a high positive shape parameter is synonym of fat tail.

The estimation procedure will allow us to make the link between the distribution's tail and the set of external factors. To do so, we assume both parameters to be a linear combination of external factors.

$$\sigma(x^\sigma) = \beta_0^\sigma + \sum_{j=1}^{n} \beta_j^\sigma \, x_j^\sigma \ and \ \gamma(x^\gamma) = \beta_0^\gamma + \sum_{j=1}^{n} \beta_j^\gamma \, x_j^\gamma$$

Therefore, by regressing the maxima set using a maximum likelihood approach, we will be able to determine the impact of each external variables on the parameters of the GPD.

However, as underlined by Hambuckers and al. (2018), such procedure fails to capture the most significant variables. Indeed, by including a large set of variables, the model will be overfitted, reducing its effectiveness. To avoid this drawback, we will be using a penalized regression technique.

Two kinds of penalized approach will be considered in this work: the least absolute shrinkage and selection operator, or LASSO, and the adaptive least absolute shrinkage and selection operator, or adLASSO.

The LASSO coefficients are specified as follows:

$$\hat{\beta} = \arg min_{\beta_0,\beta} \left[\sum_{i=1}^{n} (y_i - \beta_0 - x_i\beta)^2 + \lambda \sum_{j=1}^{p} w_j |\beta_j| \right]$$

With β_0 the intercept, λ being the penalty parameter and $w_j = 1$ if we shrink the coefficient and 0 otherwise.

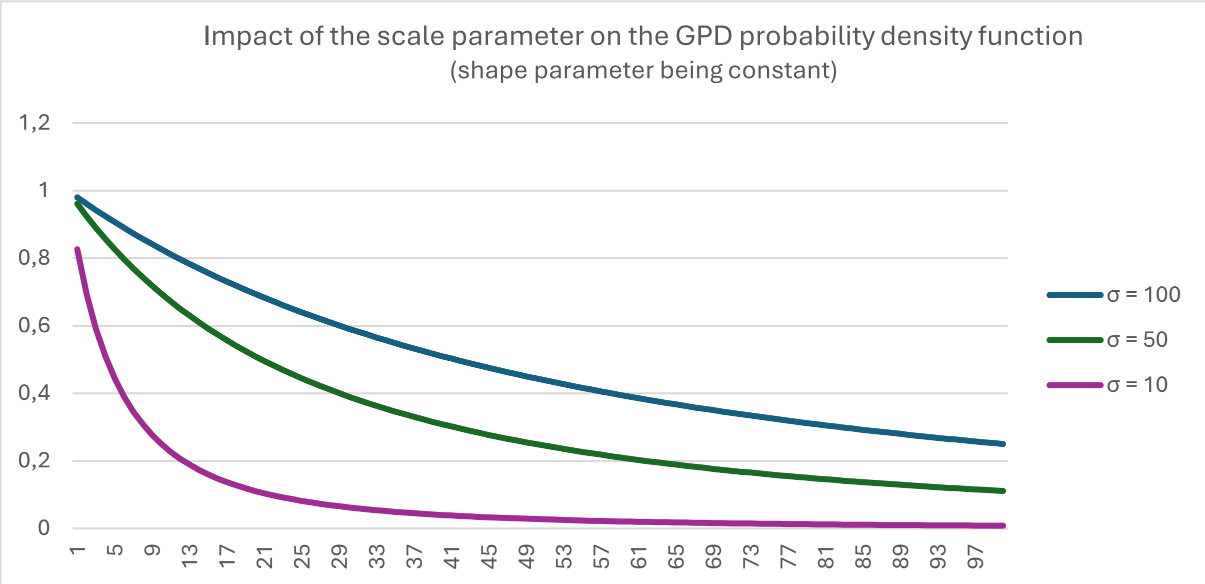
Zou (2006) have demonstrated that, in some cases, the LASSO regression is consistent only under certain conditions. That's why he has developed a modified version of this penalizing procedure: the adLASSO. The major difference lies in the definition of the element w_j. Indeed, in the adLASSO regression it is equal to the inverse of the absolute value of the unpenalized regression coefficient. The latter approach gives more stable results.

In our case, we will consider both and compare them thanks to the Bayesian information criteria since both are based on the negative likelihood approach. Once we have defined the most efficient penalized regression, we will obtain a set of relevant variables. We will proceed to an unpenalized regression on this effective set to determine the impact of those variables on the study variable.

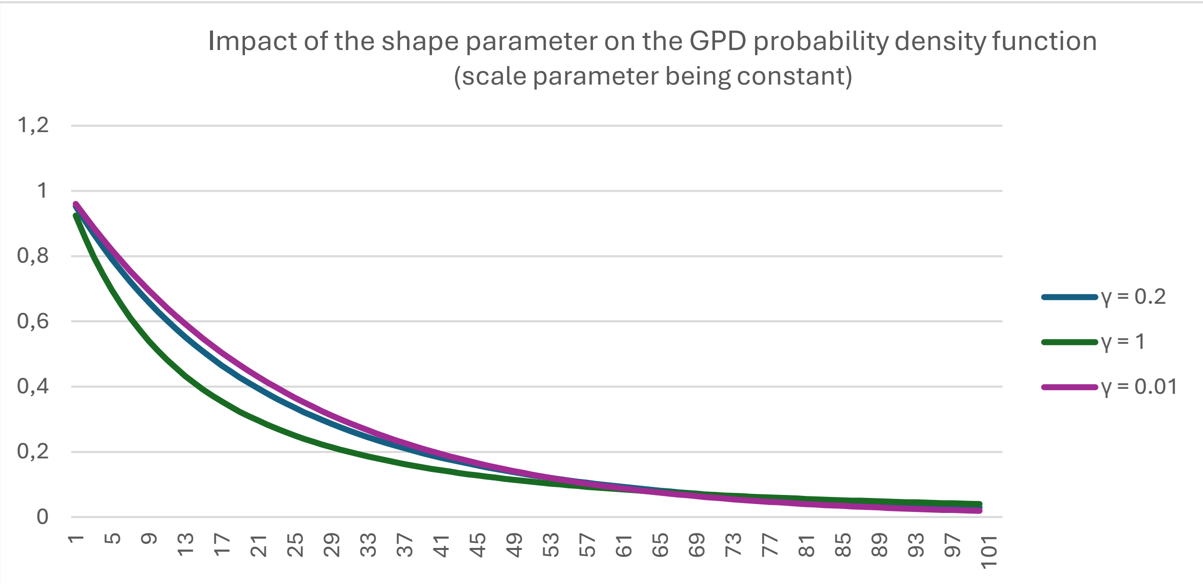
Interpretation of the General Pareto Distribution shape and scale parameters

Before demonstrating how this theory applies to the context of this work, it is important to understand the impact of the two parameters on the distribution. Indeed, the application of the methodology will not enable us to determine the number of abnormal returns, but rather an estimate of the probability of their occurrence. This estimate will be largely influenced by these two parameters, hence the importance of correctly estimating their impact on the statistical distribution.

To do this, we have simulated several probability density functions to determine by varying the value of these parameters. Initially, we will consider the case where gamma is constant.



As can be seen here, an increase in the scale parameter will have an impact on the thickness of the tail of the distribution. As the graph above shows, the probability of observing extremely high returns is much higher when sigma increases.



The impact of the shape parameter is similar to that of the scale parameter. In fact, we can see that the distribution has a wider tail as the value of the gamma parameter increases. However, we qualify this by adding that the value of the parameter will mainly influence the extreme values. As can be seen from the graph, when gamma is large, the probability of observing small values is greater, while the possibility of observing extreme values is greater than when the parameter is smaller.

To summarise, the scale parameter will have an impact on the whole distribution and can be considered as a measure of the spread of values, whereas the shape parameter will mainly influence the behaviour of the tails of the distribution and will therefore mainly impact the appearance of extreme values.

Application to the Hedge Funds context

Now that EVT has been clearly defined, we need to apply it to our context: that of hedge funds. This application has already been carried out in the work of Mhalla et al (2021) where they investigate the extremal connectedness of hedge funds. However, we will deviate from this work by considering only extreme gains, whereas Mhalla et al. (2021) focus exclusively on losses.

In this work, we will consider the set of returns generated by hedge funds as the variable of interest, Y, and we will assume that these follow a certain cumulative probability function, F. Since we are interested in positive observations, we would define the threshold, u, as an upper quantile of the distribution. We will then be able to estimate the distribution of profits above the threshold using a general Pareto distribution, provided there are enough observations. We will then estimate the parameters of the distribution as a linear combination of factors. In order to retain only those factors that have a significant impact on the result, we will apply a regularisation procedure.

Obviously, all this methodology depends on a major element: data. In the following section, we will describe the data used.

II. Description of the data

All statistical models are based on a set of data. In this work, we will use several databases. These will play a key role within our work since they represent the study's environment. We will divide our set into two subcategories: internal and external variables.

The first refers to those intrinsic to the fund, such as its returns, the strategy used, the location, the fee structure, etc. The latter will be used to model the structure of the speculative vehicles considered. However, as we highlighted in our literature review, this first category might suffer from biases that need to be corrected. The second subset refers to the data aimed at describing the environment in which the funds operate. Among these are variables accounting for the macroeconomic environment and the situation of the financial markets. In the following paragraphs, we will describe the data, demonstrate their value, and justify the restrictions imposed on the database.

The internal variables: the hedge funds database

The first category therefore includes all the data characterizing the fund. We will use a database from Eurekahedge, a company specializing in collecting data on hedge funds. This database was extracted and cleaned for the most part by Mr. Philippe Hübner, whom I would like to thank.

The database used comprises 6,600 funds, spread over 78 countries and created between 1994 and 2021. Each fund is associated with 24 variables that characterize it. These include returns, the amount of assets under management, whether the fund uses leverage, etc. These variables are summarized and described in the appendices.

The first thing we notice is the uneven geographical distribution of the funds. As shown in the graph below, the majority of funds are based in the United States, where 3,463 of the 6,600 funds are located, representing 52% of the funds considered.

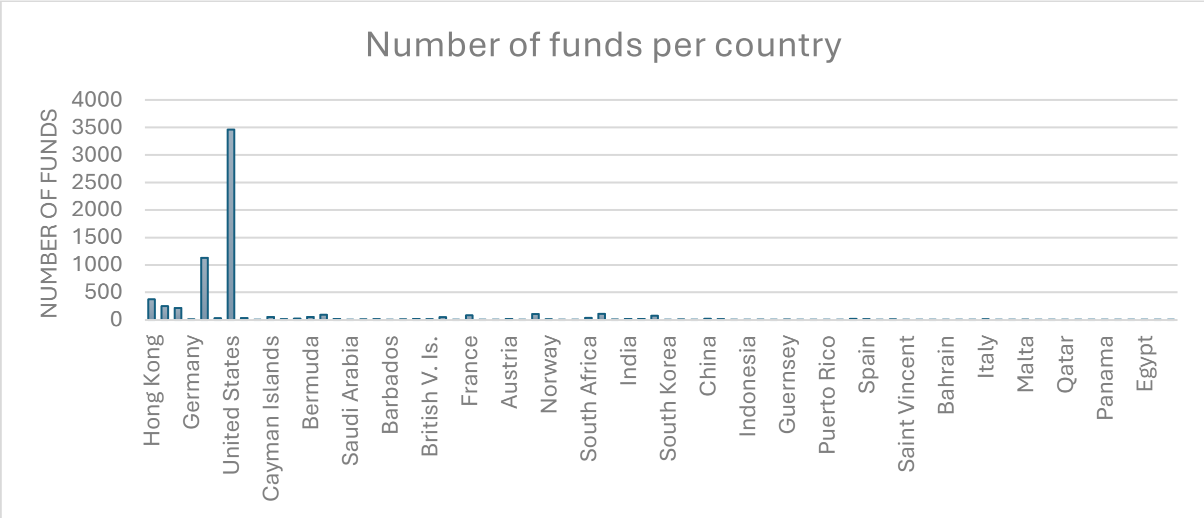
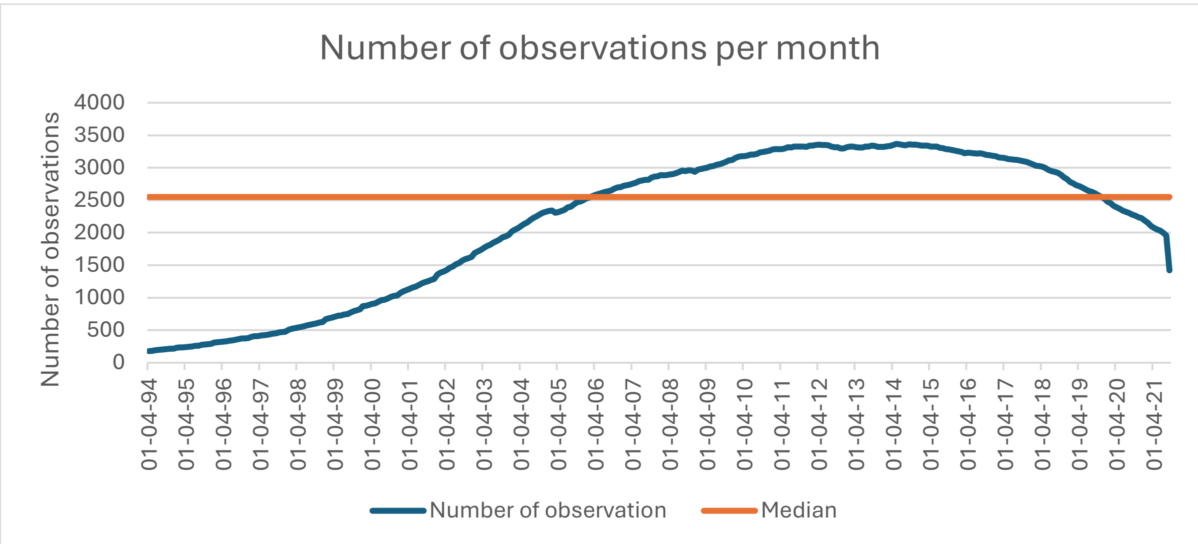


Figure 1 – Number of hedge funds recorded per countries

This inequitable distribution will lead us to consider only American funds. This approach has several advantages. Firstly, given the number of funds from this country, the sample we use will be sufficient for our analysis, which would not be the case if we were limited to funds from other countries. Additionally, later in the work, we will need to gather macroeconomic and financial data for the various countries considered. It will naturally be more straightforward to do this for a single country, the United States. Indeed, as a developed country, it provides access to a vast amount of reliable data without too much difficulty. This restriction will reduce the number of funds and observations to 6,600 and 333,719, respectively.

Along with this geographical dimension, we can also examine the period over which all our observations are made. The chart below shows the number of observations per month. It can be noted that over time, the number of observations per month increases, reaching a maximum on 31 May 2014, and then decreases until 2021. When looking at the different statistical moments in this series, we see a median of 2,550, which means that more than 50% of the time, there were more than 2,550 observations per month.



These considerations will permit us to restrict our data set a second time. We will only consider US hedge funds and returns between 2004 and 2021. This decision is justified by the lack of data available before 2004. As stated earlier, we will need external data to complete our analysis. These include the S&P 500 cumulative dividend yield and a proxy for the US bond market, which are only available from

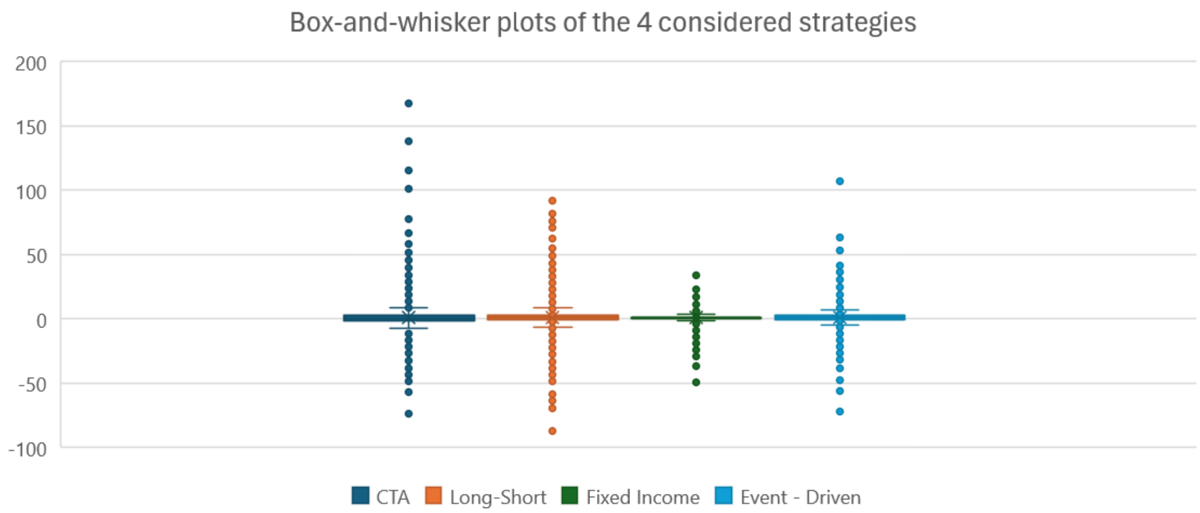
2004 onwards. By restricting the database, we will be able to benefit from these data. Additionally, this limitation allows us to retain more than half of the observations and focus on the most recent ones. Thus, we are not compromising the accuracy of our study. Our analysis will be based on contemporary data and will focus on a specific geographical area, the United States, to improve accuracy and efficiency.

After cleaning the data and applying all the restrictions set out above, we end up with a set of 211,221 observations. An observation corresponds to the return generated by a fund each month. To achieve such performance, funds organize their activities according to certain strategies; the list of those included in the database is detailed in the appendices. As shown in the graph below, the most common strategy is long-short equities (102,895 observations for 1,302 funds), followed by CTA or managed futures (57,705 observations for 590 funds), multiple strategies (23,734 observations), event-driven (23,583 observations for 220 funds), and fixed income (22,782 observations for 275 funds).

Only the most represented strategies will be considered in our analysis. As noted in the first part of the methodology, the Peak-Over-Threshold (POT) approach only holds if the sample size is sufficient. Given that we will not be considering the entire distribution but only the positive extremes, i.e., 6 to 9% of the total population, the sample must be adequately broad. Arbitrarily, we will not include strategies with fewer than 21,000 observations in total. According to this criterion, only the four strategies mentioned above are eligible.

However, as the name suggests, the 'multi-strategy' category includes all funds using various strategies. Unfortunately, the database does not provide enough information to determine which strategies are involved. It therefore seems more coherent to focus exclusively on single strategies, as funds using multiple strategies can be thought of as a "weighted sum" of all the strategies.

The chart below shows the different statistical moments for the four strategies under consideration: Long-short equities, CTA, Fixed income, and Event-driven strategies. Although all four have approximately the same mean, we notice the occurrence of extreme profits and losses. This is particularly true for the CTA strategy but less marked for the Fixed Income strategy. In fact, the interquartile range is much lower in the Fixed Income category compared to the other three. This seems logical given that fixed income instruments are described as relatively low-risk assets. However, while the scientific literature emphasizes the ability of hedge funds to outperform other asset classes, one might be surprised by the empirical averages demonstrated by the strategies considered. Nevertheless, it should be borne in mind that our database also includes inactive funds and the negative flows that may precede their extinction. These flows will pull the average down, which explains why all four strategies have averages close to zero.



Finally, we obtain four different data sets, each containing more than 21,000 observations and each characterised by 21 variables. Six of these variables have only an identification role or are redundant and must therefore be removed.

Potential biases

Part of the literature is devoted to a recurring problem inherent in hedge fund databases: bias. Their omnipresence has been highlighted in the review literature. This is the subject of the next section.

The first bias we will address is selection bias. As a reminder, this means that the databases reflect only the best-performing funds, thereby obscuring the real performance of the market. In the case of the database under review, we note extremely negative returns ranging from -49.82% to -87.21% across all strategies studied. Although these observations are not frequent, they allow us to rule out a selection bias. These large negative flows may also correspond to the liquidation of a fund, ruling out the hypothesis of a liquidity bias.

In the same spirit, we can rule out survivorship bias. This would require all liquidated or “dead” funds to be removed from the databases. However, keeping only the good performers does not reflect real performance and is therefore biased. In our case, there is a binary variable indicating whether the fund is still active or not. We can therefore reasonably rule out any survivorship bias.

Finally, instant history bias is much more complex to quantify than other types of bias. The complexity stems from the fact that it can be influenced by a varied set of contextual and circumstantial factors, making it tricky to assess precisely. In the context of this work, undertaking an in-depth analysis of this bias seems out of place, as it would require specific tools and methods that go beyond the objectives and limitations of our current study. However, it is important to highlight the possible existence of this bias, as it could have an impact on the interpretation of the results. I would therefore like to draw the attention of readers and analysts to this potential bias so that they can take it into account when interpreting the data and drawing conclusions.

The external variables : macroeconomic and financial market data

The purpose of all the above data is to describe the fund's performance, its size, its structure, and the way it organizes its investments. Our study could be limited to this set of internal factors. However, like any fund, hedge funds operate in a certain context. In addition to influencing the fund's performance, this background also serves to demonstrate the fund manager's skill. A good manager will be able to take advantage of the environment around him, revealing his true skills. We therefore need to take the context into account in our analysis by modeling it. To do this, we will add a series of macroeconomic and financial variables to our database to gain a better understanding of the conditions in which the manager operates and to determine its impact on performance.

We will base our thoughts on various scientific papers. The first one is the work of Bali et al. (2014), which encompasses eight variables measuring macroeconomic risks. They have been able to show the significant impact of these on the fund’s performance, which is the reason why we will consider them in our analysis. Among these eight variables, three have been acknowledged in the work of Lambert and Platania (2020), namely the US GDP growth rate, the relative treasury bill rate, and the aggregate dividend yield on the S&P500. Since we aim for a representative sample, we will use all eight variables. These are summarized and described in the second table of the appendix. As with the fund database variables, these are expressed in percentage on a monthly basis. Lambert and Platania (2020) introduced other relevant variables such as the volatility index of the S&P500, the VIX.

In a preliminary version of this work, these data were the only ones used to represent the study environment. However, following the presentation of this earlier version as part of the oral examination for the Advanced Statistical Methods in Finance course, Professor Julien Hambuckers and doctoral student Mr Philippe Hübner pointed out the lack of purely financial data. These observations are in line with the work of Agarwal and Naik (2003), who advocate the use of financial data such as the Russell index, the MSCI World and Emerging, the US government and corporate bond index, the World bond index, etc., to explain hedge fund performance. Once again, all these data are described in the second table of the appendix.

Therefore, after the data treatment, we obtain four sets, each following a specific strategy and characterized by 15 variables describing the fund and 17 variables describing the environment in which the fund evolves.

III. Implementation of the methodology

The previous developments have enabled us to obtain a coherent and exhaustive set of data. This will serve as the basis for developing our statistical model. This implementation will be carried out in three stages. First, we will select the threshold that determines the extremum samples. Next, we will determine which regression, penalised or not, is the most efficient before putting in place control mechanisms, with the aim of evaluating the performance of our model. But before we look at these three stages, we'll take a brief aside to explain the fate of the binary variables and the variable representing the manager's skills.

Binary variables

When it comes to implementing our methodology, the first question that arises is what to do with the binary variables. Indeed, the set of variables under consideration is a mixture of binary and non-binary variables.

The singular configuration of these variables can have an impact on the regression used. We therefore need to decide whether to include these binary variables in the regularisation procedure. Of course, this question has arisen before and has been the subject of several studies.

The most convincing is that of Meier et al. (2008), who introduce an extension of LASSO regression, known as the LASSO group. This allows variables to be penalised by separating them into groups so that the optimisation problem can be carried out in two stages. This could therefore be a solution to our problem.

However, its implementation is complex and requires a large amount of work to analyse the impact of 7 variables out of a total of 33. I therefore think that the LASSO group is an interesting approach that could be considered in later work. Unfortunately, I thought it more useful not to dwell on this procedure, and I therefore decided to exclude the binary variables from the penalty regression and to add them to the set of significant variables obtained via this same regression.

Managerial skills

One of the objectives of this work is to quantify the ability of fund managers to generate surplus profits. In the literature review, we have already discussed this subject, highlighting the various studies carried out on the subject.

These include various measures, one of which we will retain. For this analysis, we will start from the assumption that everything that cannot be explained by our model results from the managers' ability. Consequently, the interception of the various regressions performed will be considered as an approximation of the value added by the manager.

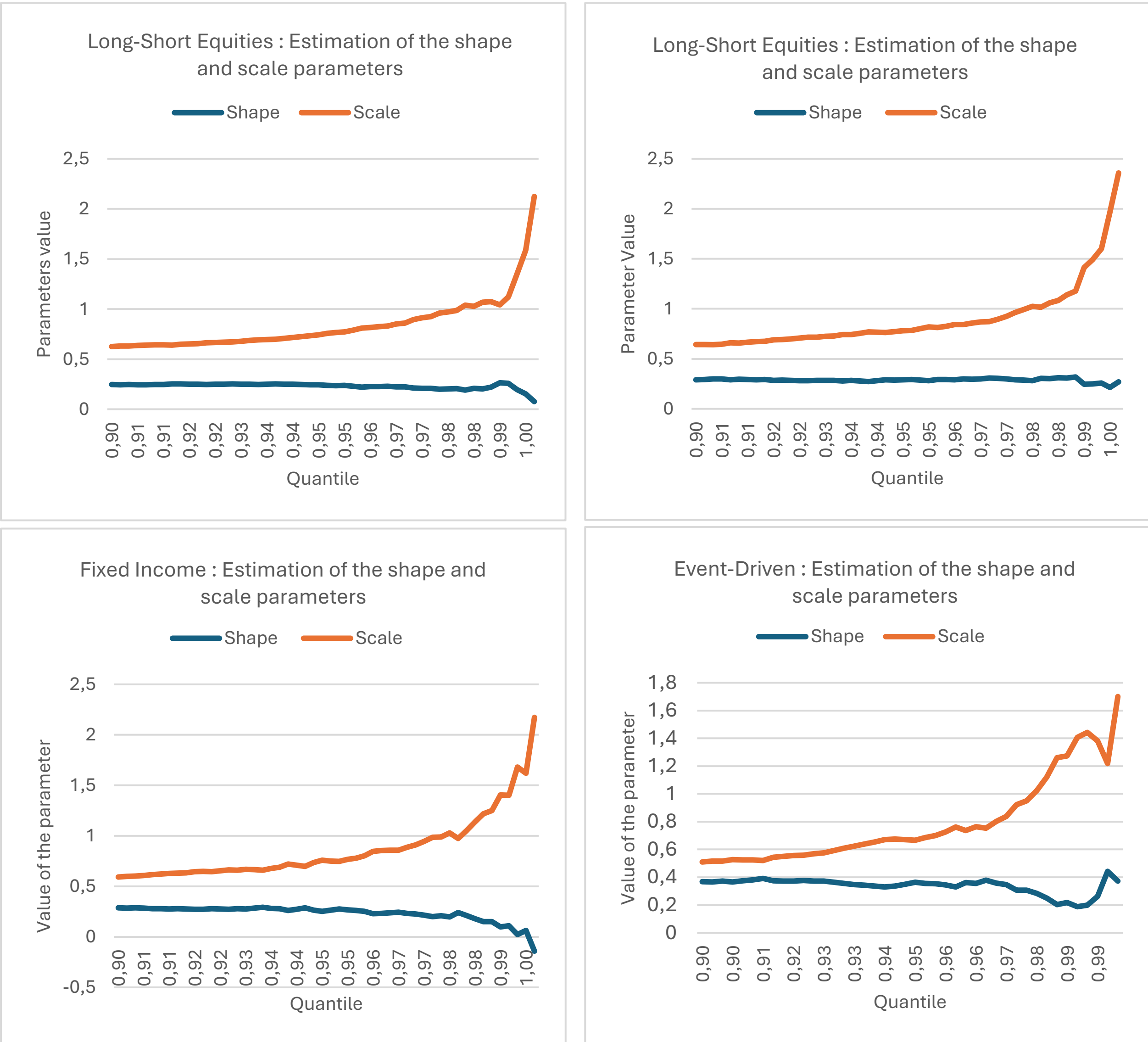
This assumption is in line with many in the literature that recognise alpha as the unexplained part of returns, generally due to managerial skills.

Threshold selection

As stated in the section on extreme value theory, according to the peak-over-threshold approach, if the initial sample is large enough, the sample of maxima will follow a general paretto distribution. However, this implies that the choice of threshold is of capital importance and will therefore influence the result. If the threshold is too high, there is a risk that the sample will be too small, while conversely, there is a risk that the sample will be too large and will induce too much variance. We therefore need to find a way of defining the threshold in such a way as to optimise the efficiency of the model.

To determine which threshold value is optimal, we will estimate the GPD parameters on different samples obtained by varying the threshold value. In this way, we obtain the variation of the parameter estimate as a function of the threshold value. The optimum value will be obtained in the interval where the parameters are stable. The graphs below show the estimation of the shape parameter as a function of the threshold.

Table 1 - Graphs of the estimated scale and shape parameters as a function of the quantile of the overall distribution, for the four strategies considered.



These graphs highlight the impact that the threshold has on the parameters of the distribution of the extremum sample. For each strategy, we have selected a specific threshold, shown in the table below:

Strategies	Threshold selected	Number of observations in the remaining sample	Mean	Standard deviation
Long-Short Equities	96%	4116	13,43%	6,71
CTA	95%	2482	13,09%	8,53
Fixed income	92%	1823	4,91%	2,94
Event-driven	90%	1961	7,58%	5,44

This initial descriptive analysis allows us to demonstrate the correct proportioning of our four samples. We can observe that each of the samples contains a minimum of 1,500 data points, ensuring an unbiased analysis. At the same time, we note that the variances remain within reasonable values and are in the same order of magnitude, a sign of good sampling.

Regularization procedure

With the various samples defined, we now need to analyse them. As explained earlier, this analysis will be carried out using two types of penalised regression. These will enable us to restrict the set of variables by keeping only the significant ones. In this way, our statistical approach will enable us to highlight the factors that have a genuine impact on returns, as well as the magnitude of this impact.

LASSO regression

The two types of regression considered are LASSO and adaptive LASSO. To fully understand the aim of our approach and justify the use of these methods, we will briefly review these two concepts. The LASSO regression has been introduced by Tibshirani in 1996. The aim of this technique was to provide a more effective way to solve the several drawbacks exhibited by the ordinary least square and maximum likelihood estimation regression methods. Indeed, the latter technique provides estimates with a too large variance and complexify the interpretation of the results by taking too many variables into account. The LASSO regression is specified as follows:

$$\widehat{\boldsymbol{\beta}}^{\,LASSO} = \mathbf{arg}\,\boldsymbol{min}_{\boldsymbol{\beta_0},\boldsymbol{\beta}}\left[\sum_{i}^{n}(\boldsymbol{y_i}-\boldsymbol{\beta_0}-\boldsymbol{x_i}\boldsymbol{\beta})^2 + \boldsymbol{\lambda}\sum_{j}^{p}\boldsymbol{w_j}\left|\boldsymbol{\beta_j}\right|\right]$$

Similar to: $\widehat{\boldsymbol{\beta}}^{\,LASSO} = \mathbf{arg}\,\boldsymbol{min}_{\boldsymbol{\beta_0},\boldsymbol{\beta}}\left[\sum_{i}^{n}(\boldsymbol{y_i}-\boldsymbol{\beta_0}-\boldsymbol{x_i}\boldsymbol{\beta})^2\right]$ ***such that*** $\sum_{j}^{p}\boldsymbol{w_j}\left|\boldsymbol{\beta_j}\right| < \frac{\mathbf{1}}{\boldsymbol{\lambda}}$

The LASSO regression adds a penalizing term to the classical regression procedure. Thanks to this addition, the regression will shrink the unsignificant factors' coefficient to zero and retain only the most impactful variables, which OLS techniques are simply incapable of doing (Tibshirani, 1996).

adLASSO regression

Although this represented a real advance in the field, there is still room for improvement in penalised regression techniques. A great deal of work has been done on the subject over the years. These have led to numerous adaptations of the original model (Tibshirani, 2011). Among all those variants, one will catch our attention: the adaptive LASSO (Zou, 2006).

In the field of statistical techniques, it is desirable for a fitting procedure to display oracle properties, i.e. consistency in the selection of variables and assymptotic normality of the coefficients. This ensures the reliability of the coefficients obtained by the regression (Fan and Li, 2001). Unfortunately, this is not the case for those obtained by a LASSO procedure. Consequently, adLASSO was developed to provide coefficients with these features. The major difference between these two approaches lies in the specification of the penalizing term, as demonstrated below.

LASSO penalizing term	adLASSO penalizing term
$\lambda \sum_{j}^{p} w_j \lvert \beta_j \rvert$ with $w_j = 1$ if the coefficient is regularized and 0 otherwise	$\lambda \sum_{j}^{p} w_j \lvert \beta_j \rvert$ with $w_j = {}^{1}\!/_{\lvert \theta_j \rvert}$, if the coefficient is regularized and 0 otherwise. θ_j being the coefficient of the unpenalized regression.

Both procedures have real advantages. Nevertheless, only one of them will be selected. To determine the optimum one, we will base on the BIC, which is in line with the method used by Hambuckers et al. (2018). Once the penalty parameters optimized, we will select the template that shows the lowest BIC. We detail this selection in the next paragraph.

Selection of the regularization procedure and penalty parameters optimization

Even though they differ on some notions, they both introduce a penalty parameter, λ, which will control the shrinkage power of the regression. This parameter must be tailored since it will obviously influence the output of the procedure. In our case, we have two penalty parameters to determine: for gamma and sigma. Two methods might be put in place to reach optimal penalty parameters values (Hambuckers and al., 2018).

Firstly, these two parameters can be defined simultaneously using a two-dimensional grid. This involves fitting the procedure to the data several times, varying the sigma and gamma penalty parameters at each iteration. At the end of this procedure, the model specification that minimises the BIC is adopted. This procedure is carried out for each type of regression considered.

A second approach is to carry out this optimisation in series. This means optimising one of the two parameters, for example sigma, and then the other, in this case gamma. The reverse should also be done, i.e. optimise gamma then sigma. Once again, we select the approach and regression that minimise the BIC.

Hambuckers et al (2018) chose the first option. To provide an innovative perspective on the issue and to allow a comparison between the two approaches, I decided to adopt sequential optimisation which, to my knowledge, has not yet been implemented. The value of the optimal parameters is summarized in the “results” section.

Analysis of the results and discussion

In this section, we will analyze the results obtained during our analysis.

Before starting the analysis, it is worth recalling the impact of the scale and shape parameters on the distribution of extremums. The first parameter, the scale, affects the dispersion of the observations. The larger it is, the larger the tail of the distribution will be, meaning that we are more likely to observe extreme positive returns. The second parameter, the shape, also affects the size of the tail, but in a slightly different way. Indeed, the shape parameter will impact the tail of the general Pareto distribution. Since the latter is used to estimate the tail of the hedge funds' return distribution, the shape parameter will thus impact the "tail of the tail," meaning that a large value of gamma will influence the probability of occurrences of extremely large returns. A representation of the effect of these parameters on the density curve of a GPD is shown in the methodology section.

I. Optimal regularization procedure

Now we turn our attention to the results. The tables in Appendix 3 show all the optimized values of the parameters and the Bayesian Information Criterion (BIC) obtained. The most obvious observation is that the most convincing regularization procedure is the adaptive LASSO. Indeed, once the penalty parameters have been optimized, it minimizes the BIC for the four strategies considered. This observation aligns with the scientific literature on the subject. Hambuckers et al. (2018) had already highlighted the advantages of the coefficients obtained by this procedure, while Zou (2006) emphasizes the beneficial properties of these coefficients.

We notice that, in half of the cases considered, the optimal model is obtained by penalizing gamma before sigma. In these cases, the value of the penalty parameter for gamma is always significantly higher than that for sigma. This results in a drastic reduction in the number of variables labeled as significant for gamma. For the other two strategies, fixed income and CTA, the model is optimized by penalizing sigma before gamma. This approach results in much more balanced and relatively high penalty parameters compared with those obtained for other strategies. However, once again, this results in a very small number of variables retained for gamma. Several conclusions can be drawn from this observation.

Firstly, this means that all the variables considered will impact the occurrence of high profits, but this impact will be less pronounced for extremely high profits, those at the tail end of the GPD distribution. As a result, the appearance of these extreme values will tend to be more independent of the environment and background structure. Furthermore, assuming that the intercept is a good approximation of fund managers' abilities, this would mean that the frequency of abnormally high returns is mostly due to their skills in managing funds, since the only significant variable is the intercept. These conclusions are valid for both cases—where gamma is penalized before sigma and where sigma is penalized before gamma—demonstrating the robustness of this interpretation.

In conclusion, the analysis of the penalty parameters and the number of variables retained leads us to an initial conclusion: the context in which hedge funds operate and their structure have a larger impact on positive returns, but not on abnormal positive returns. These seem to be due to the investment skills of fund managers.

II. Significant variables interpretations

We will now examine the variables selected, their interpretation, and their sign. Indeed, a significant number of relevant covariates would indicate a clear influence of the context and structure of the funds on their performance. Conversely, a lack of significant covariates would demonstrate a certain degree of independence from these factors. Moreover, the significance of each variable is crucial for a precise understanding of the impact of the context on the probability of occurrence of abnormally positive returns. Finally, the sign of the variables will help us determine whether the variable promotes the occurrence of positive observations or not. To do so, we will analyze each strategy independently before summarizing the findings and highlighting potential global trends.

Before embarking on this individual analysis, it should be noted that the seven binary variables were added during the non-penalized regularization procedure. We will therefore not consider them significant, although they will be analyzed separately.

Long-short strategies

As a reminder, equity hedge, or long-short, strategies involve constructing a portfolio that holds both long and short positions in equities, with the goal of achieving market neutrality. This widely used approach in non-traditional investing hinges on the manager's expertise in selecting the right stocks for both sides of the portfolio. The strategy is designed to generate profits regardless of whether the market is rising or falling, as gains from one side (long or short) should offset losses from the other. However, the success of this strategy is heavily dependent on the manager's ability to make accurate predictions; poor stock selection or misjudgement in market conditions can result in substantial losses, particularly due to the complexity and risks associated with the hedging techniques employed. In this section, we will delve into the results obtained for this strategy, analysing its performance and the factors that contribute to its success or failure.

The table below shows the selected variables, the binary variables and the non-penalised coefficients for the non-standardized variables.

Table 2 - Coefficients of the unpenalized regression on the unscaled significative and binary variables for the CTA strategy.

	Sigma
Significant variables	
Intercept	1,421
RGPD per capita	-22,798
T10Y3M	-0,007
Binary variables	
Closed	-0,207
Dead	-0,044
Hurdle rate	0,142
High Water Mark	0,066
Leverage	0,100
UCITS Compliant	-0,819
Lock up	-0,082

	Gamma
Significant variables	
Intercept	0,088
T10Y3M	-0,042
Binary variables	
Closed	0,117
Dead	-0,025
Hurdle rate	-0,048
High Water Mark	0,120
Leverage	0,055
UCITS Compliant	-0,515
Lock up	0,116

The first observation to be made is the limited number of variables selected. Only three were selected (two for sigma and one for gamma) out of a sample of 32. This seems in line with the desire of a manager practicing this type of strategy to achieve a neutral risk exposure. More precisely, we note the absence of all the variables used to model the global and US equity markets, meaning that they have no impact on the parameters of the distribution. This fits with the market neutrality of the strategy. Additionally, all the variables relating to the bond and commodities markets have been eliminated. This implies that long-short hedge fund neutrality is not limited to the equity market.

Some might point to the presence of the "T10Y3M" variable, the term spread, to justify the exposure of funds following this strategy to the fixed income market. Here, the term spread is defined as the difference between yields on 10-year and 3-month Treasury securities. Although it is based on government bonds, the term spread is often considered a macroeconomic variable reflecting investors' predictions of economic growth. If the 10-year yield is higher than the 3-month yield, this means that investors believe that government bonds will strengthen in value, which is synonymous with high economic growth. Conversely, if this relationship is reversed, it indicates investors are expecting an economic recession.

We note that this variable is significant for both sigma (shape) and gamma (scale) parameters, underlining its importance. Additionally, both coefficients obtained in the non-penalized regression on the non-standardized variables are negative, demonstrating a downward relationship between the term spread and the value of the general Pareto distribution parameters. An increase in the spread will therefore reduce the probability of positive and extremely positive returns. This may seem contrary to what one might imagine. Indeed, it is not unreasonable to suppose that a growing economy favors the emergence of excess profits. This is without considering the risk-neutralization mechanisms put in place by long-short fund managers. In fact, as we said earlier, they can make profits even when the market is in bearish conditions, notably through short selling. Therefore, the fact that this value influences both parameters and is included among the significant variables would suggest that hedge funds make greater excess profits when markets are down and, assuming that a downturn in the economy inevitably leads to a bearish cycle in the equity market. We can attribute these unusual performances to their short-selling operations. The sign of the coefficients shows that the opposite is not true, or at least not to the same extent, because if the excess profits made in bullish conditions via traditional operations were greater than those made in bearish conditions, the coefficient would be positive. We can also underline the difference between the absolute values of the two coefficients. The one relative to gamma is greater than its sigma counterpart, indicating that even though the spread term has an influence on both, its impact on gamma will be more pronounced.

Therefore, our analysis demonstrates the ability of market-neutral hedge funds to generate abnormally high profits when the macroeconomic system in which they operate slows down. These conditions have an impact both on the positive and extremely positive returns and can be attributed in particular to short-selling operations.

This hypothesis is supported by the second variable included in the estimation of sigma by the adaptive LASSO regression: real gross domestic product (GDP) growth per capita. In fact, the negative sign of the coefficient obtained means that, like the term spread, an increase in GDP will reduce the probability of occurrence of extremely positive values. The same reasons can be put forward to justify this relationship. However, unlike the term spread, GDP per capita does not influence gamma, which seems to indicate that it will have a more pronounced impact on the positive return and not on the extremely positive.

This difference can be explained by the fact that the term spread is an indicator more focused on the financial sector and is seen as a predictive measure of the economic health of a system rather than a factual sign of recession. On the other hand, GDP is a more global indicator showing the dynamism of an economy but does not predict it. Therefore, the fact that the term spread influences gamma and not GDP would suggest that fund managers would use the spread term to predict a cycle of economic recession to execute lucrative short-selling operations. Once GDP falls, the economy is effectively in recession, and the profits generated by these operations remain high but less. The inclusion of real GDP per capita growth shows the ability of hedge funds to generate profit from a downward economic context and to time the market.

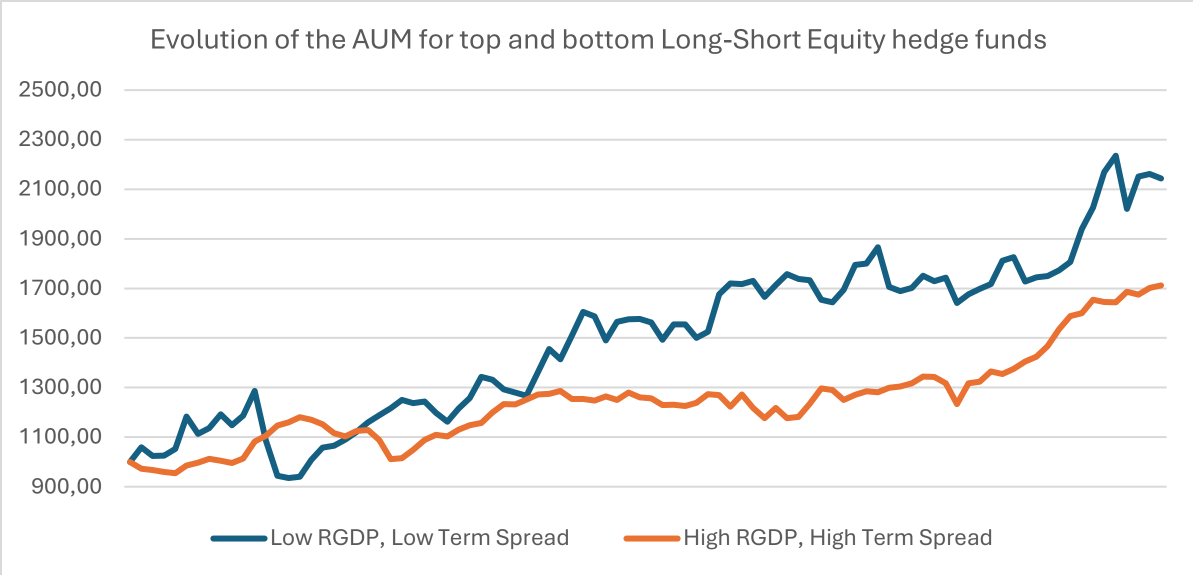
This ability to predict the state of financial markets can be seen as part of the range of skills of hedge fund managers. In any case, this is what our model suggests through the value of its intercept. As previously stated, this value will be used as an approximation of managers' investment capabilities. Based on this hypothesis, we note that the intercept of the two parameters is positive each time, which is in line with the conclusions of scientific work on the subject. However, their absolute values differ widely. Since the coefficient of the sigma intercept is greater, this would imply that managers' abilities have less impact on higher positive returns.

It should also be noted that, irrespective of investment skills, the structure of the fund does not seem to have a significant impact on its performance.

In summary, our analysis highlights the ability of market-neutral hedge funds to generate abnormal profits during economic downturns, primarily through short-selling. The significant negative coefficients associated with the term spread and GDP per capita growth indicate that these funds perform better in bearish markets, with the term spread also affecting the magnitude of these profits. This underscores the skill of hedge fund managers in timing the market and capitalizing on economic predictions, particularly during recessions, where their investment strategies seem to maximize excess profits.

To assess the precision of our results, we selected all hedge funds using a market-neutral strategy whose RGDP variable was below the first decile or above the ninth decile. This allowed us to create two groups of data: one with low and one with high RGDP variations. In the group with low RGDP values, we chose the fund for which the term spread was below the first decile. The fund with the highest number of observations among those remaining is the one that, according to our model, should perform best. We applied similar manipulations to the second group of data, the one with the highest RGDP, in order to identify a fund assumed to be underperforming. We then calculated the value of the assets under management for the two funds selected. Assuming that only the fund's performance influences its value, and that both funds start with an initial value of $1,000, the comparison of the curves obtained provides a visual representation of the respective performances, making it easier to compare them.

We have decided to start with the variable “RGDP” since it has the greatest unstandardized regression coefficient in absolute value.



	Low RGDP, Low Term Spread	High RGDP, High Term Spread
Mean of the returns	0,9491%	0,6237%
Variance of the returns	0,002	0,0006

This comparison illustrates the better performance of the fund operating in a context where the spread term and the growth rate of real gross domestic product are high. In addition, there is also greater variance in the series of returns for this hedge fund, indicating a higher probability of generating exceptionally positive returns. This copes with the results obtained by our model.

CTA or Managed futures strategies

As a reminder, managed futures strategies, also known as commodity trading advisors (CTAs), rely on sophisticated quantitative models to guide their trading decisions. These models, rooted in fundamental and econometric data, enable CTAs to actively trade futures contracts across a diverse range of markets, including commodities, currencies, and financial instruments. A key characteristic of these strategies is their focus on price trends, which sets them apart from macro strategies. While CTAs primarily follow and respond to market price movements, macro strategies are driven by broader macroeconomic analyses and forecasts, rather than direct price trends. This distinction highlights the unique approach CTAs take in the hedge fund industry, leveraging systematic models to identify and capitalize on emerging trends in various markets. We will analyse the results given by our model regarding this strategy.

Here are the two tables summarising the coefficients obtained by the unpenalized regression of unscaled significant variables and those of the binary variables.

Table 3 - Coefficients of the unpenalized regression on the unscaled significative and binary variables for the CTA strategy.

	Sigma
Significant variables	
Intercept	-0,227
Redemption notification period	0,002
RGDP per capita	-22,798
RREL	-0,026
Binary variables	
Closed	0,018
Dead	-0,095
Hurdle rate	-0,088
High Water Mark	0,019
Leverage	-0,064
UCITS Compliant	-0,060
Lock up	-0,033

	Gamma
Significant variables	
Intercept	0,300
Binary variables	
Closed	-0,009
Dead	-0,008
Hurdle rate	0,002
High Water Mark	0,014
Leverage	-0,033
UCITS Compliant	0,155
Lock up	-0,095

The first observation that stands out is the absence of significant variables for gamma. Indeed, although the penalization parameters have similar values, no relevant variable has been identified for the shape parameter, gamma. We can therefore logically conclude that neither the context nor the organizational structure of the hedge fund has any impact on abnormally positive observations. For the scale parameter, on the other hand, there is an influence of both, although the context seems to be less influential. As a result, the selection of variables by our regularization procedure shows that external variables can significantly increase the occurrence of positive profits but have no impact on the most extreme ones.

This last observation indicates that external variables do not allow managers to predict the movements of the various markets in which they are positioned. Indeed, the only way to increase the value of abnormal observations is to rely on the skills of the hedge fund manager, as our model shows by demonstrating a non-zero positive value for the gamma intercept. If we assume that advanced quantitative models and their use are part of the skill set, it is logical that they should lead to recurring profits. The constant effort that goes into perfecting these mechanisms will enable users to increase their ability to generate abnormal profits.

Despite the lack of significant variables for gamma, we note that our model selects three variables influencing the value of the sigma parameter. Of these, two are macroeconomic variables and one is structural. The latter is the notification period that has to be respected prior to the withdrawal of funds. This mechanism is one of those put in place to ensure the illiquidity of the hedge fund, a key feature of these investment pools. Our model indicates that the longer this period, the greater the fund's ability to generate positive returns. The fact that this variable is positively correlated with the emergence of positive returns suggests that the sooner managers are warned of outflows, the better they can prepare for them and reallocate the remaining capital. This period will thus reduce uncertainty about the funds available. As well as having an impact on excess profits, it shows that too short a notification period will greatly harm the fund and reduce the probability of extreme performances. This finding allows us to formulate another recommendation for hedge fund managers: increase the notification redemption period.

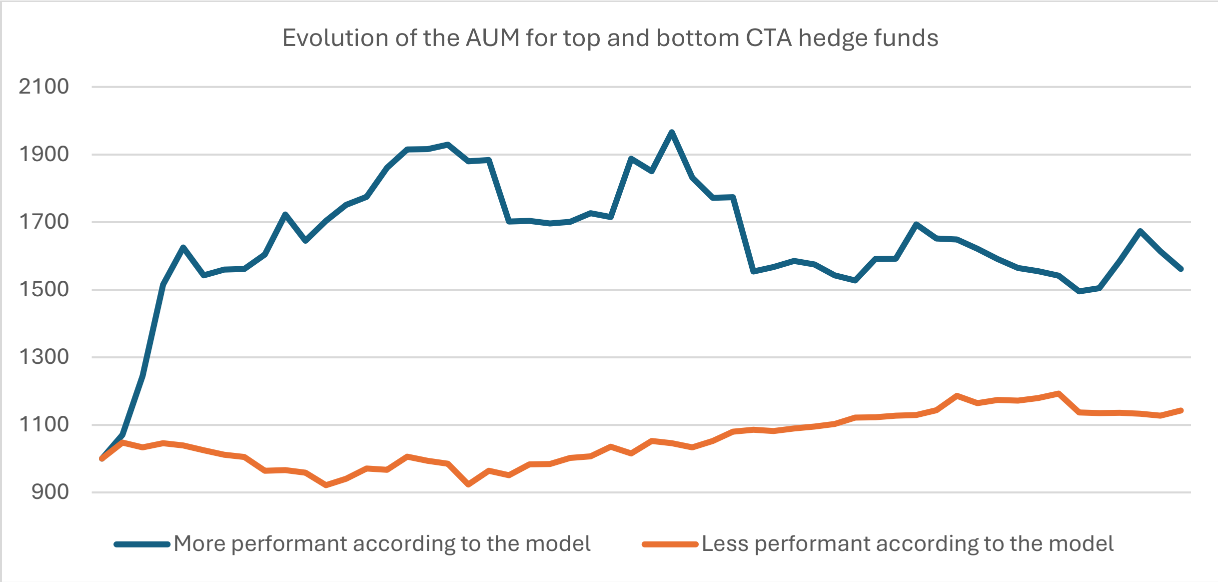
Alongside this structural variable, the relative treasury bill rate and real gross domestic product (GDP) per capita growth are also defined as significant. The first one, "RREL," is defined as the difference between the 3-month T-bill rate and the 12-month backward moving average and can be interpreted as a measure of short-term interest rate trends and shifts in market sentiment or monetary policy. The second one, well known, is a proxy for economic activity. An increase in both would be synonymous with a dynamic economy. Our model highlights the negative relationship between these two variables and the dispersion of extreme observations. Consequently, a period of recession would be conducive to the realization of excess profits. A recession is often accompanied by a general downturn in all markets (bonds, stocks, commodities, etc.). The negative sign of the two coefficients demonstrates the ability of hedge funds to generate profits, particularly in a down cycle. In the case of a CTA strategy, this can be explained by the use of quantitative models. These models will detect all overvalued assets and allow hedge funds to profit from them during a market correction generated by the macroeconomic context. The opposite is also true, although the effect is presumably smaller given the sign of the coefficient. This conclusion is in line with the previously made analysis of the results for the long-short equity strategy and seems to confirm that hedge fund returns are particularly important when the economic situation of the environment in which the hedge fund operates deteriorates.

Finally, we can point to a more surprising observation: the sign of the coefficient of the intercept obtained through the regression of the sigma parameter. If the coefficient for gamma is positive, we might expect that the coefficient for sigma would also be positive, which would mean that the manager has the ability to increase the amount and frequency of positive returns. However, this is not what our model indicates, as it highlights the negative impact of the intercept on the shape parameter. Assuming that the intercept is a consistent measure of the capacity of fund managers, this difference in sign between the two intercepts allows us to assert that managerial skills have a negative impact on positive observations but still manage to enhance extreme positive profits. According to this observation, managerial skills seem to increase the convexity of the GPD. Further investigation must be carried out to fully understand this observation.

In conclusion, our analysis reveals that while the context and organizational structure of a hedge fund have little impact on the extreme abnormal profits of CTA hedge funds, external factors like macroeconomic conditions influence the probability of positive profits. Notably, recessionary periods, characterized by declining economic indicators such as the relative T-bill rate and GDP growth, tend to enhance hedge funds' ability to generate excess returns, potentially through quantitative models. However, managerial abilities seem to impact the appearance of positive returns negatively. This nuanced understanding highlights the importance of both external economic conditions and internal fund structures in shaping hedge fund performance.

To test the results of our model, we applied the same procedure as that used to evaluate the performance of long-short equity funds. However, here we had to consider three variables. To define the two final sets, we first restricted on the basis of the 'RGDP' variable, then 'RREL' and finally 'Redemption notification period'. We proceeded in this order based on the absolute value of the standardized regression coefficients, selecting the highest coefficient first and then in descending order.

The results obtained are displayed hereunder



	More performant	Less performant
Mean of the returns	0,99%	0,15%
Variance of the returns	0,003	0,0005

Once again, we see higher returns for the hedge fund labelled as performing by our model. Moreover, the variance of the latter is also higher, which is consistent with the hypothesis that it is more conducive to the appearance of high returns.

Fixed income strategies

The fixed income strategy, as the name implies, seeks to capitalize on price discrepancies between various fixed income assets, such as bonds, treasury bills, and other debt securities. The effectiveness of this strategy is largely due to the unique characteristics and structure of the fixed income market. Unlike equities, where more standardized valuation models exist, the fixed income market lacks a universally accepted absolute pricing model. This absence of consensus creates opportunities for identifying mispriced securities. Furthermore, there are multiple relative pricing relationships among different fixed income instruments, such as interest rate differentials, credit spreads, and yield curves, which can be exploited for profit. The market is also influenced by irrational yet predictable factors, such as fluctuations in supply and demand driven by investor behavior, regulatory changes, or macroeconomic shifts, all of which can lead to temporary mispricings. Additionally, the complexity of certain fixed income securities, such as mortgage-backed securities or structured products, further contributes to the market's inefficiencies. This complexity, combined with the aforementioned factors, makes fixed income arbitrage a particularly attractive strategy for hedge fund managers who are adept at navigating the nuances of this market and identifying arbitrage opportunities that may not be immediately apparent to other market participants.

Here are the table summarizing the results of the analysis:

Table 4 - Coefficients of the unpenalized regression on the unscaled significative and binary variables for the Fixed Income strategy.

	Sigma
Significant variables	
Intercept	-0,172
UNRATE	-0,019
S&P Emerging BMI	-0,011
Binary variables	
Closed	-0,067
Dead	-0,103
Hurdle rate	-0,049
High Water Mark	-0,112
Leverage	0,036
UCITS Compliant	-0,083
Lock up	0,029

	Gamma
Significant variables	
Intercept	0,254
S&P Emerging BMI	-0,007
Binary variables	
Closed	-0,042
Dead	0,018
Hurdle rate	-0,002
High Water Mark	0,014
Leverage	0,001
UCITS Compliant	-0,032
Lock up	-0,001

A particularly surprising observation in our analysis is the absence of the US bond market performance index among the significant variables. For a strategy that relies primarily on fixed income transactions, we would expect indicators linked to the debt market to play a crucial role. However, none of these indicators were identified as having a significant impact. This raises the hypothesis that hedge funds may trade relatively independently of the bond market conditions in which they operate. In other words, the price differences they exploit appear to stem from factors other than the economic phases of the bond market, suggesting a relative independence from market cycles.

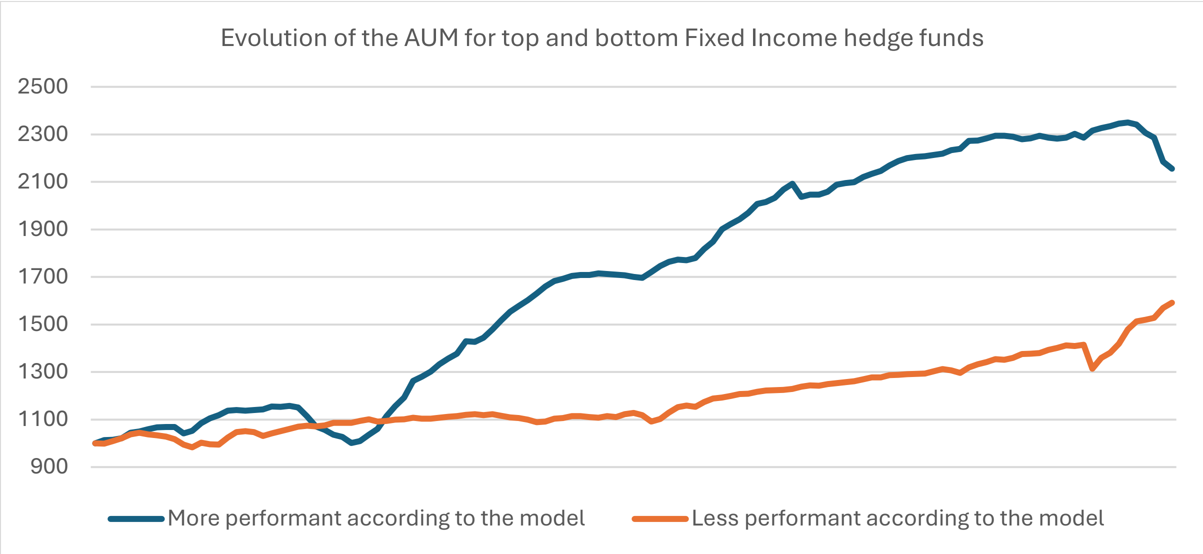
Another significant observation is the negative effect of the emerging market equity index on the two parameters studied. This result could potentially highlight a shortcoming in our model. Indeed, the databases from which we extracted the macroeconomic data did not provide sufficiently contemporaneous information to accurately model the performance of emerging country bond markets. Therefore, the fact that the variable representing the equity market in emerging countries proved to be significant could indicate that a variable representing the bond market in these same countries would also be significant. This suggests that it would be relevant to derive these specific data and include them in our model in order to test their effect on the observed parameters.

It is also interesting to note that the value of the dispersion parameter (scale) is inversely proportional to the variation in the unemployment rate. This observation marks a first in our study, showing a positive relationship between one of our parameters and a country's level of macroeconomic activity. Thus, it seems that a flourishing economy has a direct positive effect on surplus earnings, pulling them upwards and increasing their value significantly. This underlines the importance of economic conditions in optimizing hedge fund performance.

Finally, the analysis reveals that the value of the intercept is positive only for the gamma parameter, which could indicate that the fund manager has the ability to positively influence the extreme returns occurrences. The fact that the manager can influence high returns, but that the likelihood of overall positive returns remains dependent on external factors, reinforces the idea that hedge fund strategies need to be closely aligned with global economic dynamics to maximize their effectiveness.

In conclusion, our analysis reveals some intriguing results, notably the absence of the US bond market performance index among the significant variables, suggesting a relative independence of hedge funds from bond market cycles. Additionally, the negative effect of the emerging market equity index highlights a possible gap in our model, underlining the need to incorporate contemporary data to better assess the impact of emerging bond markets. Furthermore, the relationship between the dispersion parameter and the unemployment rate shows the positive influence of a booming economy on excess returns. Finally, the ability of managers to influence massive returns, rather than the occurrence of positive returns, reaffirms the importance of macroeconomic conditions in hedge fund fixed income performance. These observations suggest that hedge fund strategies need to be finely tuned to economic dynamics to maximize their success.

The evaluation of our model confirms that hedge funds adopting a fixed income strategy and operating in a context of high employment and weak equity markets in emerging countries tend to outperform other funds in terms of performance.



	More performant	Less performant
Mean of the returns	0,64%	0,39%
Variance of the returns	0,0002	0,0001

We note that although the variance of the performing fund is higher than that of its counterpart, the difference is less marked than that observed in the other strategies. This is in line with previous observations that the Hedges Funds Fixed Income market is less volatile.

Event-driven strategies

When hedge fund managers seek to capitalize on price inefficiencies that arise from corporate events, they employ what is known as an event-driven strategy. This approach encompasses a broad range of corporate events, including but not limited to mergers and acquisitions, spin-offs, financial distress situations, and other significant business developments. Within this strategy, several sub-strategies exist. One such approach is distressed debt investing, where managers invest in the debt of companies that are facing financial difficulties. Another is event-driven arbitrage, where managers exploit discrepancies in equity prices triggered by corporate events. Additionally, there are multi-strategies that combine both equity and debt investing, allowing managers to leverage opportunities across multiple asset classes affected by these corporate events. This diversified approach enables hedge fund managers to adapt their strategies to the specific nature of the corporate events and the market conditions they encounter.

The table below summarise the results of the analysis:

Table 5 - Coefficients of the unpenalized regression on the unscaled significative and binary variables for the Event Driven strategy.

	Sigma
Significant variables	
Intercept	-0,172
AUM	-0,019
Minimum investment size	-0,011
Subscription frequency	0,001
DEF	0,023
UNRATE	-0,023
S&P 500 dividend yield	0,390
Binary variables	
Closed	-0,192
Dead	-0,049
Hurdle rate	-0,125
High Water Mark	-0,062
Leverage	-0,014
UCITS Compliant	-0,589
Lock up	0,220

	Gamma
Significant variables	
Intercept	0,318
Binary variables	
Closed	-0,016
Dead	0,225
Hurdle rate	-0,101
High Water Mark	-0,088
Leverage	0,055
UCITS Compliant	-0,518
Lock up	-0,065

This last strategy differs from the others in the number of variables impacting the scale parameter. In fact, six variables have been selected: three relating to the structure of the funds, one relating to the macroeconomic context, and the last two modeling the state of the financial markets. The number of variables selected shows that the event-driven strategy is more sensitive to its environment and to the intrinsic characteristics of the investment vehicle applying it. This can be attributed to the dependency of returns on a series of corporate events, which are themselves influenced by the economic situation in which the fund operates. Logically enough, bankruptcy recovery operations are more likely to occur during a recession. Our study therefore provides a link between macroeconomic conditions and the profits generated by hedge funds applying an event-driven strategy.

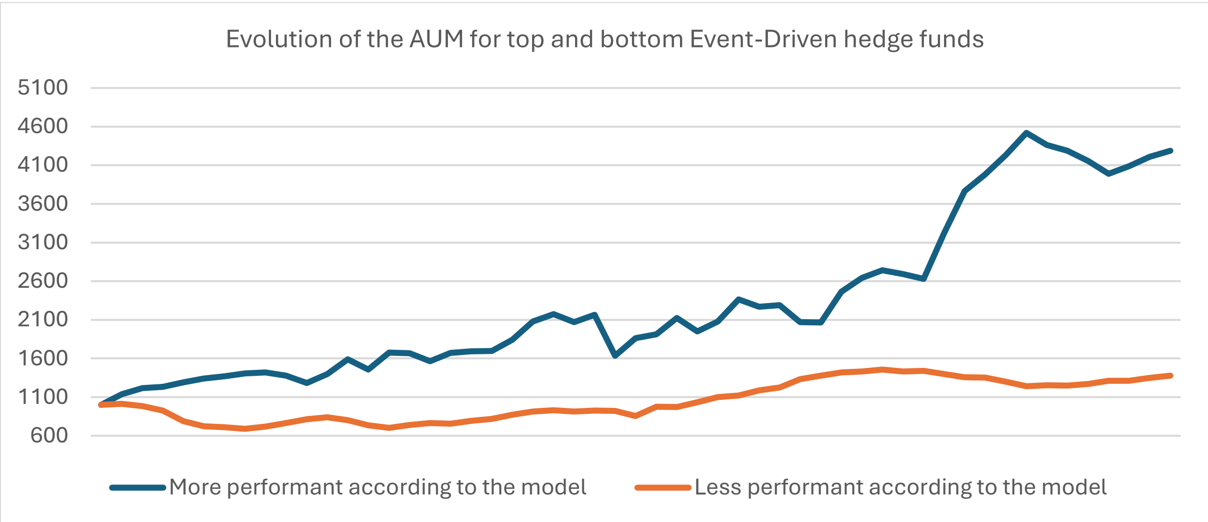
One of the external variables is the unemployment rate. The latter is negatively correlated with the sigma parameter, which means that the dispersion of profits seems to be greater when the employment rate increases. Assuming that this measure accurately reflects a country's level of activity, we can assume that, just like funds applying a fixed income strategy, funds operating using an event-driven strategy will potentially perform better if we find ourselves in a period of economic expansion.

Financial variables include the S&P500 dividend measure and the default spread. The two variables are consistent in this context because they are both derived from financial instruments issued by companies. Indeed, the cumulative measure of S&P500 dividends can be a good indicator of the economic health of US large caps, as the more they prosper, the more likely they are to pay a dividend to their shareholders. The positive relationship between the two can therefore be interpreted as evidence that the abnormal performance of event-driven hedge funds increases during an economic expansion. This can be explained by the fact that this good health will generate a series of corporate events, giving hedge fund managers an opportunity to make profits. Although this trend seems clear, the sign of the default spread coefficient is a reminder of the ability of hedge funds to make profits in all possible configurations. An increase in the default spread means that investors see an increase in credit risk. Consequently, the positive coefficient means that abnormal profits will be more likely when investors are more risk averse. This risk aversion can be a warning sign of a series of corporate events such as bankruptcies or liquidations, a boon for hedge fund managers.

Now that we have analyzed the impact of the external variables, let's look at those modeling the fund's structure. The negative coefficients allocated to both the quantity of assets under management and the minimum investment size seem to imply that it is more difficult for large funds to generate significant revenues by following an event-driven strategy. The advantage of smaller funds may come from their greater flexibility, although this remains to be defined. Finally, our analysis also reveals that the frequency with which investors enter the market also has an impact on excess returns. The longer the period of entry, the greater the likelihood of high positive profits.

In conclusion, the event-driven strategy is distinctly influenced by a combination of external macroeconomic factors and the structural characteristics of the hedge funds employing it. The selection of six significant variables underscores the strategy's sensitivity to both market conditions and the fund's inherent attributes. The analysis reveals that while economic expansion tends to enhance the performance of these funds, particularly through increased corporate events, the strategy also capitalizes on periods of heightened risk aversion, as indicated by the positive correlation with the default spread. Additionally, smaller funds with greater flexibility, as well as those with longer investor entry periods, seem better positioned to achieve higher abnormal profits, highlighting the nuanced nature of this investment approach.

We used the same procedure as previously presented to test the results obtained. However, as the number of significant variables was too high, the restrictions applied did not allow all the variables to be taken into account. We therefore had to restrict ourselves to the four variables with the highest absolute value of the standardized regression coefficient, namely the variable modeling the size of the fund, the default spread, the size of the minimum investment, and the measure of S&P500 dividends.



	More performant	Less performant
Mean of the returns	3,18%	0,74%
Variance of the returns	0,007	0,002

Our verification mechanism shows that the funds that our model indicates as being better performers outperform the others.

III. Significant variables interpretation

The final paragraphs describe the impact of the variables selected by the regularisation procedure on the various parameters, for each strategies considered. In parallel to this set of variables, our model also allowed us to highlight the impact of the binary structural variables on these same parameters. Although these variables have not been regularised, for reasons explained above, we will still analyse them, even though it is impossible to determine whether their impact is significant or not. It is for this purpose that the paragraphs below are intended.

Table 6 - Signs of the coefficients of the binary variables obtained by non-penalised regression

SIGMA	Long-Short	CTA	Fixed Income	Event Driven
Closed	-	+	-	-
Dead	-	-	-	-
Hurdle Rate	+	-	-	-
High Water Mark	+	+	-	-
Leverage	+	-	+	-
UCITS Compliants	-	-	-	-
Lock up	-	-	+	+

GAMMA	Long-Short	CTA	Fixed Income	Event Driven
Closed	+	-	-	-
Dead	-	-	+	-
Hurdle Rate	-	+	-	-
High Water Mark	+	+	+	-
Leverage	+	-	+	+
UCITS Compliants	-	+	-	-
Lock up	+	-	-	-

Firstly, we note that the fact that a fund is inactive will almost always have a negative impact on both parameters, as will compliance with the UCITS regulatory framework. This highlights the constraint that the regulatory framework represents for hedge funds, which is consistent with the scientific literature. Additionally, the overall negative effect of using hurdle rates is noted. This effect, already highlighted in the literature review, is confirmed by our analysis. It shows that when managers are constrained by legal or managerial restrictions, the probability of high returns tends to decrease.

Conversely, the application of leverage generally exhibits positive effects, though the extent of this influence varies significantly depending on the strategy employed. For instance, leverage is typically advantageous for long-short equity and fixed income strategies, as it enhances returns by amplifying gains through strategic borrowing. However, for event-driven strategies, the impact of leverage is more nuanced. While leverage positively influences the recurrence of surplus profits, indicating a higher likelihood of consistently strong returns, it also contributes to an increase in the dispersion of those returns. This heightened volatility introduces an element of risk, where profits may be less predictable, even if they are recurrent. On the other hand, for funds utilizing CTA (Commodity Trading Advisor) strategies, leverage is purely detrimental, reducing the chances of generating abnormally high returns. This suggests that borrowing in this context can exacerbate losses rather than magnifying gains.

Additionally, mechanisms such as the high-water mark and lock-up period play distinct roles across various strategies. In market-neutral strategies, the implementation of a high-water mark—designed to ensure that fund managers only earn performance fees after surpassing previous peaks—has a positive effect, raising the probability of overperformance by aligning incentives and promoting risk management. This mechanism ensures that fund managers are incentivized to recover from losses before they can benefit from further gains, which stabilizes the fund's long-term performance.

However, for event-driven strategies, the high-water mark presents a contrasting effect. It appears to curb the occurrence of surplus profits, potentially due to the nature of these strategies, which are often dependent on specific corporate events that may not always align well with the high-water mark's performance-based incentives. The lock-up period—during which investors are restricted from redeeming their shares—also exerts a noticeable influence on the dynamics of fund performance. Notably, the lock-up period has been found to have a negative impact on the emergence of extreme returns across multiple strategies. This could be attributed to the fact that restricting redemptions tends to stabilize the fund, preventing sudden liquidity crunches, but at the same time, it curtails opportunities for outsized gains by limiting the fund manager's ability to respond swiftly to market changes.

In conclusion, our analysis confirms that inactive funds, compliance with the UCITS regulatory framework, and the use of hurdle rates have an overall negative impact on hedge fund performance, in line with existing literature. On the other hand, leverage tends to boost returns, particularly for long-short and fixed income strategies, although its impact is more nuanced for event-driven and CTA strategies. Finally, the effects of the high-water mark and lock-up period mechanisms vary from one strategy to another, highlighting the complexity of their influence on extreme returns.

Conclusions and discussions

Hedge funds have always been a conundrum for the scientific community, balancing between exceptional profits and staggering losses. They present a risk-return profile unmatched by any other asset class. At the core of these varying performances is the hedge fund manager. While often criticized, scientific literature acknowledges their critical role, demonstrating that they alone have the potential to significantly enhance fund performance. Our research aimed to explore whether the fund’s environment and structure enable the manager to generate surplus profits. We approached this by modeling these factors and estimating managers' capacities using Extreme Value Theory—a key component of our study. This methodology allowed us to link the distribution of extreme returns with various explanatory variables. By employing penalized regression and optimizing penalty parameters, we were able to identify a set of significant variables for each hedge fund strategy. Analyzing these variables provided several insights.

Firstly, our analysis indicates a limited impact of the fund's environment and structure on the occurrence of abnormally positive returns, emphasizing the manager's predominant role. However, the influence of these variables on the distribution of returns is considerably more pronounced. This observation led us to evaluate each strategy separately, revealing the intricate relationship between macroeconomic conditions and fund-specific factors.

Market-neutral funds, for instance, excel during economic downturns, leveraging short-selling strategies to benefit from negative economic indicators such as term spreads and GDP growth. Similarly, CTA hedge funds thrive in recessionary environments, where factors like declining T-bill rates and GDP growth boost their ability to generate excess returns through quantitative models. However, this comes with a trade-off between the frequency of profits and their dispersion. Fixed Income strategies, surprisingly, show a degree of independence from traditional bond market cycles but remain sensitive to broader economic conditions, with economic booms positively influencing returns. Event-driven strategies, on the other hand, demonstrate a nuanced dependence on both macroeconomic factors and fund structure, flourishing in periods of economic expansion and heightened risk aversion, particularly through smaller, more agile funds.

These insights underscore the importance of aligning hedge fund strategies with prevailing economic dynamics to maximize performance. Our analysis covers four prevalent hedge fund strategies, providing a broad perspective that will be valuable to many hedge fund managers. Beyond highlighting the critical roles of fund structure and macroeconomic context, our findings delve into specific variables that significantly influence fund performance. By identifying these variables and assessing their magnitude and nature, this work offers practical guidance for managers aiming to refine their strategies and improve the likelihood of achieving extreme positive returns.

Our research lays a solid foundation for future studies in the hedge fund industry. Managers can use these insights to develop more accurate prediction models or enhance existing ones, enabling better anticipation of market movements and more strategic resource allocation. This work provides both theoretical insights and practical tools for optimizing hedge fund performance in an increasingly complex financial environment.

Moreover, this study enriches the existing literature on hedge funds by focusing specifically on abnormal positive returns—a niche area not extensively covered in current research. This targeted exploration sheds new light on the dynamics and strategies that lead to excessive profits, contributing to a deeper understanding of hedge fund performance.

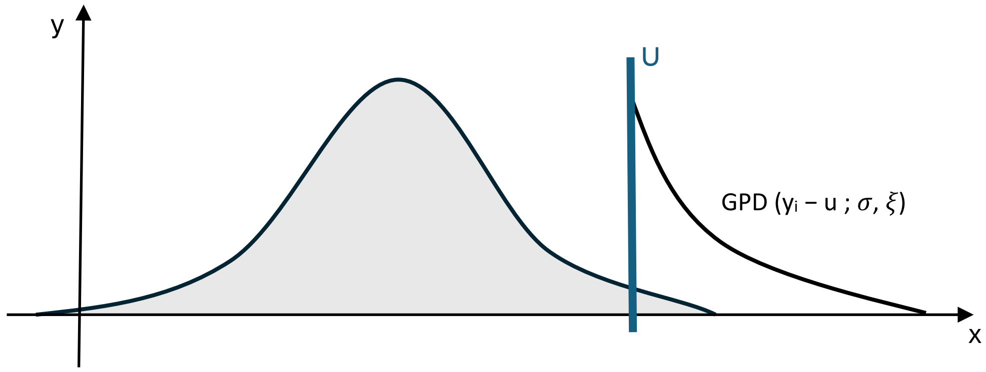
Despite its contributions, this research has limitations. The lack of observations for less common strategies prevents generalization of the conclusions drawn. Hedge funds, being among the most complex asset classes, are influenced by a multitude of factors, and it is presumptuous to claim that the variables considered capture all relevant dynamics perfectly. Additionally, a mechanism to rigorously test the accuracy of our results and conclusions was not available.

Nevertheless, this work opens avenues for further research. For instance, employing bootstrap procedures could extend the number of observations and explore less common strategies. Investigating the symmetry of the impact of significant variables on both extreme returns and large losses could provide additional insights. Finally, the methodology used could be applied to quantify the impact of fund managers' skills more precisely.

I hope this research will inspire other students at HEC Liège and beyond to delve into the fascinating world of hedge funds.

Appendix

Chart 1: The Peak-Over-Threshold approach



With U being the threshold selected, the blue curve being the sample distribution and the black curve on the right being the general pareto distribution. This graph is a representation of the situation and is not accurate.

Chart 2: Boxplot of the number of observations per month

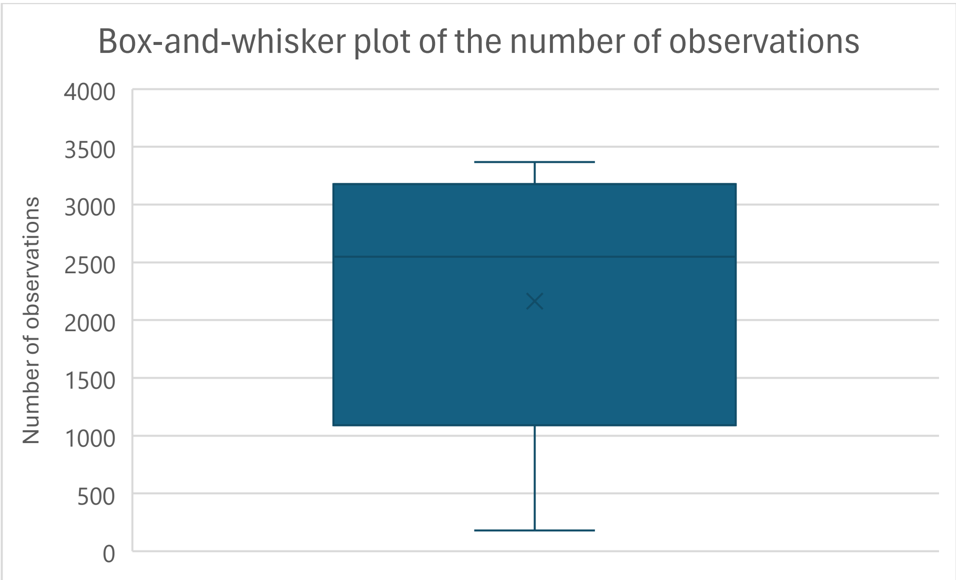


Table 1: Variables describing the hedge fund structure:

Variables	Description
AUM	Asset under management, total market value of the assets managed by the funds expressed in 1000$.
Closed	Binary variable: If "Yes": Fund doesn't accept investment anymore.
Dead	Binary variable: If "Yes": Fund is active.
Hurdle Rate	Minimum rate of return that needs to be reached if the manager wants to apply performance fees.
High Water Mark	Binary variable: If "Yes" the hedge funds use high water marks.
Min. Investment Size	Size of the minimum investment.
Leverage	Binary variable: If "Yes": The hedge fund uses leveraged strategies.
UCITS Compliant	Binary Variable: If "Yes": The Hedge Fund complies with UCITS rules
Subscription frequency	Refers to the period when the investor can invest capital into the fund
Redemption Notification period	If an investor wants to withdraw its capital from the Hedge Fund, he has to notify the manager and then wait the redemption notification period before being able to do so.
Redemption frequency	The frequency at which the investor can withdraw their capital from the hedge fund.
Lock up	Period during which the investor cannot withdraw its capital.
Management fees	It is a fee applied by the manager to manage the fund. It is a percentage of the total value of the fund.
Performance fees	It is a fee based on the fund performance. These are applied only when a certain level of return is reached (hurdle rate) and must be computed on the total added value.
Age	Difference between the date and the inception date in months.

Table 2: The different strategies considered in the initial database :

Strategy	Description
Bottom-up	The strategy is to assess the value of a single investment which could perform well, regardless of the macroeconomic situation.
Dual approach	The fund investment strategy approach is dual.
Diversified Debt	The fund invests exclusively in debt instruments (bonds, MBS, ABS, …)
Long Short Equities	The manager takes opposite positions on the same stock. It allows him to reach market neutrality.
Value	Investment is made in stocks that are assumed to be under/overvalued.
Event-Driven	The strategy is to capitalize on events such as merger, arbitrage, bankrupt, spin-off, …
Top-Down	Kind of the opposite of the bottom-up strategy. The focus is made on the investment sector and then on the individual investment.
Arbitrage	The aim of the strategy is to take advantages of pricing discrepancies
Fixed Income	Investment exclusively made on fixed income securities.
Relative Value	The manager Takes advantage of perceived mispricing or valuation discrepancies between related financial instruments.
CTA/Managed Futures	It involves trading financial and commodity futures contracts based on systematic and quantitative models.
Macro	Investment decisions based on macroeconomic conditions.
Distressed Debts	The focus is made on investment in the debt securities of companies or issuers facing financial distress or undergoing a restructuring process.
Others	The fund uses another strategy than the ones referenced in the database.
Multi Strategy	The funds use a mix of one or many strategies.

Table 3: Description of the macroeconomic and financial variables:

Variables	Description	Source	Motivations
DEF	Default spread: Difference between yields on BAA and AAA rated corporate bonds. A high default spread means that investor want higher compensation to take extra risk. Good proxy of the investor risk aversion.	Federal Reserve Bank of Saint Louis	This variable has been introduced based on the article of Bali et al. (2014).
DIV	Aggregated Dividend Yield on the S&P500. Widely use in the literature. It has been proven that it has an impact on other macroeconomic measurements (Serfling & Miljkovic, 2011).	S&P Capital IQ	This variable has been introduced based on the article of Bali et al. (2014) and Lambert and Platania (2020).
GDP	Quarterly Growth rate of the US GDP per capita. Widely use in the literature. Proxy of the economic activity of the US.	Federal Reserve Bank of Saint Louis	This variable has been introduced based on the article of Bali and al. (2014) and Lambert and Platania (2020)
INF	Inflation rate based on the consumer price. Widely use in the literature. Inflation will have an impact on the real rates of the different markets and therefore need to be considered.	Federal Reserve Bank of Saint Louis	This variable has been introduced based on the article of Bali et al. (2014).
UNEMP	Unemployment rate in the US. Widely use in the literature. Indicator of the economic health of a country.	Federal Reserve Bank of Saint Louis	This variable has been introduced based on the article of Bali et al. (2014).
RREL	Relative T-Bill rate: difference between the 3-month T-bill rate and the 12-month backward moving average. Provides valuable insights into risk-free returns, monetary policy, economic outlook, market sentiment, liquidity conditions, and inflation expectations.	Federal Reserve Bank of Saint Louis	This variable has been introduced based on the article of Bali et al. (2014) and Lambert and Platania (2020).
VIX	Volatility of the S&P500 index. Captures the volatility of one of the largest indices in the US. Good proxy for the financial situation of the biggest US capitalization.	Yahoo finance	Aim to fulfil the lack of financial variables in the model.
SP500	Returns of the S&P500 index. Captures the performance of one of the largest indices in the US. Good proxy for the financial situation of the biggest US capitalization.	Investing.com	Aim to fulfil the lack of financial variables in the model and to cover US equity market.
RU2000	Returns of the Russel 2000 index. Stock market index that measures the performance of approximately 2,000 small-cap companies in the United States	Yahoo finance	Aim to fulfil the lack of financial variables in the model and to cover US equity market.
INDPROD	Monthly industrial production of the US. This variable is used to complete the proxy of the economic health of the US.	Federal Reserve Bank of Saint Louis	Aim to offer a better coverage of the macroeconomic situation.
MSCI WORLD	Return of the MSCI World Index. It is an index encompassing mid and large capitalization across the world. It is used as a proxy of the world financial market.	Morningstar direct	Aim to fulfil the lack of financial variables in the model and to cover world equity market.
T10Y3M	It is the term spread expressed as the difference between yields on 10-year and 3-month Treasury securities. It is a crucial indicator since a positive value is often sign of economic growth whereas a negative value might indicate recession.	Federal Reserve Bank of Saint Louis	This variable has been introduced based on the article of Bali et al. (2014).
S&P GSCI	Goldman Sachs Commodity Index. Widely recognized benchmark for investment in the commodity markets.	S&P Capital IQ	Aim to model the global commodity market.
US bond	Index tracking performance of the US bond market	S&P Capital IQ	Aim to model the US bond market.

S&P Emerging BMI	S&P Emerging Broad Market Index. Equity index that measures the performance of publicly traded companies in emerging markets.	S&P Capital IQ	Aim to model the emerging countries equity market.

Table 4: Results of the analysis

Regularization procedure

LASSO REGRESSION				
Gamma then Sigma	Long-Short Equities	CTA	Fixed Income	Event-driven
λ(σ)	*0,10*	*0,31*	*0,43*	*0,25*
λ(γ)	*0,16*	*0,16*	*0,21*	*0,16*
BIC	*8378,02*	*5302,11*	*3125,47*	*2833,93*
Sigma then Gamma				
λ(σ)	*0,01*	*0,19*	*0,18*	*0,20*
λ(γ)	*0,07*	*0,30*	*0,5*	*0,24*
BIC	*8452,54*	*5255,33*	*3095,56*	*2803,65*
ADAPTIVE LASSO REGRESSION				
Gamma then Sigma	Long-Short Equities	CTA	Fixed Income	Event-driven
λ(σ)	*0,006*	*0,039*	*0,0077*	*0,0065*
λ(γ)	*0,02*	*0,0091*	*0,011*	*0,35*
BIC	***8353,28***	*5263,98*	*3088*	***2786,30***
Sigma then Gamma				
λ(σ)	*0,0099*	*0,025*	*0,022*	*0,044*
λ(γ)	*0,0051*	*0,022*	*0,03*	*0,039*
BIC	*8377,98*	***5232,31***	***3073,3***	*2803,758*

Coefficients of the regression executed on the standardized variables

Long-short equity:

For the scale parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	-0,268
Growth RGDP per capita.	-0,052
T10Y3M	-0,006
BINARY	
Closed	-0,207
Dead	-0,044
Hurdle.Rate	0,142
High.Water.Mark	0,066
Leverage	0,100
UCITS.Compliant	-0,819
Lock.up	-0,082

For the shape parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	0,006
Term spread	-0,041
BINARY	
Closed	0,117
Dead	-0,025
Hurdle.Rate	-0,048
High.Water.Mark	0,120
Leverage	0,055
UCITS.Compliant	-0,515
Lock.up	0,116

CTA:

For the scale parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	-0,236
Redemption Notification Period	0,064
Growth RGDP per capita.	-0,143
Relative T-bill rate	0,093
BINARY	
Closed	0,362
Dead	0,073
Hurdle.Rate	-0,029
High.Water.Mark	-0,083
Leverage	0,085
UCITS.Compliant	-1,199
Lock.up	0,182

For the shape parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	0,256
BINARY	
Closed	-0,067
Dead	-0,013
Hurdle.Rate	-0,210
High.Water.Mark	0,066
Leverage	-0,037
UCITS.Compliant	0,192
Lock.up	-0,081

Fixed Income:

For the scale parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	-0,087
UNRATE	-0,107
SP Emerging BMI	-0,112
BINARY	
Closed	-0,076
Dead	-0,020
Hurdle.Rate	0,014
High.Water.Mark	-0,104
Leverage	-0,010
UCITS.Compliant	-0,013
Lock.up	-0,014

For the shape parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	0,095
SP Emrging BMI Index	0,101
BINARY	
Closed	-0,035
Dead	0,093
Hurdle.Rate	-0,007
High.Water.Mark	-0,004
Leverage	0,051
UCITS.Compliant	-0,025
Lock.up	0,086

Event Driven:

For the scale parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	-0,27389376
AUM	-0,26406411
Minimum.Investment.Size	-0,13968146
Subscription.Frequency	0,03975705
DEF	0,14931625
UNRATE	0,02277880
S.P500.div.yield	-0,11313775
BINARY	
Closed	0,18568019
Dead	0,14533573
Hurdle.Rate	-0,41019392
High.Water.Mark	-0,31258893
Leverage	0,39152472
UCITS.Compliant	-1,61533524
Lock.up	0,07555528

For the shape parameter:

Variables	Coefficients
SIGNIFICANT	
Intercept	0,202507
BINARY	
Closed	-0,386845
Dead	0,243604
Hurdle.Rate	-0,043188
High.Water.Mark	0,086248
Leverage	-0,087578
UCITS.Compliant	-1,106277
Lock.up	-0,105705

Bibliography

Ackermann, C., McEnally, R., & Ravenscraft, D. (1999). The Performance of Hedge Funds: Risk, Return, and Incentives. *The Journal of Finance (New York)*, *54*(3), 833–874. https://doi.org/10.1111/0022-1082.00129

Agarwal, V., & Naik, N. Y. (2000). Multi-Period Performance Persistence Analysis of Hedge Funds. *Journal Of Financial And Quantitative Analysis*, *35*(3), 327. https://doi.org/10.2307/2676207

Agarwal, V., & Naik, N. Y. (2003). Risks and Portfolio Decisions Involving Hedge Funds. *The Review Of Financial Studies*, *17*(1), 63-98. https://doi.org/10.1093/rfs/hhg044

Agarwal, V., Mullally, K. A., & Naik, N. Y. (2015). Hedge fund performance. Dans *The Economics and Finance of Hedge Funds* (1re éd., Vol. 10).

Agarwal, V., Ruenzi, S., & Weigert, F. (2017). Tail risk in hedge funds: A unique view from portfolio holdings. *Journal of Financial Economics*, *125*(3), 610–636. https://doi.org/10.1016/j.jfineco.2017.06.006

Amenc, N., Martellini, L., & Chardin, T.D. (2003). The Alpha and Omega of Hedge Fund Performance Measurement.

Amin, G. S., & Kat, H. M. (2003). Hedge Fund Performance 1990-2000 : Do the « Money Machines » Really Add Value ? *Journal Of Financial And Quantitative Analysis*, *38*(2), 251. https://doi.org/10.2307/4126750

Ammann, M., & Moerth, P. (2005). Impact of fund size on hedge fund performance. *Journal Of Asset Management*, *6*(3), 219-238. https://doi.org/10.1057/palgrave.jam.2240177

Anson, M. J. P. (2002). A Primer on Distressed Debt Investing. *The Journal of Private Equity*, *5*(3), 6–17. http://www.jstor.org/stable/43503295

Auleta, O., & Stefanini, F. (2017). Directional Equity Strategies of Hedge Funds. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0011

Aurum. (2023). Hedge fund industry deep dive. Dans *Aurum*. Consulté le 8 juin 2024, à l'adresse https://www.aurum.com/wp-content/uploads/Aurum-Hedge-Fund-Industry-Deep-Dive_Q3-2023-Review.pdf

Baker, H. K., & Filbeck, G. (2017). Hedge funds Fundamentals. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0001

Bali, T. G., Atilgan, Y., & Demirtas, K. O. (2013). Hedge fund strategies. Dans *Investing in Hedge Funds* (p. 25-49). https://doi.org/10.1016/b978-0-12-404731-0.00002-7

Bali, T. G., Brown, S. J., & Caglayan, M. O. (2014). Macroeconomic risk and hedge fund returns. *Journal Of Financial Economics*, *114*(1), 1-19. https://doi.org/10.1016/j.jfineco.2014.06.008

Bollen, N. P. B., & Pool, V. K. (2008). Conditional Return Smoothing in the Hedge Fund Industry. *Journal Of Financial And Quantitative Analysis*, *43*(2), 267-298. https://doi.org/10.1017/s0022109000003525

Bui, T., & Ganguly, A. (2017). Issues in Hedge Fund Research. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0021

Cai, B., Cheng, T., & Yan, C. (2018). Time-varying skills (versus luck) in U.S. active mutual funds and hedge funds. *Journal Of Empirical Finance*, *49*, 81-106. https://doi.org/10.1016/j.jempfin.2018.09.001

Cao, C., Goldie, B. A., Liang, B., & Petrasek, L. (2016). What Is the Nature of Hedge Fund Manager Skills ? Evidence from the Risk-Arbitrage Strategy. *Journal Of Financial And Quantitative Analysis*, *51*(3), 929-957. https://doi.org/10.1017/s0022109016000387

Capocci, D. (2013). *The Complete Guide to Hedge Funds and Hedge Fund Strategies*. Palgrave Macmillan.

Capocci, D., & Hübner, G. (2004). Analysis of hedge fund performance. *Journal Of Empirical Finance*, *11*(1), 55-89. https://doi.org/10.1016/j.jempfin.2002.12.002

Capocci, D., Corhay, A., & Hübner, G. (2005). Hedge fund performance and persistence in bull and bear markets. *European Journal Of Finance*, *11*(5), 361-392. https://doi.org/10.1080/1351847042000286676

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal Of Finance*, *52*(1), 57. https://doi.org/10.2307/2329556

Cave, A., Hubner, G., & Sougne, D. (2011). The market timing skills of hedge funds during the financial crisis. *Managerial Finance*, *38*(1), 4-26. https://doi.org/10.1108/03074351211188330

Chavez-Demoulin, V., Embrechts, P., & Hofert, M. (2015). An Extreme Value Approach for Modeling Operational Risk Losses Depending on Covariates. *The Journal Of Risk And Insurance*, *83*(3), 735-776. https://doi.org/10.1111/jori.1205

Chen, Y., Cliff, M., & Zhao, H. (2017). Hedge Funds : The Good, the Bad, and the Lucky. *Journal Of Financial And Quantitative Analysis*, *52*(3), 1081-1109. https://doi.org/10.1017/s0022109017000217

Connor, G. and Lasarte, T. (2003) An Overview of Hedge Fund Strategies. https://www.academia.edu/69221656/An_Overview_of_Hedge_Fund_Strategies

Connor, G., & Woo, M. (2004). *An Introduction to hedge funds*. Financial Markets Group, The London School of Economics and Political Science. https://eprints.lse.ac.uk/24675/

Cumming, D., Dai, N., & Johan, S. A. (2013). Introduction to Hedge Funds. Dans *Hedge Fund Structure, Regulation, and Performance Around the World* (p. 1-20). https://doi.org/10.1093/acprof:oso/9780199862566.003.0001

De Luca, P. (2023). Private Equity, Venture Capital, and Hedge Fund. Dans *Corporate Finance* (p. 608-611). https://doi.org/10.1007/978-3-031-18300-3

Dikanarov, G., McBride, J., & Spieler, A. C. (2017). Relative value hedge fund strategies. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0014

Edwards, F. R., & Caglayan, M. O. (2001). Hedge Fund Performance and Manager Skill. *The Journal Of Futures Markets*, *21*(11), 1003-1028. https://doi.org/10.1002/fut.2102

Eurekahedge. (s. d.). *Long short Equities Hedge fund index*. Consulté le 8 juin 2024, à l'adresse https://www.eurekahedge.com/Indices/IndexView/Eurekahedge/480/Eurekahedge-Long-Short-Equities-Hedge-Fund-Index

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal Of Financial Economics*, *33*(1), 3-56. https://doi.org/10.1016/0304-405x(93)90023-5

Fama, E. F., & French, K. R. (1998). Value versus Growth : The International Evidence. *The Journal Of Finance*, *53*(6), 1975-1999. https://doi.org/10.1111/0022-1082.00080

Fan, J., & Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association*, *96*(456), 1348–1360. https://doi.org/10.1198/016214501753382273

Farrelly, C., & Lhabitant, F. (2017). Event-Driven hedge Fund strategies. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0012

Fung, W., & Hsieh, D. (2006). Hedge funds : an industry in its adolescence. *Federal Reserve Bank Of Atlanta Economic Review*, *91*, 1-34. https://www.atlantafed.org/-/media/documents/research/publications/economic-review/2006/vol91no4_fung-hsieh.pdf

Fung, W., & Hsieh, D. A. (1997). Empirical Characteristics of Dynamic Trading Strategies : The Case of Hedge Funds. *The Review Of Financial Studies*, *10*(2), 275-302. https://doi.org/10.1093/rfs/10.2.275

Fung, W., & Hsieh, D. A. (1999). A primer on hedge funds. *Journal Of Empirical Finance*, *6*(3), 309-331. https://doi.org/10.1016/s0927-5398(99)00006-7

Fung, W., & Hsieh, D. A. (2000). Measuring the market impact of hedge funds. *Journal of Empirical Finance*, *7*(1), 1–36. https://doi.org/10.1016/S0927-5398(00)00005-0

Fung, W., Hsieh, D. A., Naik, N. Y., & Ramadorai, T. (2008). Hedge Funds : Performance, Risk, and Capital Formation. *The Journal Of Finance*, *63*(4), 1777-1803. https://doi.org/10.1111/j.1540-6261.2008.01374.x

Getmansky, M., Lo, A. W., & Makarov, I. (2004). An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal Of Financial Economics*, *74*(3), 529-609. https://doi.org/10.1016/j.jfineco.2004.04.001

Goetzmann, W. N., Ingersoll, J. E., & Ross, S. A. (2003). High-Water Marks and Hedge Fund Management Contracts. *The Journal of Finance (New York)*, *58*(4), 1685–1718. https://doi.org/10.1111/1540-6261.00581

Hambuckers, J., Groll, A., & Kneib, T. (2018). Understanding the economic determinants of the severity of operational losses : A regularized generalized Pareto regression approach. *Journal Of Applied Econometrics*, *33*(6), 898-935. https://doi.org/10.1002/jae.2638

Hedge Fund Research. (2024, 6 mai). *HFR Hedge Fund Strategy Classification System | HFR*. HFR. Consulté le 8 juin 2024, à l'adresse https://www.hfr.com/hfr-indices/hfr-hedge-fund-strategy-classifications/

Hossain, A. T., Saadi, S., & Treff, M. (2017). Characteristics of Successful Hedge Fund Managers. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0004

Huang, J., Liechty, J., & Rossi, M. (2018). Return smoothing and its implications for performance analysis of hedge funds. *The Journal Of Finance And Data Science*, *4*(4), 203-222. https://doi.org/10.1016/j.jfds.2018.05.002

Ibbotson, R. G., Chen, P., & Zhu, K. X. (2011). The ABCs of Hedge Funds : Alphas, Betas, and Costs. *Financial Analysts Journal*, *67*(1), 15-25. https://doi.org/10.2469/faj.v67.n1.6

Jagannathan, R., Malakhov, A., & Novikov, D. (2010). Do Hot Hands Exist among Hedge Fund Managers ? An Empirical Evaluation. *The Journal Of Finance*, *65*(1), 217-255. https://doi.org/10.1111/j.1540-6261.2009.01528.x

Jansen, A., Mathieson, D., Eichengreen, B., Kodres, L., Chadha, B., & Sharma, S. (1998). Hedge Funds and Financial Market Dynamics. Dans *Occasional paper/Occasional paper*. International Monetary Fund. https://doi.org/10.5089/9781557757364.084

Jones, M. (2007). Examination of fund age and size and its impact on hedge fund performance. *Derivatives Use, Trading & Regulation*, *12*(4), 342-350. https://doi.org/10.1057/palgrave.dutr.1850052

Karagiorgis, A., & Drakos, K. (2022). The Skewness-Kurtosis plane for non-Gaussian systems : The case of hedge fund returns. *Journal Of International Financial Markets, Institutions & Money*, *80*, 101639. https://doi.org/10.1016/j.intfin.2022.101639

Kat, H. M., & Brooks, C. (2001). The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors. *Social Science Research Network*. https://doi.org/10.2139/ssrn.289299

Kosowski, R., Naik, N., & Teo, M. (2007). Do hedge funds deliver alpha ? A Bayesian and bootstrap analysis. *Journal Of Financial Economics*, *84*(1), 229-264. https://doi.org/10.1016/j.jfineco.2005.12.009

Lambert, M. (2012). Hedge Fund Market Risk Exposures : A survey. *Finance*, *33*(1). https://hdl.handle.net/2268/117581

Lambert, M., & Platania, F. (2020). The macroeconomic drivers in hedge fund beta management. *Economic Modelling*, *91*, 65-80. https://doi.org/10.1016/j.econmod.2020.04.016

Lhabitant, F.-S. (2006). *Handbook of hedge funds*. John Wiley.

Liang, B. (1998). On the Performance of Hedge Funds. *Social Science Research Network*. https://doi.org/10.2139/ssrn.89490

Liang, B. (2000). Hedge Funds : The Living and the Dead. *Journal Of Financial And Quantitative Analysis*, *35*(3), 309.

Ling, Y., Satchell, S., & Yao, J. (2023). Decreasing returns to scale and skill in hedge funds. *Journal Of Banking & Finance*, *156*, 107009. https://doi.org/10.1016/j.jbankfin.2023.107009

Lo, A. W. (2002). The Statistics of Sharpe Ratios. *Financial Analysts Journal*, *58*(4), 36-52. https://doi.org/10.2469/faj.v58.n4.2453

Longo, J. M. (2022). Hedge Funds : Overview, Strategies, and Trends. Dans *Encyclopedia of Finance* (p. 1249-1271). https://doi.org/10.1007/978-3-030-91231-4_54

Maginn, J. L., Tuttle, D. L., McLeavey, D. W., & Pinto, J. E. (2007). *Managing investment portfolios : A Dynamic Process*. John Wiley & Sons.

Malladi, R. K. (2020). Luck versus Skill in Evaluating Hedge Fund Managers' Performance. *Journal Of Business And Management*, *26*(1), 22-39. https://doi.org/10.6347/jbm.202003_26(1).0002

McMillan, M., Pinto, J. E., Pirie, W. L., & Van de Venter, G. (2011). *Investments : Principles of Portfolio and Equity Analysis*. Wiley.

Meier, L., van de Geer, S., & Bühlmann, P. (2008). group lasso for logistic regression. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *70*(1), 53–71. https://doi.org/10.1111/j.1467-9868.2007.00627.x

Metzger, N., & Shenai, V. (2019). Hedge Fund Performance during and after the Crisis : A Comparative Analysis of Strategies 2007–2017. *International Journal Of Financial Studies*, *7*(1), 15. https://doi.org/10.3390/ijfs7010015

Mhalla, L., Hambuckers, J., & Lambert, M. (2022). Extremal connectedness of hedge funds. *Journal Of Applied Econometrics*, *37*(5), 988-1009. https://doi.org/10.1002/jae.2900

Nabilou, H. (2017). The Conundrum of Hedge Fund Definition. *European Company And Financial Law Review*, *14*(1). https://doi.org/10.1515/ecfr-2017-0006

Pozen, R., & Hamacher, T. (2015). Hedge funds. Dans *The fund industry*. John Wiley & Sons.

Racicot, F., Théoret, R., & Gregoriou, G. N. (2014). The Hedge Fund Alpha Puzzle with an Application to Asian Hedge Funds. Dans *Handbook of Asian finance* (p. 343-362). https://doi.org/10.1016/b978-0-12-800986-4.00018-2

Samuelson, P. A. (2015). Proof that Properly Anticipated Prices Fluctuate Randomly. Dans *World Scientific handbook in financial economic series* (p. 25-38). https://doi.org/10.1142/9789814566926_0002

Savona, R. (2016). Hedge fund performance. Dans *Asset Management and Institutional Investors* (p. 355-371). https://doi.org/10.1007/978-3-319-32796-9_12

Schneeweis, T., & Spurgin, R. B. (1998). Multifactor Analysis of Hedge Fund, Managed Futures, and Mutual Fund Return and Risk Characteristics. *The Journal Of Alternative Investments*, *1*(2), 1-24. https://doi.org/10.3905/jai.1998.407852

Schuhmacher, F., & Eling, M. (2011). Sufficient conditions for expected utility to imply drawdown-based performance rankings. *Journal Of Banking & Finance*, *35*(9), 2311-2318. https://doi.org/10.1016/j.jbankfin.2011.01.031

Serfling, M. A., & Miljkovic, D. (2011). Time series analysis of the relationships among (macro) economic variables, the dividend yield and the price level of the S& ; P 500 Index. *Applied Financial Economics*, *21*(15), 1117-1134. https://doi.org/10.1080/09603107.2011.562167

Shin, J., Kim, M., Oh, D., & Kim, T. S. (2018). Do hedge funds time market tail risk ? Evidence from option-implied tail risk. *Journal Of Futures Markets*, *39*(2), 205-237. https://doi.org/10.1002/fut.21972

Smith, D. M. (2017). Evaluating hedge fund performance. Dans *Hedge Funds: Structure, Strategies, and Performance*. https://doi.org/10.1093/oso/9780190607371.003.0023

Stafylas, D., Anderson, K., & Uddin, M. (2016). Recent advances in hedge funds' performance attribution : Performance persistence and fundamental factors. *International Review Of Financial Analysis*, *43*, 48-61. https://doi.org/10.1016/j.irfa.2015.11.001

Stulz, R. M. (2007). Hedge Funds : Past, Present, and Future. *The Journal Of Economic Perspectives*, *21*(2), 175-194. https://doi.org/10.1257/jep.21.2.175

Swartz, L. M., & Emami-Langroodi, F. (2018). Relative Value Hedge Funds : A Behavioral Modeling of Hedge Fund Risk and Return Factors. *The Journal Of Behavioral Finance*, *19*(4), 462-482. https://doi.org/10.1080/15427560.2018.1434654

Teo, M., Koh, F., & Koh, W. T. (2003). Asian Hedge Funds : Return Persistence, Style, and Fund Characteristics. *Social Science Research Network*. https://doi.org/10.2139/ssrn.416960

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B, Methodological*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *73*(3), 273–282. https://doi.org/10.1111/j.1467-9868.2011.00771.x

Van, P. N., & Duong, K. D. (2022). The Performance Measurement of Generalized Sharpe Ratio and Economic Performance Measure : A Hedge Funds Example. *Universal Journal Of Accounting And Finance*, *10*(1), 124-130. https://doi.org/10.13189/ujaf.2022.100113

Zou, H. (2006). The Adaptive Lasso and Its Oracle Properties. *Journal Of The American Statistical Association*, *101*(476), 1418-1429. https://doi.org/10.1198/016214506000000735