Copenhagen Business School

Master Thesis

Master of Science in Business Administration and

Mathematical Business Economics

Hedge Fund Replication Before, During and After the Financial Crisis

Author:

Michelle Westphal Soerensen *(Student: 68254)*

Emil Vingaard Flindt Hansen *(Student: 76366)*

Supervisor:

Marcel Fischer, Associate Professor

Number of characters: 183,477

Number of Pages: 106

Date of Submission: January 15^{th}, 2020

CBS COPENHAGEN BUSINESS SCHOOL HANDELSHØJSKOLEN

Resumé

Hedge fonde påstår at kunne levere et højt afkast givet lederens evnen til at slå marked. Lederens evne til at slå markedet sammen med en bestemt hedge fond strategi, giver det alpha som investor betaler meget høje afgifter for, for at kunne investere i hedge fonde. Denne afgift er dog kun en af ulemperne ved hedge fonde som investorer vil opleve. Derfor er hedge fond replikation blandt andet blevet et interessant emne at undersøge. Hedge fond replikation ønsker at genskabe et produkt der kan levere et afkast der er ens med et hedge fond index eller højere. Dette kan undersøges ved brug af flere forskellige replikations tilgange.

Dette studie vil fokusere på den faktor-baseret tilgang, hvor der prøves at genskabe det samme månedlige afkast for en hedge fund over en given periode. Dette studie vil desuden tage udgangspunkt i tre forskellige perioder, opdelt i forhold til den finansielle krise i 2008, det vil sige før, under og efter den finansielle krise. Den første periode vil replikere ved hjælp af data før den finansielle krise, "Pre-crisis" perioden. Den anden periode vil replikere hvor data observationerne for den finansielle krise er medtaget, "Crisis" perioden. Den sidste periode vil replikere efter den finansielle krise, "Post-crisis" perioden.
Genskabelsen af det månedlige afkast over en given periode, gøres i dette studie ved brug af metoden "rolling window" og multiple lineær regression. Denne metode genskaber det månedlige afkast for en hedge fond ved at opstille en multiple lineær regression, hvor hældningskoefficienterne justeres over tid. Den multiple lineær regression har ingen skæring, da denne ønskes elimineres, grundet skæringen defineres som hedge fond lederens alpha.
Studiet tager udganspunkt i Eurekahedge hedge fund indekses, og prøver at replikere 11 forskellige strategier. Bl.a. medtages strategierne macro, arbitrage og event driven.

Resultaterne i dette studie, viser ikke entydigt om det er muligt at replikere hegde fond indeks, og det er derfor valgt at udvide metoden "rolling window" til en udvidet form. Denne udvidet metode giver mulighed for at justere både faktorer og beta undervejs, hvilket giver forbedret resultater, dog uden at være entydige. I studiet ændres der desuden på flere variable indenfor metoden, såsom længden på "vinduer" eller in-sample perioder, men stadig uden entydige resultater om replikation af hedge fonde i perioderne før, under og efter den finansielle krise.

Contents

1 Introduction

Hedge funds have been around for a while, but have over the past decades been more common as the requirements for investing into them have been loosened up. Before 2016 it was only investors making or having a certain amount of money that could invest in hedge funds. However, in 2016 the Fair Investment Opportunities for Professional Experts Act. changed to open up for more investors.[1] This has made hedge funds even more interesting to investigate, as their product is now available to more people and industries.
An area that has become especially interesting to look into is what happens within the hedge funds, as they are very limited with information and many investors after the latest crisis have become more aware of fully understanding their investment to ensure none too heavily correlation and to big risk exposure in their portfolios. The second is to investigate whether the high fees are actually worth paying - if the manager's skills are actually contributing to a better performance.

To enlighten these problems this study will focus on replicating different hedge fund indices. Replication of hedge fund index has been around for a while and studied by many different researchers, to see if it would be possible to provide investors with information about hedge funds, and if it could be possible to generate a similar return by the use of different replication method, to get rid of the high fees.
Even though this area has been studied by multiple researchers and by just as many different methods, we still find that certain areas of this field could be further investigated.

Although hedge fund replicating has been around for a while and Jaeger and Wagner(2005) found that only 20% of the hedge fund returns are due to the manager's skills or other risk factors that are unknown, and 80% of the return are due to beta exposure, hedge fund clones have still not been 100% accepted by the investors. Some still doubt that clones will actually deliver the same product as hedge funds.

[1] Baker, H. & Filbeck, G. (2017)

This study will therefore include an analysis on the factor-based approach but also include historical performance on hedge fund and different hedge fund replication methods to make a conclusion on whether hedge fund clones are trustworthy or not.

1.1 Problem Statement

This study will focus on hedge fund replication in the search for creating a product that is more transparent and generates a similar return to a hedge fund index. This is done to create more insight into the black box of hedge funds and discuss whether hedge fund clones can be an alternative to the expensive hedge funds. This study will therefore seek an answer to the following problem statement:

> *Is it possible to replicate hedge fund strategies by using factor-based replication methods and thereby deliver similar performance?*

To be able to answer this problem statement this study will seek answers to the following sub-statements:

> *(1) Is it possible to generate a similar month-to-month return by the use of the replication method rolling-window, based on Eurekahedge hedge fund indices?*
>
> *(2) How can the characteristics of rolling window be changed to deliver a possible better performance?*
>
> *(3) Is factor-based replication a better possibility before, during or after the financial crisis, for the different strategies within Eurekahedge hedge fund indices?*

1.2 Delimitation

To answer the above problem statement, the statistical methods and analysis limits to the factor-based replication model rolling-window. This study will also include three other replication methods in the theory section but these are limited to the overall description of the methods and the associated historical findings based on the listed literature.

The method in this study are using multiple linear regression and ordinary least square, no other regression models will be used in this study. We will test for some statistical properties time-series might have, to make sure we are aware of these when making the conclusions on the results. We will not test for multi-co-linearity because time-series more often than not are containing this, and can therefore be seen as part of nature.

The analysis limits to the 11 hedge fund indices and 25 factors described in the data section. The factor and hedge fund indices has been chosen based on historical research and to get a board cross-section of risk factors that typically can be found within hedge fund. All observations for this study are monthly data, because Eurekahedge are only proving monthly data points.
The data observation used for the analysis will be from December 1999 through August 2019. The data are obtained from Bloomberg terminal. The Bloomberg terminal were launched in 1981, and since then have they been able to give the financial market complete transparency.[2] This study analyze will only take the month-to-month return into account.

1.3 Structure of the study

This study is divided into five main parts.
(1) Theory. This section will focus on hedge fund and hedge fund replication. Within the hedge fund section, the theory that explains why hedge fund replication are attractive to understand. This is based on the benefits of hedge funds that clones aim to replicate and the downside of hedge funds that clones try to overcome. This section will also include the most common hedge fund strategies to provide some insight into hedge funds.

[2] www.bloomberg.com

The section for hedge fund replication are including a description for three different replication methods, and their associated findings. This section will also include the previous findings for the method used in the analysis of this study. The historical findings are to give a better foundation for further discussions and conclusions within the field.

(2) Data Validation. This study is using a big amount of return data, and this section will give an insight into what return data has been chosen for this study.

(3) Method. This section will include the replication methods used in this study and the statistical methods used to perform the replication. It will also include the performance measurements that will be the main focus for the analysis.

(4) Analysis. The results from the replication method will be presented, by performance measurements and graphically, if needed. This will be the main foundation for the discussion on the results and other historical findings.

(5) Discussion. This section will reflect on the presented results in the analysis and the findings under the theory section.

Lastly, the conclusion will recap all of the above to answer the problem statement of this study.

2 Theory

2.1 Hedge Fund

A hedge fund is an investment institution that, by the use of various strategies, invests in a variety of asset classes and on different markets. The investments are managed by managers that use different strategies and tools to reduce the volatility and risk meanwhile, attempting to preserve capital and deliver a positive return under any market conditions.[3]

To get an even deeper understanding of hedge funds the next four sections will explain how hedge fund operates, by going into some of the overall strategies and the benefits and downsides of hedge funds.
This are included to give a general understanding of the investment institution that this study are focusing on, but also to create an understanding of why hedge fund replication are an interesting tool to further analysis.

2.1.1 Background & Structure

The first hedge fund was established in 1949 by Alfred W. Jones. He wanted to create a new investment model that were combining mathematical analyzing tools and forecasting to create the first long/short strategy, and thereby create good investment results at a lower risk. Alfred W. Jones is also known for introducing the fee structure that hedge funds are well known for today.
The fee structure in most hedge funds today usually include a flat manager fee of 2% plus an additional incentive fee equal to 20% of the net performance. For the hedge funds to demand this fee the managers are set to deliver an abnormal performance and rate of return, which is normally described by alpha. Alpha is defined as the strategy plus the manager's ability to beat the market. It is this parameter that hedge fund replication wishes to eliminate.

The strategies and how hedge funds operate have, however, changed since 1946. These changes are made in order to maximize the return. This is done by adopting more risky strategies. Some of those strategies will be described in a later sections.

[3]Baker, H. & Filbeck, G. (2017)

Some other main characteristics that describe hedge funds are:

- Only qualified investors can invest.
- Hedge funds are not limited to what they can invest in.
- They are using leverage to create a better performance.
- They provide limit information.

The limit access to information can be a downside and risk when investing in a hedge fund. However, this is just one of the downsides. The many downsides of hedge funds are also one of the things that hedge fund replication aims to eliminate. The next section will therefore look into some of the downsides of hedge funds and how hedge fund replication might be able to get rid of some of them. The next section will also include some of the benefits that might be more difficult to create by using hedge fund replication.

2.1.2 Downside & Benefits

Not much information about hedge funds is publicly available and one of the downsides is therefore the lack of **transparency**. However, over the past decade hedge funds have improved the transparency as institutional and retail investors are starting to look for new investments and higher performance. Despite this, it is rarely that the investors know exactly what is going on inside the hedge funds. Investors are, because of this, forced to invest a lot time into doing background checks of the hedge fund to ensure good investment.
The managers are also the ones choosing to keep or change the strategy, which can then cause changes in the profile. The investor will not have to be notified, and it can take them a long time to notice that something has changed, due to the lack of transparency. The investors will however properly be able to see some changes in the monthly reported return.
This can be a huge risk for investors, as they can end up with high correlated investments within their own portfolio.
Hedge fund replication is creating more transparency by making the investment information available for the investors and they are thereby able to create a more risk neural portfolio, based on more knowledge than is offered by the hedge funds.

Increasing the transparency is important, especially after what most investors went through doing the 2008 financial crisis. But it is not seen as the biggest downside of hedge funds or either the greatest advantage of replication products.
Despite hedge fund replication creating more transparency, some investors still mean that this is not enough. When buying or establishing a hedge fund replication strategy, this strategy is a product that is based on a limited knowledge, and because of this, some investors mean that the investments are too risky. But as replication models have higher transparency compared to hedge funds, this is still seen as one of the benefits of hedge fund replication.

A great advantage of hedge fund replication products is higher **liquidity**. Most hedge funds have a lock-up period, where the investors are tied to the fund for a certain time period.
Some funds also demand a leaving fee of up to 5% to compensate remaining investors for the costs of having to re-balance the fund portfolio.[4] The higher liquidity in replication products can therefore both be cheaper for the investors but also make it easier to adjust the exposure quickly.

Another main downside of hedge funds are the massive **fees**, which were described in a previous section - especially the fees that are based on the performance can end up being a sizeable payment back to the hedge fund for the investors. This is avoided by the use of hedge fund replication that will provide the investor with a cheaper alternative.

Even if there are some major disadvantages from hedge funds that hedge fund replication can remedy, it has been found that there is still a great deal of skepticism about hedge fund replication - especially because there are also some benefits hedge fund replication cannot fulfill, such as **performance**.
Hedge funds do overall deliver a good performance which can be difficult to replicate properly because hedge funds are not meant to be replicated. Hedge funds demand huge fees because they promise to deliver an abnormal return based on the manager skills, the true alpha, which should not be a factor that is able to be replicated.

[4]Kat, H.M. & Palaro, H.P. (2005)

Most replications are based on monthly return. This means that the replication tools could be missing out on the daily work a manager puts in. Some asset managers does therefore mean that hedge fund replication might work in theory but in reality is it just an approximation of the real hedge fund return, that not all the time are good enough.[5] This study wants to investigate if these opinions are true, by analyzing if we by the use of monthly returns from various factors can overcome the daily work from a manager, the true alpha. We will in a later section go over what some historical findings and what has been done to replicate different hedge funds indices.

There exist different kinds of hedge fund replication methods. This study will mainly focus on the method described by Hasanhodzic & Lo (2007), where the hedge fund return can be shown in a linear form. But this might not always be the case; it could be that more complex models to replicate the hedge fund return are needed. But today there are still **missing advanced statistical models** and associated new technology to offer a more complex replication model.

However, based on the many downside hedge funds have, it seems as an interesting subject to investigate hedge fund replication on the linear form. Even though this does also have its flaws, is it still seen as a good strategy to overcome some of the downsides that hedge funds have. This study will, in the analysis, see if it is possible to offer a deeper transparency of the hedge fund indices before, during and after the 2008 financial crisis, by being able to replicate different hedge fund strategies using the linear form of factor-based replication. The next section will therefore describe some of the strategies that can be found in hedge fund indices. The section will focus on the strategies that are included on the analysis. This is done to shed light on the missing information and create a better ground for later selection of replication factor to include.

[5]EDHEC Risk & Asset Management Research Centre publication. (2008)

2.1.3 Strategies

To have the best prerequisites to select the factor for the replication, this section will go over some of the common strategies in hedge funds. This are a good way of getting some inside into hedge funds.
The strategies that will be included are *Equity*, *Event Driven*, *Macro*,*Relative Value* and *Multi-Strategy*, and some of the associated sub-strategies.

Equity

As the name implies, this type of fund trades listed equities, taking both long and short positions. The main task in this strategy is to identify equities that are over performing or under performing and thereby set up an equity portfolio with returns that are not correlated to market performance. The individual manager can create their own variations of this strategy by focusing on a particular market cap, sector or industry, whether the equity offers a dividend or by analyzing individual securities against a set of balance sheet fundamentals.[6] Within this strategy is also *Equity market-neutral*. This strategy is taking a long and short position in equities so the created portfolio is market-neutral. The target is to get the beta-adjusted net exposure close to zero.

$$Beta\ adjusted\ net\ exposure = \sum_{i=1}^{L+S} \omega_i * \beta_i \tag{1}$$

Where ω_i is the stock's weight in the portfolio as a percentage of the fund's net asset value and β_i is the stock's beta relative to the market.[7]
Equity Long/Short and Equity market-neutral are closely related but while long/short equity funds are exposed to systematic risk, equity market-neutral funds are trying to eliminate this risk and mainly being exposed to idiosyncratic risks. By eliminating this risk they are also eliminating the chance of over performance. Equity market-neutral funds are therefore making money in an upwards tending market if the long positions increase more rapidly than the drop in value incurred by the short positions in the portfolio and vice versa in a downwards tending marked.

[6] McCann, B. (2014)
[7] Baker, H. & Filbeck, G. (2017)

At the extreme end is the strategy variation *Short-biased*. Here the managers specialize in short selling, which means that they seek equities that are expected to decline, e.g. companies that are negatively affected by technology changes.
Hedge fund managers using this strategy are trading on the same market as the traditional equity managers. What differentiates the two types of managers are short selling, leverage and the manager's incentive system structurally.[8]

Event Driven

This strategy is looking for pricing inefficiencies that could occur because of a corporate event. A corporate event could e.g. be mergers, bankruptcies and leadership transitions. But events could for some also include earnings surprises and dividend announcements; this form of the events are called soft events where the corporate events are called hard events. Event driven can be divided into sub-strategies such as *merger arbitrage* and *distressed securities*.
Merger arbitrage takes advantage of the possibility of success and failure of a merger or acquisition. The basis technique is to take advantage of share pricing, which will be traded at a discounted prices compared to the offered price when a merger or acquisition is made public. The difference in price is called arbitrage spread, and reflects the time value of money and uncertainty about the final outcome.
Distressed Securities are primarily the purchase of debt securities of issuers that are in financial difficulties or are going through a reorganization or liquidation because of bankruptcy law. The securities are often priced at a lower price than their implicit fundamental value, because of forced selling, high degree of uncertainty faced by the market, etc.

Macro

Macro or global macro are funds that are mainly driven by large macroeconomic conditions. Funds that use this strategy want to capture dislocations in the market that are caused by macroeconomic events.
Funds that using macro strategy tend to operate on a longer term and focus on exogenous economic and political-risk events, compared to event driven that focus on internal corporate events.

[8] Baker, H. & Filbeck, G. (2017)

Macro funds can hold both long and short positions in different asset classes, such as equities, fixed income, cash & cash equivalents, etc., but many positions are generally over a long-term view. But before deciding what assets to invest in, a top-down approach is used to look at different economy and political risk overlays to resolve what position to take.
Economy and politics can change quickly and it can therefore be harmful for macro funds to take an illiquid position. Limiting the illiquid assets will result in a portfolio with a lower risk. This trading strategy can also be divided into sub-strategies - for instance, *systematic diversified* and *Discretionary Thematic*.
Systematic diversified rely on mathematical, algorithmic and technical models to identify markets trends; very little or no human interaction is used under this sub-strategy.
Discretionary thematic relies on individuals to make an investment decision by analyzing market data, relationships and influences. This makes the strategy more flexible to market changes than systematically diversified.

Relative Value

The main for this strategy is to identify and exploit pricing discrepancies among the same or related securities using long and short positions.[9] The trades are made profitable if the securities in the portfolio are closing in on the fundamental value. Relative values are built on the idea of arbitrage, but the trade related to this strategy will not be risk-free. The funds are, however, trying to create a market-neutral portfolio by combining long and short positions so beta is as close to zero as possible, but not all sub-strategies can fulfill this, e.g. *convertible arbitrage*.
Convertible arbitrage focuses on pricing discrepancies in the convertible bond market. The main idea is to take a long position in an undervalued convertible bond and short position in the related stock. To keep a market-neutral portfolio under this sub-strategy, mangers must increase their hedge, or sell more stocks short if the price goes up and vice versa if the prices goes down. Convertible arbitrage can also create profit by using volatility trading, by rebalancing the hedge ratio when stock prices are changing in value.
Sub-strategies that are considered a market-neutral strategy are *fixed-income arbitrage*. Fixed-income arbitrage focuses on pricing discrepancies in fixed-income instruments, such as govern-

[9] Baker, H. & Filbeck, G. (2017)

ment bonds, corporate bonds and mortgage-backed securities and goes both long and short in the exportation that the overpriced and underpriced security will get closer together.

Multi-Strategy

To have a wider approach to the market it can be a good idea to place the capital where the best investments are; this type of strategy are called multi-strategy. This approach to the market is profitable because sometimes the opportunity within one strategy can be too narrow. Multi-strategy can be divided into two types *multi-strategy single funds* and *multi-strategy funds of Funds*. In *multi-strategy single funds* are a fund combining at least two strategies to offer a more diversified product, they combine the managers knowledge to create a wider portfolio.
Multi-strategy funds of funds have the same basic idea, combining different strategies, but instead of doing it internally, this strategy invests in other funds.

This section has described some of the most common hedge fund strategies, but a hedge fund strategy can come in many variations; each manager has their own variation of some classic trading strategy. But the common strategies have been described to get a better understanding of what can be found within the different hedge fund index selected to be replicated. The selection of factors to the replication are based on the knowledge of hedge fund strategies plus historical findings and will be described in section three, data validation.

2.1.4 General Performance

To get more understanding of why hedge fund replication has become an interesting topic to study and why multiple methods have been developed and implemented in some of the world's largest banks, this section will look at the general performance for hedge funds, which is one of the benefits that can be difficult to replicate.

As mentioned before, hedge funds are extremely limited with information about their performance. This study will therefore describe the overall performance of hedge funds by mainly using the study made by Metzger, N. & Shenai, V. (2019). Their study is seen as a good indication of hedge fund performance because they use hedge fund indices from a reasonable

large database. The data are from Hedge Index and include over 9500 funds, which is considered a reasonable amount to provide an indication of the general performance.

To be able to conclude something about the hedge fund's performance a benchmark is needed. The article Metzger, N. & Shenai, V. (2019) are using the S&P500 to show how the hedge funds are performing compared to the market (S&P 500).

The period analyzed will be from June 2007 to January 2017.[10] It is almost the same period that this study will be using during the analysis. There will only be a difference of two-and-a-half years as our analysis will include numbers through August 2019. The difference in period length does not matter, because this section wants to show why it has become an interesting thing to replicate hedge funds.

The performances for 10 different hedge fund strategies are visually shown in figure 1 below

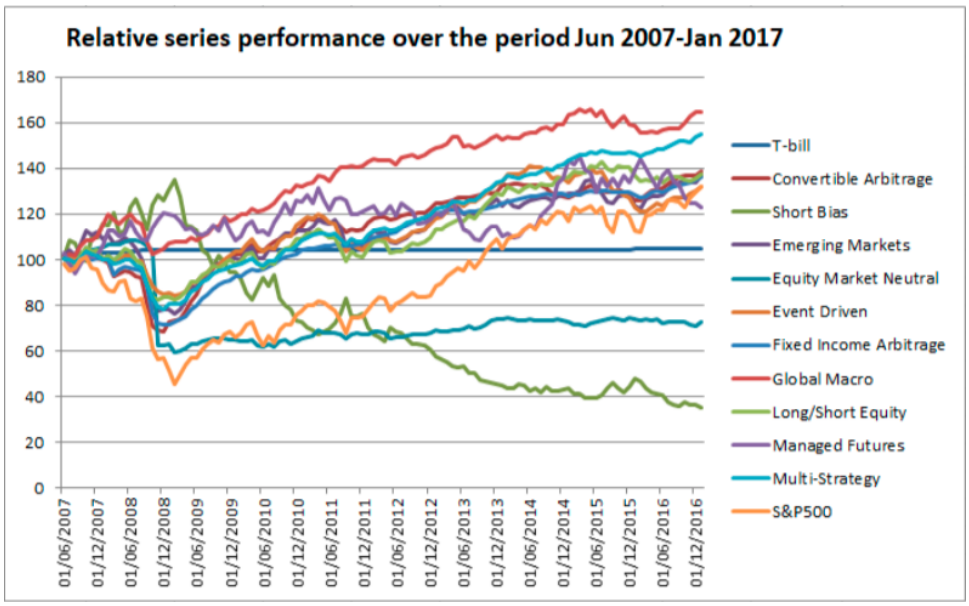


Figure 1: Performance June 2007 - January 2017, Source: Metzger, N. & Shenai, V. (2019)

[10] In Metzger, N. & Shenai, V. (2019)

Over the whole period are three strategies performing under the benchmark and two of them under index 100 and the risk-free rate. The remaining seven strategies are performing over both the risk-free rate and the benchmark, with Global Macro strategy with the highest index at 164.47.
During the financial crises all strategies over performed the market that dropped to index 49.42. If we only focus at the period just after the crises the market is creating the highest returns, but this is because of the extreme low price the market had just after the crises. Hedge funds are therefore a more stable investment by creating less volatility over time to lower the risk and generate maximum return - which is what they are said to do, and also why they are charging such high fees.
This result shows why hedge fund replication has become such an interesting topic to investigate, to be able to deliver a good performance and a lower volatility than the market at a lower cost.

This overall performance also implies that the performance depends on the manager's ability to forecast and make the right choices in the right time to avoid the high volatility. These properties can be hard to replicate, especially by the use of passive investment strategy. This is also one of the reasons why we have chosen the method rolling window, which will be described later on, and also why we have deselected some other replication methods. This is also the main reason that we, in this study, will further develop the replications method rolling window - to maybe deliver a possible better replicate. If we by the use of these methods will be able to do so will be shown in the analysis.

According to this research, the best performing strategies are Global Macro and Multi Strategy and the worse performing strategy is Short Bias. Short Bias delivers an index at the end of the period at 35.42 and is very negatively correlated to the market performance.
It is also a very interesting subject to look into what strategy that is most likely to replicate successful. This will be analyzed and discussed in this study later on.

2.2 Hedge Fund Replication

To get a deeper understanding of the main topic for this study, this section will describe some of the methods and findings within hedge fund replication. This section will include overall descriptions of some different replication methods, including methods that we will not use in the analysis. This is done to give an understanding of different approaches that could have been chosen, that might have given a different end result, and provide knowledge for further discussion and development within this field.
This section will also include the historical findings within hedge fund replication; this is to see what has previously been discovered and to help create a discussion and conclude on our results later on.

2.2.1 Replication Methods

The main purpose of this study is to examine whether it is possible to replicate different hedge fund strategies. This can be done by using statistical models or algorithm trading strategies to replicate a hedge fund index returns, to create attractive hedge fund clones.
An attractive hedge fund clone, is either a clone that provides the same month-to-month return or the same statistical properties as the return of the hedge fund index. The differences relies within what approaches or methods that are being used.

The hedge fund clones may be found by the following two methods:

- Factor-based replication
- Payoff distribution replication

The two methods mentioned in this study are performed in relative different ways, and from this comes the different outcome. Factor-based replication gives the month-to-month return by finding a relationship between the hedge fund return and different risk factors and their associated sensitives, where Payoff distribution replication uses the performance to create a clone that has the same properties as the hedge fund.
The Factor-based replication is a bit easier to perform than payoff distribution replication. Because of the simplicity compared to payoff distribution replication, this approach is chosen

for this study. This decision is made because factor-based creates more transparency, which gives us more to base our conclusion on when looking at the ability to replicate hedge fund performance.[11] But this choice has also been made as we are missing the advanced statistical models and associated technology.

As factor-based replication creates more transparency, is easiest to perform and is cost-effective, the approach has also become the most used when replicating hedge fund returns. But within this approach are developed different methods. This study will, in the analysis, use the method **Rolling Window**, which will then be described in the method section and used in the analysis. This well-know method will therefore not be described in this section, as this section is to provide insight in what other replication methods exist to create a better grounds for further discussion on our results.

Within the factor-based approached the method **Fixed-Weight** and **Kalman filter** also exists.

The factor-based methods are built around the same theory as we are using in our analysis, and can therefore be found described in the section method. Overall the factor-based approach using the hedge fund return can be shown in linear or non-linear form, for the linear form, see equation 5. We will, as mentioned, stick to including and describing Fixed-Weight and Kalman filter, which are based on the linear form, as rolling window is also shown on the linear form.

The method fixed-weight is widely known from the same study as we are using as background work for rolling window. This study is done by Hasanhodzic, J. & W. Lo, A. (2007), and are named *"Can hedge-fund return be replicated? The linear case"*. This method are a passive investment strategy because an in-sample period are used to construct a regression that contains fixed factors and their associated sensitives, β, there then are used on the whole out-sample period, meaning that for the clone returns are only the factor changing relative to time, t.

$$R_t^{Clone} = \sum_{i=1}^{I} \beta_i * F_{it} \tag{2}$$

[11] Jawadi, F & Khanniche, S. (2011)

The linear form of the return can be found by the use of linear regression and Ordinary Least Square; see more details in the method section.

The second method within factor-based replication that we will include is **Kalman filter**. This is a computational algorithm that better captures the time changing dynamics of hedge fund returns.[12]. This method is included, as the foundation for developing the extended version of the rolling window that will be described in the section method. The Kalman filter method has therefore been chosen to include in this study to give a deeper understanding of why we implement the extended version of the rolling window, but also to be able to compare if a method that better captures the time changing dynamics of hedge fund returns performs on a different level than a method with fixed factors.

The standard Kalman filter is split into two steps: the prediction and the updating step. These are given by the following equations:

Prediction step:

$$R_t = \sum_{i=1}^{k} \beta_{it} Fit + \epsilon_t \tag{3}$$

Updating step:

$$\beta_{it+1} = \beta_{it} + b_{it+1} \tag{4}$$

where ϵ_t and b_{it+1} is an error term that is independently normally distributed with mean zero and a constant variance, just as the assumptions under ordinary least square, which is described and used in the method and analysis.
b_{it+1} change over time according to random walk but is also depending on the previous value.

Kooli, M. & Sharma, S. (2012) did a study that included a new version of the Kalman filter. They implemented that the factors could change over time, meaning that the factors

[12]Kooli, M. & Sharma, S. (2012)

are added to the replicated portfolio if they become significant for the model and removed if they become insignificant.
This addition does therefore take into account that hedge funds do not have constant exposure over time by changing both the risk factors and their associated sensitive.

This development gave the idea to extend the standard rolling window with the same property, to see if we will be able to make a better fit on poor performing clones and if this extension will have any financial and statistical significance on the results.

The difference between Kooli, M. & Sharma, S. (2012) method and the one we will be using in the analysis of this study is based on the statistical method behind the two approaches. We have chosen to use the same statistical method as in rolling window to make our analysis more comparable, but have also found it important to include Kalman filter in our study to use the historical finding from this method to compare with our results from the extended version of rolling window, and for further discussion within the field hedge fund replication.

A completely different approach within hedge fund replication is **Payoff Distribution Replication**. Payoff distribution replication does not want to clone the month-to-month return, because some investors will maybe not be interested in looking at every months'. Some investors are more interested in finding a return with the same statistical properties as the return generated by the hedge fund. Another difference in this method compared to the factor-based approach is that this approach also takes the investors' existing investment into account.

Kat, H.M. & Palaro, H.P. (2005) showed that it was possible to replicate the statistical properties of the joint return between the investor's portfolio and a hedge fund by creating a joint return between the investor's portfolio and some risk factors.
This is done by the use of joint distribution and an associated payoff function that then aims to give the same probability distribution for the clone as the desired distribution(Hedge fund and investors portfolio). This will create a dynamic trading strategy.
There can be multiple payoff functions that fit the desired distribution; it is therefore also necessary to find the cheapest payoff function.

This can be done by a simple set of assumptions:

- Investors' preferences depend only on the joint probability distribution of terminal wealth derived from the investment and their existing portfolio.[13]
- Perfect capital markets.
- Investors prefer more to less.

By these assumptions would it be possible to find the cheapest payoff function for the joint distribution of the risk factors and the investor's portfolio that will have the same properties as the desired distribution.

This method involves a lot of different statistical methods, which are not the same as in the factor-based replication. The statistical methods are very complex and have a time frame of understanding and developing that exceeds the time frame of the master thesis. Therefore are this method only included in this study in the theory where it is overall described and the methods historical findings will be shown in the next section.
It has been decided to include this method in this study to explain a different approach to hedge fund replication. This can be used for further discussion on different approaches that could have been taken with hedge fund replication and on our results when comparing to historical findings from payoff distribution replication.

The next section will therefore show some of the historical findings that other researchers have found, within the two different approaches for hedge fund replication, factor-based and payoff distribution replication.
Within factor-based replication there will be both findings from fixed-weight and Kalman filter. This is for further discussion on the ability of other methods to replicate month-to-month return relative to our analysis. But also findings within rolling window will be shown in the next section, as this method will be used in our analysis and is therefore very important to include to compare our ability to replication hedge fund index by the use of rolling window relative to other research that has used the same method.

[13]Harry M. Kat & Helder P. Palaro. 2005

2.2.2 Historical Findings

As hedge fund replication has been studied by multiple researchers over the last couple of decades, they each have different approaches to the method described above; this study will focus on the findings from: (1) Hasanhodzic, J. & W. Lo, A. (2007), where they used the factor-based approached on 11 strategies. They used six predetermined factors and the methods fixed-weight and 24-month rolling window to replicate the 11 strategies. (2) Kooli, M. & Sharma, S. (2012), used the factor-based approached on 14 strategies and 22 factors but for each strategy the nonsignificant factors was deselected. Their study used fixed-weight and Kalman filter to replicate the 14 strategies. (3) Kat, H.M. & Palaro, H.P. (2005) & (2006), used the approach payoff distribution, where they wanted to replicate the statistical properties of different hedge fund returns.

For the convenience of this study the historical findings will be divided into sections of findings for factor-based replication and payoff distribution replication.

Factor-based replication

To try to make a conclusion on whether the replication of the strategy is good or bad, it is important to look at different performance measurements, just as our later analysis will. But to give an indication of the performance, the table on the next page will show the Sharpe ratio for the hedge fund and the clones for the different studies. This measurement is comparing the return of the investment to the risk taken. The higher the Sharpe ratio the better the investment has been performing compared to the risk taken. Further description of the Sharpe ratio is found in section 4.7.1. This is therefore a good indication of the performance, but a deeper explanation of the historical finding will follow after the table.

Factor-Based Replication							
Strategy	Hasanhodzic, J. & W. Lo, A. (2007)				Kooli, M. & Sharma, S. (2012)		
	Hedge Fund Index	Clone Fixed-Weight	Hedge Fund Index	Clone Rolling-Window	Hedge Fund Index	Clone Fixed-Weight	Clone Kalman Filter
All strategies	1.63	2.00	1.61	1.20	0.26	0.19	0.23
Convertible Arbitrage	2.07	1.81	1.85	1.17	0.10	0.30	0.26
Short Bias	0.28	0.09	0.12	0.19	0.06	-0.02	-0.01
Emerging Market	1.26	0.71	0.96	0.36	0.20	0.18	0.18
Equity Market Neutral	2.06	2.41	1.04	1.42	0.28	0.50	0.62
Event Driven	3.08	1.89	3.06	1.71	0.30	0.33	0.26
Fixed Income Arbitrage	2.93	1.79	2.03	1.12	0.19	0.73	0.41
Global Macro	1.73	2.43	1.39	1.48	0.40	0.23	0.29
Long/Short Equity	1.38	1.22	1.67	0.99	0.26	0.16	0.17
Managed Futures	0.83	1.80	0.82	0.78	-	-	-
Multi-Strategy	2.52	1.62	2.00	1.04	-	-	-
Fund of Funds	1.59	2.21	1.61	1.68	0.15	0.01	0.17
Relative Value	-	-	-	-	0.28	0.11	0.37
Merger Arbitrage	-	-	-	-	0.51	0.71	0.60
Distressed Securities	-	-	-	-	0.33	0.42	0.40

Hasanhodzic, J. & W. Lo, A. (2007) numbers are from February 1986 to September 2005, where Kooli, M. & Sharma, S. (2012) are from February 2004 to September 2009; this causes the difference in Sharpe ratio because of different risk-free rates, but also that the Hedge Fund Index are from two different databases, so they most likely has different portfolios. It is also important to be aware of difference in the two studies on how the factors are selected for the linear equation. If they are determined beforehand, like Hasanhodzic, J. & W. Lo, A. (2007), or are they selected by the use of linear regression and ordinary least squares, like Kooli, M. & Sharma, S. (2012).

The research from Hasanhodzic, J. & W. Lo, A. (2007) does show a over performance on 5 out of 12 strategies on fixed-weight, based on Sharpe ratio and mean return. The over performing strategies are:
(1) All strategies with an annualized mean return of +5.68pp vs. the hedge fund index.
(2) Equity market neutral with an annualized mean return of +5.83pp vs. the hedge fund index.
(3) Global macro with an annualized mean return of +9.96pp vs. the hedge fund index.
(4) Managed futures with an annualized mean return of +18.77pp vs. the hedge fund index.
(5) Fund of funds with an annualized mean return of +5,68pp vs. the hedge fund index.

Long/short Equity does have a lower Sharpe ratio than the hedge fund index but does deliver a slightly better annualized mean return.
The rest of the hedge fund strategies do deliver a lower mean return and Sharpe ratio - especially Short Bias, which delivers an annual compound mean return that is negative, and Emerging markets, which has an annualized mean return of -8.65pp vs. the hedge fund index. Looking at rolling-window it also over performs in 5 out of 12 strategies and these 5 strategies are almost the same as fixed-weight, except all strategies are over performing in fixed-weight and Short Bias are over performing in rolling-window. This indicates which strategies they have had the most success replicating. The worst performing clone is still Emerging markets, which could indicate general difficulties in replication of this strategy.

In general, Kooli, M. & Sharma, S. (2012) have less well performing fixed-weight replication, but do a better job replicating Emerging markets than Hasanhodzic, J. & W. Lo, A. (2007). One struggling strategy for them is Short Bias, which has a negative mean return on both replication methods. Both methods do over perform the hedge fund index in both Sharpe ratio and mean return in the strategies Fixed income, arbitrage and Convertible arbitrage, which were not some of the strategies that Hasanhodzic, J. & W. Lo, A. (2007) did over perform. This might indicate that the way the factors are chosen could have an effect on the results. In both strategies, fixed-weight delivers a better mean return than both the hedge fund index and Kalman filter, but Kalman filter has higher correlation with the index than fixed-weight, meaning that both strategies have their pros and cons. This is one of the reasons why we have decided to include and further develop poor preforming clones to see if we can get a better result or a better fit in different areas.

Comparing the two methods from Kooli, M. & Sharma, S. (2012) Kalman filter are on multi performance measurements more alike the various hedge fund index. This means that this method is more likely to capture dynamic time changes. This observation is also one of the reasons why the study include the extended version of rolling window, to see if we on poor performing clones better can capture changes over time, and our method is therefore built around the Kalman filter method from Kooli, M. & Sharma, S. (2012).

The historical findings for payoff distribution replication will be based on Kat, H.M. & Palaro, H.P. (2006), which are an extension of their work from 2005, which is described in the previous section.
Their research found that by evaluating the net-of-fee performance of 485 funds of hedge funds, the majority of them did not provide the investor with a return they could not have created by themselves, by investing in S&P 500, T-bond and Eurodollar futures.[14] They did the same finding one month later, where they evaluated the net-of-fee performance on 1,917 individual hedge funds, and found that only 17.7% of the funds beat the benchmark.[15] This were by investing in the same assets as the work before.

[14]Kat, H.M. & Palaro, H.P. (2006)
[15]Kat, H.M. & Palaro, H.P. (2006)

This is, when compared to previous hedge fund replication studies, an interesting finding. Most replication studies, including the once described above, only deliver a more equal result within some strategies, and are not near the performance on others. The percentage of good performing results are not as good as Kat, H.M. & Palaro, H.P. (2006) and the consistency within the findings are not strong enough to make the same conclusion as Kat, H.M. & Palaro, H.P. (2006).

Most other findings indicate that hedge fund replication can be a real thing, but do not make as strong a conclusion as Kat, H.M. & Palaro, H.P. (2006): "in terms of returns therefore, most funds of funds have failed to add value"[16]. The same conclusion can be made on the individual hedge fund.

One difference that is also worth noticing between the described hedge fund replication methods are if the strategies do both take long and short or only long positions. The methods under the factor-based approach do allow both long and short positioning, where the method under payoff distribution replication does only use long positions. However, Takahashi, A. & Yamamoto, K. (2013) did a research on payoff distribution replication with both long and short positions and the results showed that the performance improved substantially compared to Kat, H.M. & Palaro, H.P. (2006). [17]

This could indicate that this method is performing on a high level within hedge fund replication. But as it requires a lot of resources and highly advanced skills this method is not suitable for individual investors, or this study. This method does also not provide us with as much transparency as factor-based replication and is therefore not found suitable for the problem of this study. This replication method is not chosen, but the historical results are definitely worth noticing and including for further discussion within the area of hedge fund replication.

[16]Kat, H.M. & Palaro, H.P. (2006)

[17]Takahashi, A. & Yamamoto, K. (2013)

3 Data Validation

This study contains a large amount of data collected from Bloomberg. The data collected are historical return data on Hedge fund index from Eurekahegde and different factors used for the replication. The data will containing observations from December 1999 through August 2019, this period are chosen because Eurekahegde's index starts to provide data from that given time. If some factors don't provide us with information back to December 1999, are there returns set equal to zero until they start. All data sets are also collected in US dollars. Given that we have chosen to work with monthly return data and we have selected 11 hedge fund index and 42 factors, we will have around 12,700 observations to work with; we do therefore find it necessary to describe the data that we are working with. This will also give a better understanding of what we are basing our analysis on and our results on the different hedge fund strategies.

3.1 Hedge Fund Index

The hedge fund indices that this study aims to replicate is from Eurekahedge. In the process of selecting a database that provides a hedge fund index based on strategies, we also looked into data from Credit Suisse, but because of the number of constituent funds within each index, we chose the database that had the largest numbers. This is because we wish to get a broad performance to replicate in our analysis. The main index for Eurekahedge has 2447 constituent funds, compared to Credit Suisse, which limits its index to 25 funds.

To be able to investigate if some hedge fund strategies are more suitable for replication, we have collected an overall index and different strategy index based on what is accessible and the strategies that we described in the theory section. The index strategies included in this study will therefore be:

- Long/Short Equity (949 constituent funds)
- Macro (195 constituent funds)
- Multi Strategy (255 constituent funds)
- Relative Value (81 constituent funds)

- Arbitrage (82 constituent funds)
- Distressed Debt (21 constituent funds)
- Event Driven (110 constituent funds)
- Fixed Income (315 constituent funds)
- Equity Market Neutral (58 constituent funds)
- Equity Short Bias (1 constituent fund)

All indices are equally weighted of all the constituting funds, meaning that we get an overview of the average performance of all the constituent funds. This has been chosen over asset-weighted because we do not want to highlight monthly inflows and unjustly overweight performance due to good marketing or location in investor hot spots.[18]

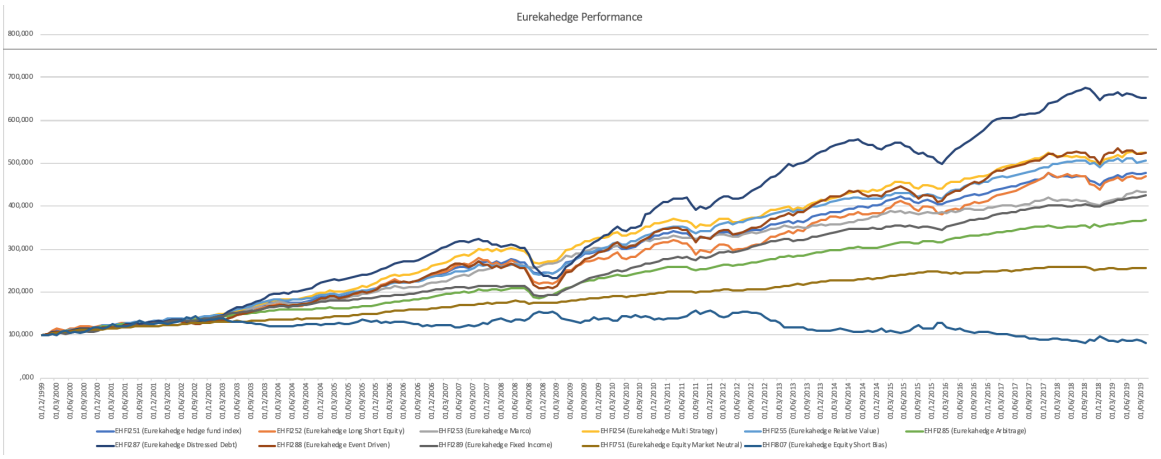


Figure 2: Eurekahedge Index Performance December 1999 - August 2019

In figure 2 the visualization of the Eurekahedge fund indices are shown, from the establishment in December 1999 until today. Looking at the development in the different indices we divide the indices into 4 groups.

[18] http://www.eurekahedge.com/

The first group only contains one index, which is the distressed debt strategy index. This index has its own development through the periods, and with the highest index-value today, it is the index with the best performance since the establishment.

The second group contains several indices, which all have a very similar pattern, and with a consistent return over time. The group contain all indices except the distressed debt, Equity market neutral and Equity short bias. In this group the arbitrage and fixed income hedge fund indices are less impressionable to volatility in the market, such as the financial crisis. This is seen in figure 2, where the financial crisis creates large fluctuations in the indices, but less in these to hedge fund indices.

The third group is the equity market neutral strategy. This index has a very steady increase through all the difference in-sample periods. Furthermore, the hedge fund is very unique compared to the other indices, which are all more or less impressionable to the market.

The last group is only one index, the Equity short bias. This is the only index which is under index 100 today, and then with a negative development since the establishment. The index is positively influenced with e.g. the financial crisis, which means that under the financial crisis the index increased where the rest of the indices decreased.

In figure 1, the performance for hedge strategies from June 2007 to January 2017 are shown with S&P500 and the risk-free rate (T-bill). To compare the result of these strategies from Metzger, N. & Shenai, V. (2019), figure 3 shows the same period with index 100 June 2007 for Eurekahedge hedge fund indices in different strategies.

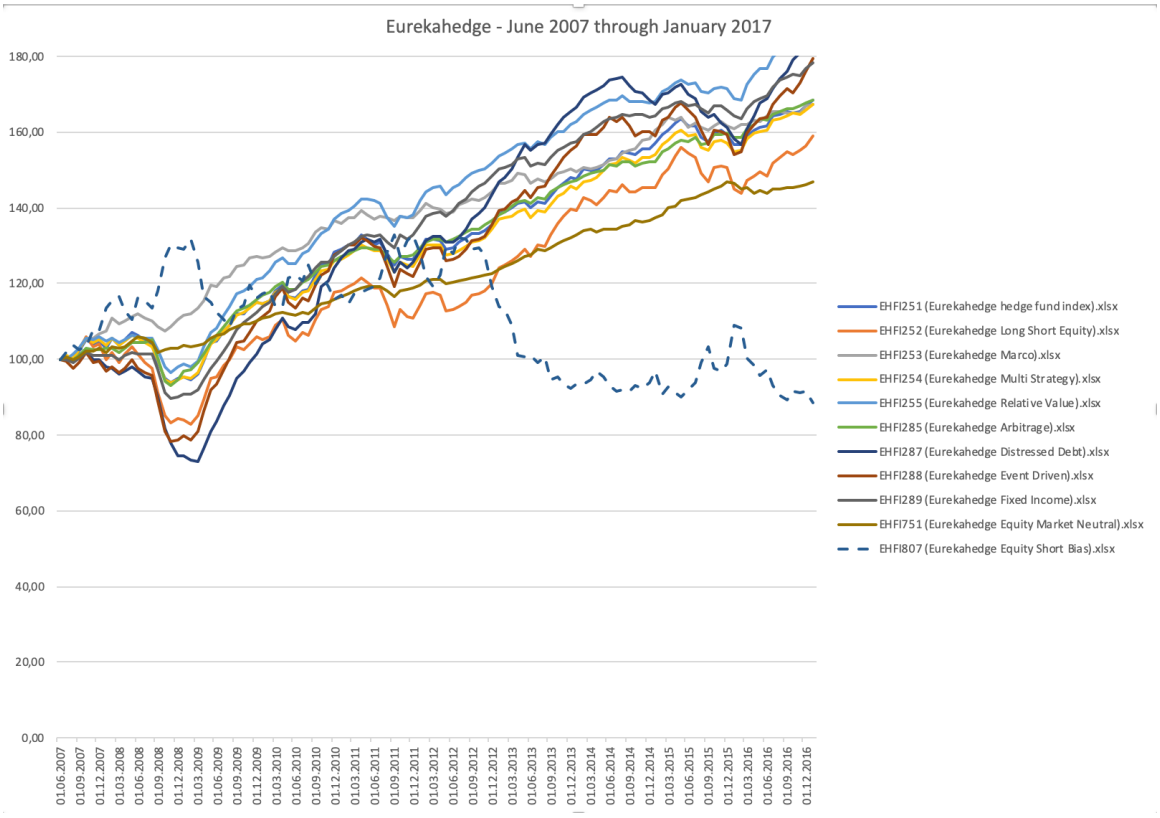


Figure 3: Eurekahedge performance June 2007 - January 2017

The most obvious strategy to compare is the short equity with the long/short equity. In figure 1 we see a huge increase in the index in the first 1.5 year, and then the index decreases drastically. This picture is equal in both examples, and shows a similarity in the strategies across the different sources.

There are also some strategies where the two sources differ, e.g. the equity market neutral strategy are ending around index 75 in the Metzger, N. & Shenai, V. (2019) source, and the Eurekahedge Equity market neutral strategy index and ending in index 147.
This shows that even though the strategies are similar, the result can differ from source to source. But given the differences the chosen data are still delivering a good performance over the whole period and therefore fit the findings described under the general performance of hedge funds.

3.2 Factors

This factors are chosen to have covered a reasonably broad cross-section of risk exposures of the typical hedge fund. All the different factors are chosen to cover all different categories that we are trying to replicate. The different strategies are covered in section 2.1.3, but here the factors will be divided into different asset classes.

> *"An Asset class is a grouping of investments that exhibit similar characteristics and are subject to the same laws and regulations"* [19]

3.2.1 Equities

The asset class with most factors is equities. This asset class is representing a share of ownership in a company, which is possible for purchasing.[20]
The asset class Equities have produced the highest returns, when investing 7 years and longer, and thereby giving the highest returns for the long term investors. Achieving high returns is not the only income an investor gets from investing in equities. In general, companies pay dividends every six month, which is a portion of the annual profit.[21]

The set of factors inside equities are 25. Of course, some will be left out of the analysis because of too high correlation with similar factors, but they are all chosen to target different strategies that this study want to replicate.
The factors are:

1. **DFA Emerging Markets Core Equity Portfolio** (DFCEX USD Equity) are the first factor in the class of equities. For DFCEX the index contains 99.68 non-US equities.[22] The risk and returns are average for the asset class.

2. **Dow Jones US Growth Index** (DJUSGR). The US Growth Index from Dow Jones and the Value index are both indices from the company, which was founded by the founder of Wall Street Journal, Charles H. Dow.

[19] www.investopedia.com
[20] www.prudential.co.za
[21] www.prudential.co.za
[22] www.morningstar.com

3. **Dow Jones US Value Index** (DJUSVA).

4. **iShares Emerging Markets index fund** (EEM). According to iShares, the goal is to achieve a return which matches the global emerging equity market.[23] iShares Emerging Market index fund has a benchmark index, this is MSCI Emerging Markets Index or MXEF index.

5. **Morgan Stanley Capital International Europe, Australasia and Far East Index** (MXEA) are capturing large and mid-cap representations, and have 920 constituents funds.[24]

6. **Morgan Stanley Capital International Emerging Market Index** (MXEF). Morgan Stanley are explain the goal for this index to be:

   > *"At MSCI, we strive to bring transparency to these dynamic yet disparate economies"*[25]

7. **Morgan Stanley Capital International Emerging Market Growth Index** (MXEF000G)

8. **Morgan Stanley Capital International Europe Growth Index** (MXEU000G)

9. **Morgan Stanley Capital International Europe Value Index** (MXEU000V)

10. **Morgan Stanley Capital International US Growth Index** (MXUS000G). The four indices MSCI EM growth, EU growth, EU value and US growth are all equity indices from Morgan Stanley international. The growth investments are defined using 4 variables: long-term forward EPS growth rate, short-term forward EPS growth rate, current internal growth rate and long-term historical EPS growth trend and long-term historical sales per share growth trend.[26]

11. **Morgan Stanley Capital International World Large Cap** (MXWOLC). With 736 constituents, MXWOLC captures large cap representation in 23 developed countries.[27]

[23] www.ishares.com
[24] www.MSCI.com
[25] www.MSCI.com
[26] www.MSCI.com
[27] www.msci.com

12. **Morgan Stanley Capital International World Small Cap** (MXWOSC). With 4.352 constituents, MXWOSC captures small cap representation in 23 developed countries.[28]

13. **Russell 1000 Growth index** (RLG) are the approximately 1000 largest US companies on the equity market.[29]

14. **S&P Global Broad Market index** (SBBMGLU). This factor includes over 11.000 stocks, which come from 25 developed markets and 25 emerging markets.

15. **S&P Emerging Broad Market index** (SCRTEM) is an index containing the companies within the emerging market from the Global Broad Market index.

16. **Spread between Dow Jones US Growth Index and Dow Jones US Value Index**

17. **Spread between iShares Emerging Markets index fund and iShares Morgningstar Large-Cap**

18. **Spread between Morgan Stanley Capital International US Growth Index and Morgan Stanley Capital International US Value Index**

19. **Spread between Morgan Stanley Capital International Europe, Australasia and Far East Index and Morgan Stanley Capital International US Large Cap**

20. **Spread between Morgan Stanley Capital International Europe Growth Index and Morgan Stanley Capital International Europe Value Index**

21. **Spread between Morgan Stanley Capital International World Large Cap and Morgan Stanley Capital International World Small Cap**

22. **Spread between Vanguard Developed Markets Index Fund and Vanguard Large Cap index Admiral Shares**

23. **Spread between S&P Emerging Broad Market index and S&P 500 index**

24. **S&P 500 Value Index** (SVX). Value stocks are measured with using three factors: The ratios of book value, earnings, and sales to price.[30]

[28] www.msci.com

[29] www.investopedia.com

[30] US spindices

25. **Vanguard Developed Markets Index Fund** (VTMGX)

> *"This index fund provides investors low-cost, diversified exposure to large-, mid-, and small-capitalization companies in developed markets outside of the United States."*[31]

3.2.2 Fixed income

The most common fixed incomes are government and corporate bonds, but other funds are also tradeable as well. The fixed income asset group is a sort of investment security which pays a fixed income (interest payments). At maturity the principal will be repaid to the investor.[32]

This type of investment is often used by conservative investors, and companies are using these investors to lend them money by issuing a bond with a given rate. These funds are, among others, used to finance larger projects or unequal economies.

In this study we have incorporated 8 factors within the fixed income asset class.

1. **Credit Suisse High Yield Bond Fund** (DHY) is a non-diversified closed-end management investment company, which invests in securities with high current income.[33]
2. **iShares J.P. Morgan USD Emerging Markets Bond** (EMB) investment strategy is US Dollar bonds from emerging markets, with a minimum of 2 year to maturity.
3. **iShares 7-10 Year Treasury Bond** (IEF) are focusing on maturities between 7 and 10 years,
4. **Barclays US Corporate Bond Index** (LUACTRUU) are from bloomberg described as:

   > *"...measures the investment grade, fixed-rate, taxable corporate bond market."*[34]

[31] https://investor.vanguard.com/
[32] www.investopedia.com
[33] www.bloomberg.com
[34] www.bloomberg.com

5. **Barclays US Mortgage Backed Securities Index** (LUMSTRUU) are described by Bloomberg as:

   "...is constructed by grouping individual TBA-deliverable MBS pools into aggregates or generics based on program, coupon and vintage."[35]

6. **S&P U.S. Treasury Bond 10 Year** (SPBDU10T) are constructed to measure the US treasury bond markets performance with 10 year maturity.

7. **S&P U.S. Mortgage-Backed Securities** (SPMBS) is covering US Dollar mortgage backed securities.

8. **iShares iBoxx High Yield Corporate Bond** (HYG) is an exchange-traded fund.

3.2.3 Cash & cash equivalents

Cash and cash equivalents are the assets of the company, i.e. cash or assets which can be converted to cash right away.[36] It could, for instance, be a T-bill with a maturity date of less than 90 days. Therefore, the risk-free rate for this study (Generic US T-Bill 1 Month), is classified as an asset group of cash and cash equivalents, because it can rather quickly be converted into cash.

1. **US Dollard Index** (DXY) is reflecting the value of the US dollar compared to the currency of US partner countries.

2. **USDEUR** is a currency index, which shows the US Dollar relative to the Euro.

3. **USDGBP** is a currency index, which shows the US Dollar relative to the United Kingdom British Pound.

4. **USDJPY** is a currency index, which shows the US Dollar relative to the Japanese Yen.

5. **Generic US T-Bill 1 Month** (GB1M). The treasury bill are a US government debt obligation, and this factor are used as our risk-free rate for this study.

[35] www.bloomberg.com

[36] www.investopedia.com

3.2.4 Commodities

Commodities are basic goods, such as gold, oil or similar good. They can also be goods such as coffee beans and sugar. These goods are all natural products with different qualities, but commodities are more similar no matter the producer. The commodities are traded in future contracts.

1. **Bloomberg Commodity Index** (BCOM) are reflecting commodity future price movements, and is calculated on excess return.[37]

2. **Goldman Sachs Commodities Index** (GI1). Are almost equal to SPGSCI and are therefore excluded for the analysis.

3. **S&P GSCI Commodities Index** (SPGSCI). The GSCI are measuring the performance of the commodity market, and are the corresponding to S&P 500 for the equity asset class.

[37] www.bloomberg.com

4 Method

4.1 Factor-Based Replication

This study focuses on the factor-based replication method, and will in the analysis aim to replicate different hedge fund indices by using rolling window replication and for some selected hedge fund indices an extension of the rolling window method. But before going into the two different variations of the method, this section will describe the general methodology about factor-based replication.

Factor-based replication models want to replicate the month-to-month return of a certain hedge fund index; this is done by reproducing a clone based on risk factors and their associated sensitives. By identifying the risk factors and the sensitivity for these factors, it should be possible to create a portfolio of traditional assets that will generate a similar return as the hedge fund index.
The traditional assets can be stocks and bonds, or exotic assets.[38] The traditional assets that have been selected for this study are described in the data validation section.

The approaches used in this study, is the linear case where the hedge fund return can be shown on the linear form. The hedge fund return can be decomposed into the following linear regression:

$$R_t = \alpha + \beta_1 F_{1t} + \cdots + \beta_i F_{it} + \epsilon_t \tag{5}$$

$$R_t = Hedge\ Fund\ Index\ Return\ at\ time\ t$$

$$\alpha = Manager - specific\ alpha$$

$$\beta_i = Risk\ exposure\ for\ risk\ factor_i$$

$$F_{it} = Return\ of\ risk\ factor\ i,\ at\ time\ t$$

$$\epsilon = Estimated\ specific\ risk\ in\ the\ return\ of\ the\ hedge\ fund$$

[38] Jawadi, F & Khanniche, S. (2011)

As α depend on the manager skills, this parameter is taken out of the linear equation when trying to replicate the hedge fund index, because we only want the clone to be based on risk factors and their sensitivities.

The clone can be found by using multiple linear regression and ordinary least squares to estimate the parameters within the clone in the in-sample period. In the in-sample period will the return be of the form:

$$R_t = \sum_{i=1}^{I} \beta_i * F_{it} + \epsilon_i \tag{6}$$

The clone will then take the following form in the out-sample period:

$$R_t^{Clone} = \sum_{i=1}^{I} \beta_i * F_{it} \tag{7}$$

The beta can be positive and negative, meaning that we can either take long or short positions. This is done to achieve the kind of risk exposure that hedge funds normally have. They are also set to sum to one. This is to make sure that the same amount is invested in the clone as there would be invested in the hedge fund. The difference in the sum of the beta from one will be adjusted by taking either long or short position in the risk-free rate. This means that the difference between 1 and the sum of the betas are equal to either a long or short position in the risk-free rate.

As described above, the intercept is excluded because this relates to the manager's skills and we want to force the factors to replicate the hedge fund return.

The betas and the factors can either be fixed or change over time. In the method fixed-weight, are both beta and the factors determined in the in-sample period and kept fixed in the out-sample period. This method suffers form look-ahead bias because the entire historical data are used to construct the clone. This method can also suffer from non-stationary, which can have an effect on the validation when comparing the clones return with the hedge funds return.[39] For this reason the fixed-weight method has been deselected in this study.

[39]Hasanhodzic, J. & W. Lo, A. (2007)

Instead this study will focus on the factor-based method that addresses these issues. This is done by changing the beta and keeping the factors fixed over time; this method is called **Rolling Window**. This study will also extend this method by changing both the factors and the betas over time. This will be performed on the clones that, under rolling window, perform the worse. This method has been included to see if this will change the outcome and provide us with a more satisfying fitted clone. In the extended version of the rolling window, the factors will be added to the replicated portfolio if they become significant for the model and removed if they become insignificant.

This extension of the rolling window does therefore take into account that hedge funds do not have constant exposure over time by changing both the risk factors and their associated sensitive; this will therefore give a more realistic picture because hedge funds portfolios are adjusted over time.

If the two differed variation and if the significant of the factors will have a financial impact will be discussed in a later section. But it has been found to be worth looking into both methods to see if the performance can became more satisfying by "just" adjusting the factors over time.

4.2 Rolling Window

This part focuses on the method described and used by Hasanhodzic, J. & W. Lo, A. (2007). They tried to replicate a number of hedge fund indices based on six factors and by using two different factor-based methods: fixed-weight and rolling window.

As mentioned before, we will use rolling window to overcome the look-ahead bias. We will also have more factors to select from when first deciding on the factors to include in the replication, but still keep the numbers of factors in the final models at a reasonable level. It can be tempting to add more factors to the model to get a better fit in the in-sample, but this can have a negative effect on the fit in the out-sample period. The explanatory power of the model will decrease with too many factors in the model, and lead to a potential over-fit. We have therefore decided to keep the variables in the final model that has a significant level at minimum 5%. This might give us a lower adjusted R^2 than we could have obtained in the in-sample period, but this will make the replicated portfolio more manageable and possibly give a better out-sample result.

The factors will be chosen and set by the use of linear regression and backward elimination on an in-sample period and kept constant thoughtout the rolling window. The overall data selection for the factors in the data collection process have also been chosen based on the hedge fund strategies.

Rolling window is as mentioned a factor-based replication method that aims to create a linear clone of the month-to-month return of a hedge fund index, by changing the sensitivity to the factors over time.
The linear clone and the associated sensitivities are found by using the linear regression model described above, equation 6, and a moving window that overlaps the sample periods. The basic idea of overlapping the sample periods is to move the "window" forward after estimation of one observation, and by the end of the rolling window all observations will be estimated with different betas and then form the clone.
The overlapping sample periods can have different length; we will, in the analysis, test different lengths, to analyze the performance and if there is a significant difference in the length of the windows.

4.2.1 Renormalization

The different approach that Hasanhodzic, J. & W. Lo, A. (2007) have to the replication theory is that they renormalize the clone returns. This is done to ensure that the clone returns have the same volatility as the original hedge fund return.[40] This will give us a fairer comparison when we later on analyze how the performance of the clone is relative to the hedge fund index. The renormalized return $\widehat{R_t}$, will be created after the linear regression and before comparing the clone with the hedge fund return, and it can be found by using the following steps:

$$\widehat{R_t} = \gamma_t * R_t^* \tag{8}$$

$$\gamma_t = \frac{\sqrt{\sum_{t=1}^{T}(R_{t-k} - \bar{R}_t)^2/(T-1)}}{\sqrt{\sum_{t=1}^{T}(R_{t-k}^* - \bar{R}_t^*)^2/(T-1)}} \tag{9}$$

[40] Hasanhodzic, J. & W. Lo, A. (2007)

$$\bar{R}_t = \frac{1}{T}\sum_{t=1}^{T} R_{t-k} \qquad \bar{R}^*_t = \frac{1}{T}\sum_{t=1}^{T} R^*_{t-k} \tag{10}$$

$$k = \ every\, month\, within\, the\, rolling\, window,\ k = 1, ..., K$$

$$\gamma = renormalization factor$$

$$R_t = Original\ Hedge\ Fund\ Return\ at\ time\ t$$

$$R^*_t = Clone\ return\ before\ renormalization\ at\ time\ t$$

When renormalizing the return of the clone, we change the leverage of the clone portfolio, meaning the that sum of the betas will no longer be equal to one, but equal to the renormalization factor. When the renormalization factor is greater than one, positive leverage is required and the clone is over invested in the factors. If the renormalization factor is less than one, negative leverage is required and the clone is not fully invested in the factors. An additional asset R_l could therefore be introduced that would represent leverage, meaning borrowing and lending, and the return of the clone will then be given by:

$$1 = \gamma_t(\beta^*_{t1} + ... + \beta^*_{ti}) + \delta_t \qquad \widehat{R_t} = \gamma_t * R^*_t + \delta * R_l \tag{11}$$

But the borrowing and lending rate depends on many factors such as the credit quality of the respective counter-parties, the riskiness of the instruments and the portfolio strategy, the size of the transaction and the general market conditions; we therefore pay attention to the fact that this factor exists, but will not be including it in this study.[41]

The next section will describe the second variation of the rolling window method that we will use in this study on the worth performing clone under rolling window.

[41] Hasanhodzic, J. & W. Lo, A. (2007

4.3 Extended Rolling Window

After studying hedge funds and the manager's way of operating, it has been found useful to look into the possibility to change factors over time, as this will reflect the way a manager would adjust his or her portfolio and model over time.[42]
The addition will, in this study, be included by extending the rolling window method described above, by adding an addition to the updating step, to reflect the adjustments in the portfolio over time.

The basic idea of the extended rolling window is the same as the method above and will also have the same restriction on significant level and money invested: beta equal to one. But instead of beforehand selecting the factors for the whole out-sample period, the factors are changed if they are no longer significant for the model and added if they become significant. Along with updating the beta's for the significant factors.
This will split the method into two steps that are slightly different from the method described above. The estimation step will be similar; when a linear clone for a moving window is found, the next observation will be estimated. The updating step will be a bit different because both factors and beta will be updated before estimating the next observation.
When all the rolling windows have been updated and all observation in the out-sample period has been estimated, the clone return will be renormalized, just as described in the previously section.

If this extension to the rolling window will have a financial improved result for the worst performing clones under rolling window when aiming to replicate the hedge fund index again, will be analysed in a later section by looking at financial measurements.

The next section will go over the statistical methods that will create ground for the rolling window, and make sure that we will have the best prerequisites to replicate hedge fund index.

[42] Kooli, M. & Sharma, S. (2012)

4.4 Multiple Linear Regression

The factor-based replication can be found by using multiple linear regression, as the clone's return can be shown as a linear regression of the significant factors and their associated sensitivities.

Linear regression is a statistical method that uses independent variables to predict the outcome of the dependent variable, by modeling a linear relationship between the independent variables and the dependent variable. If there are more independent variables describing the dependent variable, the regression is called multiple linear regression. This is the case under hedge fund replication as the clones of the hedge fund index are described by the use of more than one factor.

The multiple linear regression is given by:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_i x_i + \epsilon \tag{12}$$

$$Y = Dependent\, variable$$

$$x_1, ...x_i = Independent\, variables,\ assumed\, known\, values$$

$$\beta_0 = Intercept$$

$$\beta_1, ...\beta_i = \ slope\, coef ficients$$

$$\epsilon = Error\, term\ /\ residuals$$

This equation is similar to equation 5, which has shown that a hedge fund return can be represented on a linear form and does therefore allow us to use this statistic method to find the clone of the hedge fund index, by the use of different factors. Beside the assumption that there should be a linear relationship between the hedge fund return/clone and the factors. Should the factors also not be too highly correlated with each other. This is therefore checked before doing the replication, and can be found in section 5.2.1.

To estimate the parameters in the multiple linear model, we can use Ordinary Least Squares (OLS) or Maximum Likelihood (ML). But as Ordinary Least Squares is mathematically easier to perform, it is more appealing to use this method over maximum likelihood.

The end result for the two methods will also tend to be similar when performing a linear regression. We will therefore be using OLS when estimating the parameters in the multiple linear regression.

4.4.1 Ordinary Least Squares

To ensure that the estimated Y or return of the clone is as close as possible to the original Y/return, OLS adopts the least-squares criterion. The principle of the least-squares criterion is to estimate the betas so $\sum \epsilon^2$ is as small as possible. To have the best prerequisites for estimation the parameters does OLS also operate within the framework of the classical linear regression model and does have the following assumptions: [43]

1. The model or the parameters should be linear.
2. The independent variables X should be independent of the error term. Meaning that we require zero covariance between ϵ and each X variables.

$$cov(\epsilon_t, X_{ti}) = 0$$

3. The mean value of disturbance ϵ_i should be zero

$$E(\epsilon_t | X_{ti} = 0) \;\; for\, each\, t$$

4. There should be homoscedasticity or constant variance of ϵ

$$var(\epsilon_t) = \sigma^2$$

5. There should be no autocorrelation, or serial correlation between the disturbances.

$$cov(\epsilon_t, \epsilon_k) \;\; t \neq k$$

6. The number of observations must be bigger than the number of parameters to be estimated.
7. There must be variation in the values of the X variables.

[43] Damodar, G & Porter, D. (2009)

8. There must be no exact collinearity between the X variables. Meaning that must be no exact linear relationship between X_i and X_j where $i \neq j$

9. There must be no specification bias. The model must be correct specified

It is worth noting that ϵ is assumed to be normally distributed with a mean equal to zero and a constant variance, assumption 3 and 4. This will make OLS estimation of the coefficients equal to the ML estimation, which is also why OLS has been chosen for this study.[44]

To ensure that we have the best prerequisites to make a conclusion on the results found using the replications methods and OLS, the selected data will be tested to see if they follow the assumption under multiple linear regression. If they differ slightly from the assumption, the reader will be made aware of this, and this will be taken into account when coming to a conclusion on the ability to replicate a hedge fund index.
Before creating the linear regression, the data will be tested to see if it is normally distributed; this is done by Jarque Bera test. The second thing that is tested before creating the linear regression is stationarity. But as rolling window is a replication method that should overcome any non-stationarity in the data, we will only test for this when we use fixed-weight to select the factors to use under rolling-window. We will therefore not go into detail with the statistical test for stationarity.
The next section will describe the statistical method that will be used to test for normal distribution.

4.4.2 Normal Distribution

The data set will be tested to see if it is normally distributed. This is done because it is an assumption under multiple linear regression that would create less trustworthy results if it is not present.
The error terms will also be tested for normal distribution in this study but no action will be taken as they are more robust to violations, it will just be taken into account when concluding on the final replication.[45]

[44] Damodar, G & Porter, D. (2009)
[45] Osborne, J. & Waters, E. (2002)

If our factors are not normally distributed, it can have a negative effect on the significant test that we are using to select the best model for the out-sample period. We will test the total data set before starting the rolling window. This is done because if the total data set has a normal distribution, will a sample set taken from the total data set that we using for the rolling window also be normal distributed.[46]

To test for normal distribution we are using **The Jarque-Bera test**. This test is widely used because it takes two variables into account: skewness and kurtosis. Skewness describes the symmetry in a distribution and a normal distribution will have a skewness value of zero. Kurtosis describes how much data can be found in the tails and a normal distribution will have a kurtosis at three.
The Jarque-Bera is given by:

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \tag{13}$$

$$n = Number\ of\ observations$$

$$S = The\ samples\ skewness$$

$$K = The\ samples\ kurtosis$$

The closer the Jerque Bera test is to zero the closer the sample is to being normally distributed. The following null hypothesis and alternate hypothesis is therefore set up, where the null hypothesis can be rejected at a 5% significant level:

$$H_0 = The\ sample\ is\ normal\ distributed$$

$$H_1 = The\ sample\ is\ not\ normal\ distributed$$

Just before the Jarque-Bera test, we will look at some other basis characteristics for the data, such as mean return, median return, standard deviation and Sharpe Ratio.[47] This is done to give us a better base for selecting and deselection factors.

[46] Newbold, P. & Carlson, W. & Thorne, B. (2013)
[47] Appendix 8.18

After a closer look at the basis characteristics, the Jarque-Bera test will be performed, followed by the multiple linear regression, which is done by using backwards elimination and elimination based on significant level, which will be described in the next section.

4.5 Best Fitted Model

The best fitted model is found by backwards elimination and is further cut down by the use of elimination based on significant level. **Backwards elimination** is a stepwise regression method that finds the best fitted model based on model search using the Akaike information criterion (AIC). AIC is a measurement that describes the goodness of the fit, by taking the number of parameters into account.[48] In stepwise backward elimination, the regression starts with all factors, and step-by-step remove the factors that contribute the least to the model, and stops when the AIC has the best fit.

The model found by running the stepwise backward elimination will give us an $\boldsymbol{R^2}$ and **adjusted $\boldsymbol{R^2}$**, which describe the goodness of the fit of the regression, meaning how much in percentage of the total variation in the depending variable is described by the independent variables joint together. R^2 will lie between 0 and 1, meaning that if R^2 is equal to 1, the regression will have a 100% fit. So, in general, it can be tempting to say that the closer R^2 is to 1, the better the regression fits the dependent variable. In this case, will it be that the in-sample of the clone will fit closely to the hedge fund index in the in-sample period. But this is not always true, because R^2 is a non-decreasing function of the numbers of independent variables added to the regression. We will therefore only focus on the adjusted R^2 in this study; this gives us information about the goodness of fit but takes into account the numbers of parameters added to the model as well. We do aim to get a high adjusted R^2 in the in-sample period, but as mentioned before we also are aware that the goodness of the fit in the in-sample is not necessarily equal to a good fit in the out-sample period. We do therefore only allow independent variables that has significant level below or equal to 5%. Significant value, also called p-value, are the smallest level of significance for which the observed data indicate that the null hypothesis should be rejected. [49]

[48] Dalgaard P. (2008)

[49] Wackerly, D. & Mendenhall III, W. & Scheaffer, R. (2008)

The null hypothesis and alternate hypothesis is given by:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

This means that if the p-value is lower than 5%, the null hypothesis can be rejected and that the dependent and independent variable do not have a significant relationship. If the p-value is higher than the 5% level, we cannot reject the null hypothesis, and there is no significant relationship between the dependent and independent variable.

After the use of backward elimination using AIC and significant level, we are left with the best fitted model for the sample period. But before making the final financial analysis, based on financial characteristics and some performance measurements, the models error terms will be tested for homoscedasticity and autocorrelation.

4.6 Error Term Test

To ensure that we end up with the best multiple linear regression model, and that the OLS assumption are kept, we will test whether assumptions 4 and 5 apply to our model.

4.6.1 Homoscedasticity

The fourth assumption under OLS states that the variance for the error term should be some positive constant number equal to σ^2.[50] If this applies, there will be homescedasticity; if not there will be hetroscedasticity.
If hetroscedasticity occurs, it will have two main consequences for the model that we have found using the selection method described above. The OLS estimations will no longer be unbiased, meaning that they will no longer be the best fit. It would be possible to find a new model that describe the hedge fund index in a better way. Second could it create misleading and wrong hypothesis testing because the standard error would be incorrect.

[50] Damodar, G & Porter, D. (2009)

These two consequences will the reader be made aware of, if there would be hetroscedasticity in our models. But we will continue to work with the model, as the OLS estimations will still be linear.

To check for homescedasticity/hetroscedasticity we will use Breusch-Pagan test. This test is sensitive to the assumption of normality, but this is tested and taken into account previously. This test is therefore selected in this study because it can be performed using the statistical tools that we have.

The Breusch-Pagan test uses a variance function and sets up a null hypothesis that homoscedasticity will be present. The variance function and hypothesis are given by:[51]

$$\sigma_1^2 = \alpha_1 + \alpha_2 Z_{i2} + \alpha_n Z_{nt} \tag{14}$$

$$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_n = 0$$

$$H_1 : At\ least\ one\ \alpha\ is\ \neq 0$$

This means that if the p-value is less than 5%, we can reject that there will be homoscedasticity. We will then have to be aware of this in the analysis through this study. But we are also aware that the model fitted for the rolling window is chaining over time, and might not therefore always be the best fit in each sample period.

4.6.2 Autocorrelation

The fifth assumption under OLS is that there should not be any autocorrelation or serial correlation between the error terms. This means that ϵ_t and ϵ_k should not have any positive or negative correlation with each other. If this would be visual seen must there not be any symmetry between the error terms.
If autocorrelation occurs in the error terms, the forecast will be inefficient. This means that we have missed out of some information that we should have taken into account in the model to get the best forecast.

[51] Yobero, C. (2016)

Even though we will be missing information if there is autocorrelation, the error terms will still be unbiased, so the model can still be used, but we could have ended up with a better result if there was no autocorrelation.

Ljung-box will be used to test for autocorrelation. Ljung-box is chosen in this study because it is a modification of the Box-test that better captures the absence of autocorrelation and serial correlation, by specifying the lag k. Ljung-box and the hypothesis is given by:

$$Ljung - box = n(n+2)\sum_{j=1}^{m}\frac{r_j^2}{n-j} \tag{15}$$

$$H_0 : \ No\ Autocorrelation$$

$$H_1 : \ Autocorrelation$$

This means that if the p-value is more than 5%, we cannot reject that there will be no autocorrelation, and the model's forecast values are therefore seen as the best possible values for the linear regression model.

When the error terms have been tested, we will move on to the performance measurements that give us a foundation for our conclusion on the performance of the clone compared with the hedge fund index. The performance measurements will be described in the next section before moving into the analysis.

4.7 Performance Measurements

We will, in the analysis, conclude and discuss the performance of the clones compared with that of the hedge fund indices, using the following performance measurements: *Mean return*, *Correlation*, *Standard deviation*, *Sharpe ratio*, *RMSE*, *Tracking Error*, *Information ratio* and *Theil's Inequality coefficient*.
Not all of the measurements will be described in this section, but we will describe some of the measurements that we found worth explaining.

4.7.1 Sharpe ratio

The Sharpe ratio is a financial measurement that compare the return of the investment to the risk. Meaning that the higher the Sharpe ratio is the better has the investment been compared to the risk taken. The risk is set by the historical volatility (standard deviation) compared to the risk-free rate. To ensure that the risk has enough data to give a fair picture of the actual risk, the Sharpe ratio should not be calculated for too-small a sample size. The out-sample size for this study is seen to been fair, but should not be less.
The Sharpe Ratio is calculated by:[52]

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p} \tag{16}$$

$$R_p = Return\ of\ the\ portfolio \qquad R_f = The\ risk - free\ rate$$

$$Standard\ deviation\ \sigma_p = \sqrt{\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}\right)} \tag{17}$$

$$x_i = The\ return\ at\ time\ i \qquad \bar{x} = The\ mean\ return\ for\ the\ portfolio$$

$$n = The\ number\ of\ observations$$

When we are comparing Sharpe ratio for the clone and the hedge fund, we are looking at how good the two portfolios are to remove systematical risk by spreading out their investments. A positive Sharpe ratio means that the portfolio has taken on extra risk and achieved an excess return. If the returns for the portfolio are more stable over the sample period, the Sharpe ratio will be bigger as the portfolio will have achieved a stable return by taking on some extra risk.[53] If the Sharpe ratio is negative, the portfolio will have a lower return than the risk-free rate. Negative Sharpe ratio can be difficult to compare. It is therefore important to look at the number behind the Sharpe ratio to get an idea of the negative value.

[52] www.investopedia.com.

[53] www.finansdanmark.dk

4.7.2 Root Mean Square Error

Root Mean Square Error (RMSE) is more a statistical analysis tool that is used to show the differences between predicted values and the actual values. It is therefore often used when forecasting and under linear regression, because it provides us with information about how close the data is to the best fitted model. RMSE can therefore be incorporated as a tracking error of the clone in this study. The closer RMSE is to zero the better fit will the clone be to the hedge fund index.
RMSE is given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Predicted - Actual)^2} \tag{18}$$

4.7.3 Tracking Error

The tracking error describes the return differences between a portfolio and a benchmark - in our case, the difference between the clone and the hedge fund index.
This is an important tool for this study because it shows how our clone performed relative to the hedge fund index that we were aiming to replicate. It provides us with a measurement of who the consistency of our clone is compared to the hedge fund index over a given time period. If the Tracking Error is low our clone is closely following the hedge fund index, but if the Tracking Error is high our clone will not be consistent with the hedge fund index.
The tracking error is calculated by the following formula:

$$Tracking\ Error = \sigma(Clone_i\ Returns - Hedge\ Fund_i\ Returns) \tag{19}$$

4.7.4 Information Ratio

To analyze our ability to create excess return relative to the hedge fund index, we will use Information Ratio. Information Ratio is a measurement that is used to analyze the portfolio's ability to generate return at a given risk. It is compared to a benchmark, in our case a hedge fund index, and shows how our chosen betas are generating more or less return than the benchmark.

The Information Ratio is given by:

$$Information\ Ratio = \frac{Clone_i\ Return - Hedge\ Fund\ Index_i\ Return}{Tracking\ Error_i} \tag{20}$$

The Information Ration is calculated by the use of the Tracking Error. This means that the lower the Tracking Error is, the better the Information Ratio will be and vice versa.

4.7.5 Theil's Inequality Coefficient

Theil's Inequality Coefficient is also a statistic performance measurement that can be used under forecasting to determinate how good the forecast values are compared to the observed values - in our case, how the clone is doing compared to the hedge fund index. This measurement has been chosen as it gives us information about whether our replications are good or a random guess would have been just as good. The closer Theil's Inequality Coefficient is to zero, the better the forecast value. The closer the measurement gets to one, the more we might just as well have made a random guess.
Theil's Inequality Coefficient is given by:

$$U = \frac{RMSE}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Predicted)^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Actual)^2}} \tag{21}$$

All the methods and models used to replicate the selected hedge fund indices have now been described, as well as the performance measurements that we will use to analyze how well the clones are performing compared to the hedge fund indices. The next section will be an analysis where we aim to replicate the hedge fund indices described in the data section. We will focus on the finance performance and measurements described in the last bit of this section, but also focus the replication method and shortly make the reader aware of any statistical fluctuations there might be in the multiple linear regression model.

5 Analysis

5.1 In & out sample periods

The analysis will focus on three different periods, all within the chosen period for this study, December 1999 through August 2019. The data is split into three periods to analyze the ability to replicate before, during and after the 2008 financial crisis.

The first period is before the financial crisis in 2008 - pre-crisis period. The in-sample period is from December 1999 through November 2007. These 96 observations are the foundation of the linear regression, which is used during Rolling Window.

In figure 4 all the factors are shown as indices for the first in-sample period. Almost all the factors have a small change over time, but a few have drastic changes. The most obvious one is the spread between Morgan Stanley Europe Large Cap index and Morgan Stanley US Large Cap index; it is because of the close index-value development that a small differentiation has a huge impact. The development between the two indices is shown in Appendix 8.1.

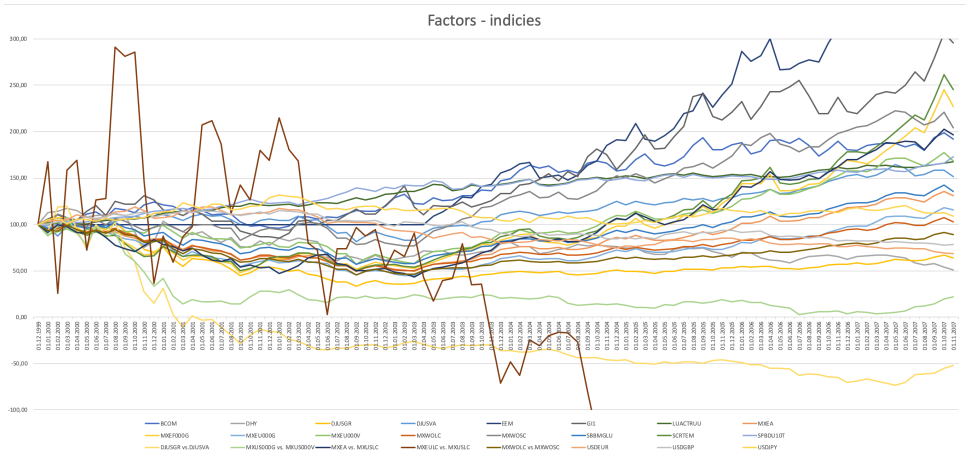


Figure 4: Factors In-sample 1 - 1999 to 2007

The out-sample period regarding this in-sample period is at 36 observations and is from December 2007 through November 2010. In this period, the global financial crisis had a huge influence on the financial markets. The crisis started with the subprime mortgage crisis in 2007, and with the Lehman Brothers Bank collapse in 2008, the crisis took the world by storm, with massive consequences for the financial market and the macroeconomic in total.[54]

[54] Rosefielde, S. & Kuboniwa, M. & Mizobata, S. (2011)

The second in-sample period includes the global financial crisis - crisis period. The volatility of the financial crisis returns therefore have a huge effect on the in-sample period. The in-sample period is from December 1999 through November 2010, and has a total of 132 observations. The out-sample period is from December 2010 through November 2013. Looking in isolation at the out-sample period, the hedge fund indices are all increasing through the out-sample period except Equity short bias.

The third and last in- and out-sample period is the post-crisis period. The in-sample period's first observation is in December 2009 and goes through August 2016 with 81 observations. The out-sample period has 36 observations as the two other test periods, and are therefore going from September 2016 through August 2019. We will later in the analysis look at the different hedge fund replications in all three periods, and will also look at the different pros and cons for replicating in the given periods with the determinate factors.

All the three periods are shown graphically in figure 5, where the timeline is shown for both in-sample and out-sample for all three periods.

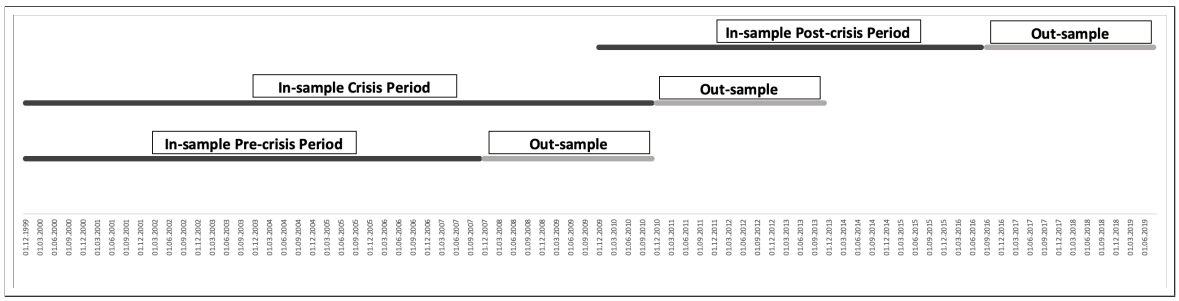


Figure 5: In & out-sample periods

5.2 Test of in-sample period

As mentioned in the method section, it is important to test the data before the linear regression, to ensure that the data is given the best possible prerequisites for replicate. We will therefore test the in-sample period before the regression.

5.2.1 Correlation

Before determination on the final factors for the analysis, we started out with 42 factors, which where all described in the data validation section. They are all checked for correlation and deselected based on too-high correlation or indices that could be described elsewhere. The 42 factors were selected based on the hedge fund strategies and to cover a reasonable board cross-section of risk exposure within the typical hedge fund. In Figure 6 it is visible how all 42 factors are correlating with each other. Where the correlation are higher than 90% or lower than -90%, the cell is highlighted with a red color.

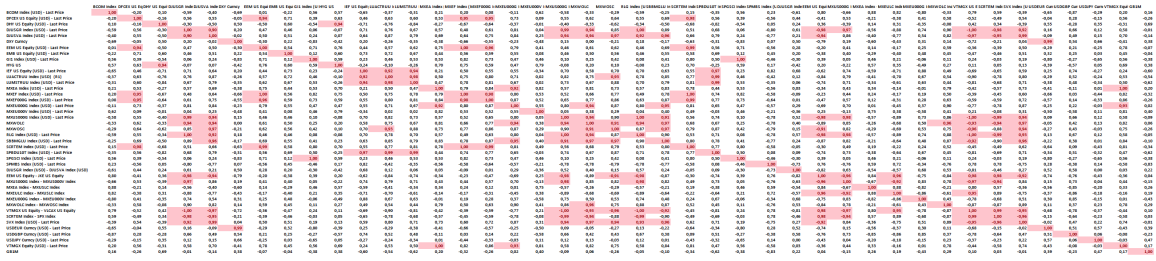


Figure 6: Correlation all factors

Most factors deselected are within the same investment area. Some factors are also deleted because of poor data, meaning not enough observations or too many outliers which are making the replication not possible. Figure 7 shows the correlation matrix after the elimination process, with the final 25 factors to be used in the analysis. There are still some correlations, but because of different investment areas those are the factors kept, knowing that the linear regression done later on will eliminate non significant factors for our model.

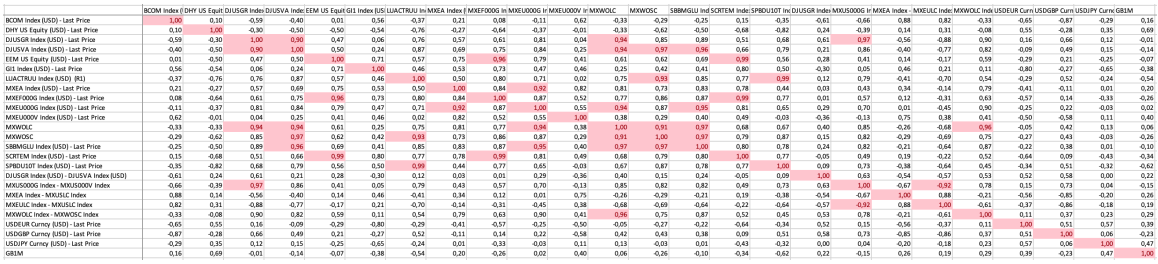


Figure 7: Correlation all factors

5.2.2 Normality

As mentioned in the method section, the in-sample periods will be tested for normality. This is done by looking at the skewness and kurtosis and the relationship with a normal distribution. In RStudio, we are using the packages *"matrixTests"*. In this package include the Jarque Bera test and is performing a goodness of fit test for normality.[55] This test returns, among others, a p-value regarding the null hypothesis for which the sample is normally distributed. 15 of the factors in the pre-crisis period have a higher p-value than 5%, and the null hypothesis can therefore not be reject of normal distribution in the factors data frame. The same applies for the crisis period and the post-crisis period, where the test for normality returns respectively 5 and 22 factors with a p-value higher than 5%.

The final clones can therefore contain some non-normal distributed factors, which is not completely satisfying. But the knowledge of normality in the factors gives the possibility of looking at the final clones after the linear regressions and then conclude on the result with the normality in mind. The crisis period has a low amount of factors that are normally distributed, which can have an effect on the performance on the clones in that period.
Appendix 8.2 shows the goodness of fit test from RStudio, on all our factors with the data from the in-sample periods. Shown here is the difference in kurtosis and skewness, which are the important values for if the data is fitting a normal distribution.

5.3 Performance Measurements

In the entire analysis we will focus on the performance measurements over the 11 different hedge fund indices and furthermore for all of the three periods. As described in the method section, we will focus on several performance measurements - respectively Mean, Standard Deviation, Sharpe Ratio, Correlation, Root Mean Square Error, Tracking Error, Information Ratio and Theil's Inequality Coefficient.

All the performance measurements for all strategy clones and hedge fund indices can be seen in appendix 8.3.

[55] cran.r-project.org

5.4 Linear Regression

5.4.1 Main index - EHFI251

The first hedge fund index we try to replicate, by the use of rolling-window, is the overall hedge fund index from Eurekahedge. The Bloomberg tricker for this Eurekahedge's flagship is EHFI251.

Pre-crisis period

Before starting the linear regression and finding the best fitted model, the risk-free asset is isolated. By isolating the risk-free asset, the models are making sure that the total amount is invested in the strategy. Practically, this means that the difference between the sum of the associated sensitivities and 100% are put into the risk-free asset. There is therefore the possibility of going both short and long in the risk-free asset.

Next, the significant level is set at 5%. To get backwards elimination to secure a minimum of 5% significant level, the corresponding k-value is set. This is found through RStudio and with the function *qchisq*. The 5%-level returns a k-value at:

$$k = 3.841459 \tag{22}$$

To find the model for replication, backwards elimination are used, and this is done now with the 5% significant level and for the pre-crisis period the linear regression model for replication is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * GI1 + \beta_3 * LUACTRUU + \beta_4 * MXEU000G + \\ \beta_5 * MXWOLC + \beta_6 * SPBDU10T + \beta_7 * MXEULC.vs..MXUSLC + \\ \beta_8 * USDEUR + \beta_9 * GB1M \end{aligned} \tag{23}$$

The corresponding significant level to this function is going from 2.6% down to less than 0.00%, by removing the factors that are not significant at a 5% level. In this model, we have both currency indices, spreads, commodities and much more. This gives a good base for replicating the main index because of the overall strategy.

The last step before creating a clone with rolling window is testing for normality in the residuals, for heteroscedasticy and autocorrelation. As mentioned in the method section, if there is heteroscedasticy or autocorrelation in the model, it is possible that there can be a better model for replicating. Furthermore, the standard error could be incorrect which could lead to wrong hypothesis testing.

In the pre-crisis period for the main index, the normality in the residuals are seen normal. The plot in Figure 8 shows that the residuals are following the normality path on an acceptable level.

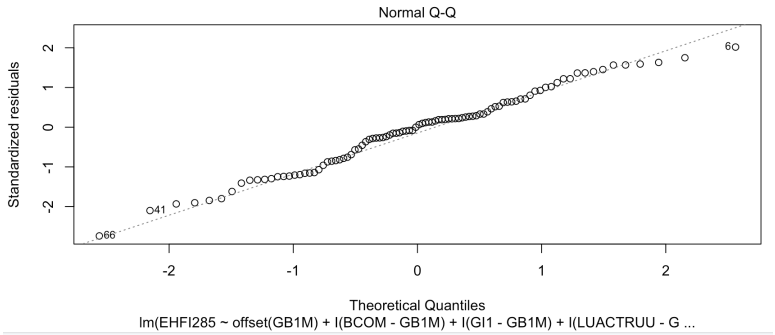


Figure 8: Residuals plot Main Index pre-crisis period

Testing for heteroscedasticy is done through RStudio by using the function *bptest*, which stands for the Breusch-Pagan test. This function returns a p-value on 60.68%, and the null hypothesis about homoscedasticy cannot be rejected, this means that the fitted model is still the best fit.
The residuals is tested for autocorrelation with the function in RStudio called *Box-Ljung test*, and for the pre-crisis period, the p-value is less than 5%, which means that the null hypothesis that there is no autocorrelation can be rejected. This means that the fit is not the possible best fit. This means that the model is missing out of some information that should have been taken into account. So, we are still able to use the linear regression model, this is just something that could be adjusted to maybe get a better end result.

After all the previous tests, the beats are now estimated for each window for the entire out-sample period, and we are aware of that the fit is possibly not the best, because of the

autocorrelation. This is however possible to change over the out-sample period, when the window is "rolling".

To estimate betas for each period, the method Rolling Window is used to respectively estimate all the periods by "rolling" a window over the out-sample period. As described in the method section, a 36 period long window is constantly rolling to estimate the coming beta.

After estimating betas for the entire out-sample period, the last step before making the clone is to renormalize the returns. Renormalizing the returns is to make sure the volatility between the clone and the hedge fund returns more equally. This step is done for all periods in the out-sample period one by one, by using the formula as mentioned in the method section, 4.2.1.

The renormalization is calculating a γ pr. period, and this value is multiplied with the returns to establish the total return for the clone.

Results Main Index pre-crisis period

In the first out-sample period, the Main index are going from index 270 to 322, and the mean return is given with 0.535%. The clone has a mean return of 0.005%, which shows that there is a great difference in the replication mean return and the actual mean return for the hedge fund index.

When looking at the graphical development of the clone and the hedge fund index, it is clear that the replication is not with a good fit, and under-performs the hedge fund. From December 2007 to September 2008, the clone and the hedge fund index are much alike, but during the financial crisis, the gap between the clone and the hedge fund is increasing, and after the financial crisis the returns of the clone are constantly lower than the hedge fund, which among others is because of the financial crisis influence in the 36-month period. This is shown in Figure 9. This illustration can also be seen in the performance measurements, where, as mentioned earlier, the mean returns are not alike and the standard deviation is more equal. The standard deviation of the hedge fund is given by 2.016% and the clone has a standard deviation on 1.974%, so looking in isolation at this, the clone seems to fit very well to the hedge fund.

This means that the clone has a volatility around its own means almost equal to the hedge fund index.

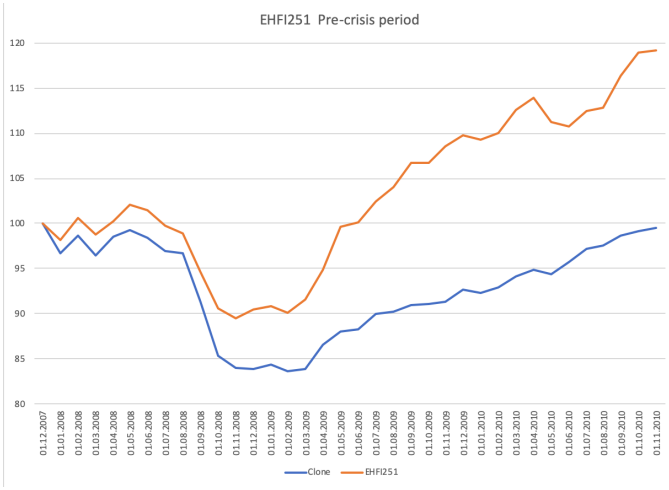


Figure 9: Pre-crisis period Main Index Hedge Fund vs. Clone

But looking at all other statistics, the clone and the hedge fund index are not similar. The hedge fund main index has a Sharpe ratio of 1.32%, compared to -25.872% for the clone. This means that the hedge fund index takes on extra risk and achieves an excess return, where the clone does not achieve an excess return, because the return is less than the return from the risk-free rate, which is not a good characteristic for a portfolio. Furthermore, the correlation between the returns is only 64.34% and with an adjusted R^2 value on 87.3% before the beta estimating, the rolling window method is not delivering as well as the adjusted R^2 value implies. Finally, the RMSE, tracking error and Theil's inequality coefficients are acceptable, at 0.01149, 1.046% and 0.7938 respectively. This implies that the fitted clone is performing acceptably when looking at how close and how consistent the forecasted values are compared to the hedge fund, but when looking at the total index development it is clear to see that this does not apply to the whole period. Just at it is shown in the index graph will the information error be negative compared to the hedge fund. As the clone are not able to generate excess return relative to the hedge fund. Totally the clone are not satisfying, and the fit with the hedge fund are not on a high level.

Crisis period

The difference between the pre-crisis period and the crisis period is that the financial crisis is within the crisis period and more betas is therefore estimating with the financial crisis within the window, than in the pre-crisis period.
The model for this period is found via backwards elimination, and is given by:

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DHY + \beta_3 * LUACTRUU + \beta_4 * MXEU000G + \beta_5 * USDEUR + \beta_6 * USDJPY + \beta_7 * GB1M \quad (24)$$

But compared to the pre-crisis period where the adjusted R^2 value was on 87.3%, the model for the replication of the clone in the crisis period has an adjusted R^2 value on 84.9%, both of which are at an acceptable level. The goodness of the fit is therefore set to be a bit worse than in the pre-crisis period, but looking at the statistic and graphically, it shows that the clone is an acceptable fit to the hedge fund index, in the out-sample period.

The mean return for the clone is 0.176%, and that is 0.272%-points less than the hedge fund index. The standard deviation is again much equal with 1.228% for the hedge fund and 0.986% for the clone, which is acceptable. The correlation is 94.532%, which is relatively high. This is also visible to the eye, when looking at figure 10, where the path of the clone and the hedge fund are very much alike, but the distance are increases slightly for the second half of the period, which indicates not similar returns.

The RMSE and the tracking error are 0.00617 and 0.519%, which are on a good level. The Theil's inequality coefficient are at 0.05406, which also are considered good. Meaning that the clones forecasted values are performing good on how close and consistent they are relative to the hedge fund index. The only statistics which shows that the clone is not an acceptable fit is the Sharpe ratio. Here, the hedge fund has a Sharpe ratio at 33.021% compared to 12.557% for the clone, meaning that hedge fund is generating more excess return relative to the risk than the clone is, but in this period compared to the pre-crisis period is the Sharpe ratio is positive, meaning that the clone is generating excess return relative to the taken risk.

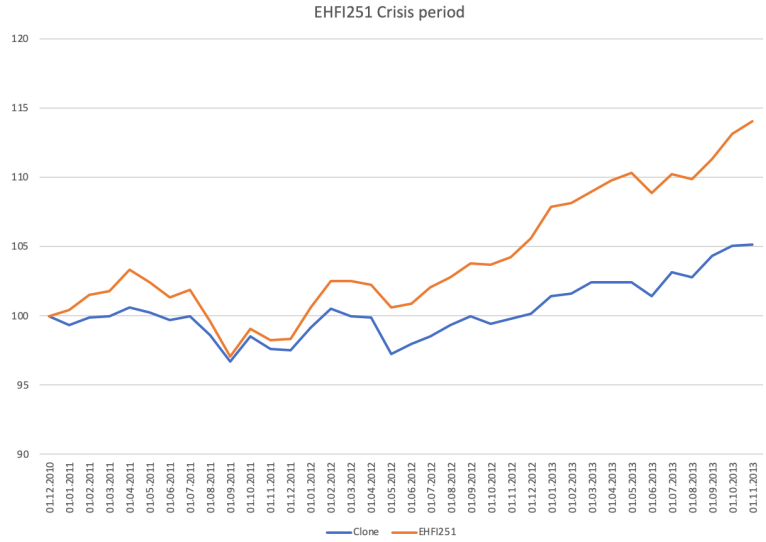


Figure 10: Crisis period EHFI251 vs. clone

Overall, the clone for the main index is an acceptable fit to the hedge fund, but it is under performing the hedge fund index, which is also shown in a negative information ratio.

Post-crisis period

In the post-crisis period does the model has an adjusted R^2 value on 87.9%. The factors used for the clone model are given by:

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * EEH + \beta_3 * GI1 + \beta_4 * MXWOLC + \\ \beta_5 * SBBMGLU + \beta_6 * SPBDU10T + \beta_7 * GB1M \tag{25}$$

Looking at the statistics, the clone has a correlation index at 79.4%, but the standard deviation is almost equal. The Sharpe ratio for both the hedge fund and the clone are very negative, but that is because of the risk-free asset. In the post-crisis period the risk-free asset are increases from a rate equal to 0 in May 2015 (0.183% in September 2016) to 2.088% in August 2019. Therefore, the Sharpe ratio is very low because of the return compared to the risk. In the portfolio for the clone,a long position on the risk-free asset is taken, with different weights through the period, and these positions make the clone over-perform the hedge fund index, which is also seen in the mean and the information ratio, which is positive, 0.4169.[56] The mean for the

[56] Appendix 8.4

hedge fund is at 0.323% and the clone has a mean of 0.592%. The higher return is also because of the stagnation in the hedge fund return, where for the last 1.5 years the returns are about 0.

The light conclusion needs to be taken for this period as the clone is having trouble fitting the hedge fund, primarily because of the risk-free asset. However, when looking at how the forecasted values are performing compared to hedge fund, the RMSE and the tracking error are at a good level with respectively 0.00693 and 0.645%. Theil's inequality coefficients are also on a good level with 0.069.

Overall strategy replication

The first replication strategy has, overall, shown some possibilities but also some weaknesses. The impacts of the risk-free asset in the post-crisis period and the financial crises in the pre-crisis period are at a high level, which is also seen in the results. This was anticipated but still it was possible to make some relative acceptable fits. The rest of the analysis will be divided into strategies, to analyze which strategies give the best replication opportunities.

Finally, the overall strategy hedge fund index is not showing any sign of specific asset class within the index, but this study's replications show that the index most likely has the currency between US Dollar and the Euro within the index at every period. Furthermore, the different markets, such as emerging, US, EAFE and developed, are possibly all represented within the index. As earlier mentioned in the data validation section, the overall strategy index contains 2447 constituent funds, and therefore it makes intuitive sense that all strategies are represented within the index.

5.4.2 Long Short Equity - EHFI252

Pre-crisis period

The replication for the long short equity strategy index is, after backwards elimination, given by the model:

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \beta_4 * GI1 + \beta_5 * LUACTRUU + \beta_6 * MXEU000G + \beta_7 * MXWOLC + \beta_8 * MXWOSC + \beta_9 * SBBMGLU + \beta_{10} * MXEA.vs..MXUSLC + \beta_{11} * MXEULC.vs..MXUSLC + \beta_{12} * USDEUR + \beta_{13} * GB1M \quad (26)$$

This is considered a large amount of factor for the clone, but as the significant level to cut off factor has been predetermined, does we stick with this model. But are aware that more factors could be cut of to make the model more manageable.
The model has an adjusted R^2 on 90.97%, which is at a good level for the future replication.

Like in the overall strategy hedge fund index pre-crisis period, the clone for the long short equity are having trouble with the replication after the financial crisis.[57] The correlation between the hedge fund and the clone is 70.575%, but if the period stopped in March 2009, the correlation would be 94%; after the financial crisis, the clone cannot establish a return on the same level as the hedge fund.
The period after the financial crisis is also the reason for the less good performance measurements for the clone. The mean return for the hedge fund for the period is 0.315% and to compare the clone have a negative mean for -0,162%. Furthermore, the Sharpe ratio for the clone is much more negative than the hedge fund, with a value of -26.686%. As the standard deviations are much alike for the clone and the hedge fund, the more negative Sharpe ratio means that the clone returns are performing worse than the hedge fund compared to the risk-free rate.
These performance measurements are indicating a fit that is not satisfying and with an adjusted R^2 on 90.97% is it clear that the clone are not as good as first anticipated, which could be an indication of too many factors in the model, as the out-sample is not performing as wished.

[57] Appendix 8.5

Crisis period

In the crisis period, the model for the replicate the long short equity hedge fund index are fit fewer factors than in the pre-crisis period, which is more manageable and is often providing a better fit in the out-sample period. The model is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSVA + \beta_3 * EEM + \beta_4 * GI1 + \\ \beta_5 * LUACTRUU + \beta_6 * MXWOSC + \beta_7 * SBBMGLU + \\ \beta_8 * USDEUR + \beta_9 * GB1M \end{aligned} \tag{27}$$

The adjusted R^2 for this model is 90.39%, which is almost equal to the pre-crisis period, but when analyzing for normality in the residuals there is seen some outliers which can have an effect on the result. The outliers form the normal distribution are seen in appendix 8.15.

This model also, in the Breusch-Pragan gives a p-value under 5%; the null hypothesis for homoscedastic can then be reject, meaning that it might be possible to find a model that explains the hedge fund index in a better way. So, in general, the statistic behind the linear regression is not good, which does not give the best prerequisites for replication using rolling window, but we are still able to continue with the model found under backward elimination, just keeping in mind the missing assumption under OLS.

Looking at the result for the replication of the long short equity hedge fund index for the crisis period, the correlation is at a very high level with the value of 98.764%. This high correlation gives some relatively good results on mean return Sharpe ratio and standard deviation, but this clone is still behind the hedge fund index, but gives more acceptable results than the pre-crisis period.
Looking at the index graph in appendix 8.5, it is visible that the clone is at some point also over-performing the hedge fund, but still most of the time under-performing, which is also seen in the results in a negative information ratio, -0.301. This shift in the performance also gives a tracking error of 0.9%, which is still acceptable.
Overall, the replication clone for the long short equity hedge fund index has an acceptable fit.

Post-crisis period

In the post-crisis period for the replication, the model is at a really high level of the adjusted R^2 on 94.01%, but is also at the higher end of the numbers of factors which could possibly lead to a bad clone.

The model for the post-crisis period is given by the equation:

$$\begin{aligned} R_t^{Clone} = \beta_1 * DJUSVA + \beta_2 * EEM + \beta_3 * MXEA + \beta_4 * SBBMGLU + \\ \beta_5 * DJUSGR.vs.DJUSVA + \beta_6 * MXUS000G.vs..MXUS000V + \\ \beta_7 * MXEA.vs..MXUSLC + \beta_8 * USDEUR + \\ \beta_9 * USDJPY + \beta_{10} * GB1M \end{aligned} \tag{28}$$

The correlation for this clone is negative, with a value of -14.659%, which is very bad. This correlation is enough to see that the replication is not good and all the performance measurements are following the correlation with bad results.

What makes the clone perform badly in the out-sample period is the spread between MXEA and MXUSLC, which have a high influence on the bad correlation. In the three months of November 2018, December 2018 and January 2019, the two indices are so close to each other that the spread is respectively positive, then negative and then positive again.[58] The rolling window method cannot capture this single negative observation, and the impact is huge.

To overcome this obstacle, the extended rolling window is used, as described in the method section. Instead of running a linear regression once, and then only estimating betas for all periods, the model is extended, so that after every period of estimating betas, the linear regression will be made.[59] This gives the possibility to capture small changes in the factors, because if the observation's negative influence on the model is high enough, the significant level for the factor will increase, and the factor will be excluded for the model because of the negative influence on the AICs.

After using the extended rolling window, the observation in December 2018 is excluded from the model. The spread was included in the model in March 2018 and then again in February

[58] See appendix 8.16

[59] Appendix 8.19, Figure 38

2019. This is changing the outcome positively, and the correlation changes from negative to 92.70%. This is a huge improvement, and furthermore the RMSE and the tracking error is changing from 0.06791 and 6.754% to 0.00852 and 0.866% respectively, which also verifies that the clone from the extended method has a much better fit than the clone from the rolling window replication.

This change in the clone is shown graphically in figure 11.

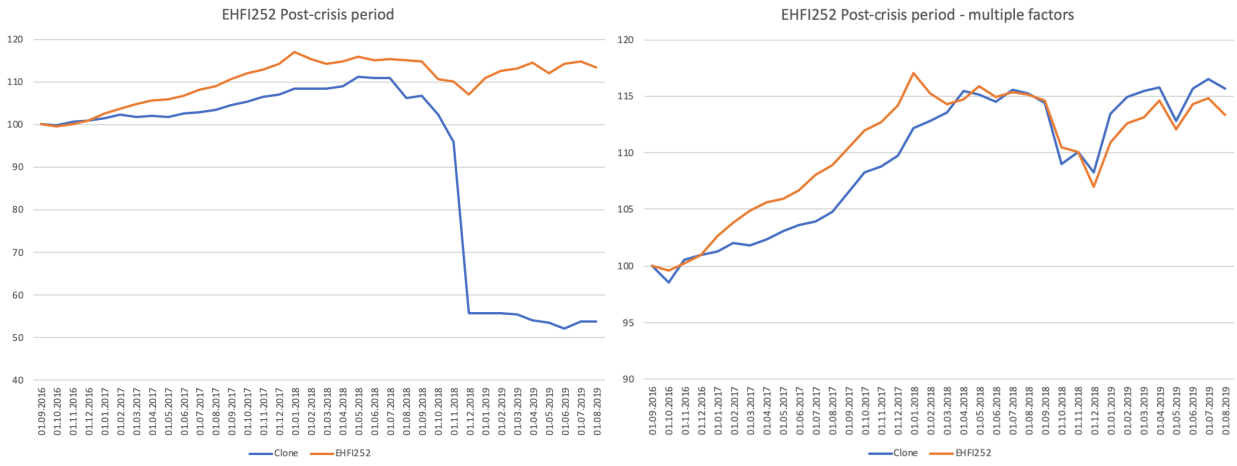


Figure 11: EHFI252 RW clone vs. extended RW clone

Long Short Equity strategy replication

After trying to replicate the Long Short Equity hedge fund index, it is seen that the method is not very consistent. From period to period, the change in quality was high, and it is therefore not possible to conclude if the strategy is replicable or not. However, we find that Dow Jones US Value Index most likely are within the index with a short position, and furthermore the S&P Global Broad Market Index are most likely also represented within the index. This is also applicable for when the method is changed to the extended rolling window, but this method also shows that there is a possibility that the Morgan Stanley Capital International Europe Value Index is represented within the index.

Looking at the different asset classes, it is not a surprise that the equities are most likely present with more than 50% of the index, and the different markets are mainly US and Broad markets.

5.4.3 Macro - EHFI253

The third hedge fund index is the index for macro strategies. As through the whole analysis three different periods will be used, and try to replicate the hedge fund index for each period to see if it is possible to obtain some knowledge about the fund, and create similar results with the replication clone.

Pre-crisis period

After running the linear regression, the model for this period regression (chosen on a 5% significant level) is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \beta_4 * EEM + \\ \beta_5 * GI1 + \beta_6 * SBBMGLU + \beta_7 * SPBDU10T + \\ \beta_8 * USDJPY + \beta_9 * GB1M \end{aligned} \tag{29}$$

The Adjusted R^2 for this model is not given the best conditions for a good fit, but still with an acceptable value on 76.88%.

The results for this period are not very satisfactory. The correlations are negative and the rest of the performance measurements are not giving any signs of a good fit. The reason for the bad result is mainly because of the risk-free asset. The first observation of the out-sample period, the beta for the risk-free asset are 7.03%, and with the risk-free asset with a rate on 2.82%, the impact are not very high. But in the end of the period, the beta for the risk-free rate increases to a long position of 86.27%, for trying to reach the high return of the hedge fund index. This would have given a good result, but because of the risk-free rate are going towards a 0% rate, the impact is inefficient, and the return for the clone is never at a high level.

This is given a mean return of -0.101% and a Sharpe ratio on -35.565%, compared to 0.68% and 14.588% for the hedge fund. This means that the clone does not generate an overall positive return or excess return, which is also seen in the graph in appendix 8.6, where the clone is clearly under-performing the hedge fund and deliveries a index below 100 at the end. This clone is therefore one of the worst performing clones in the pre-crisis period. Given that

none of the statics assumptions under the linear regression and OLS are giving any alert, we may assume that this strategy in the pre-crisis period is difficult to replicate; this might be due to the chosen method or other unknown factors.

Crisis period

The crisis period are given a much better fit than the pre-crisis period. Taken the financial crisis within the in-sample period, are given a better result, than have it in the out-sample period.

The model for the crisis period replication has many factors with a 5% significant level. The model has an adjusted R^2 on 68.46%. These two things do not indicate a good model, as too many factors can have a negative influence on the out-sample period and a lower adjusted R^2 indicates a poor fit in the in-sample period. The model is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \beta_4 * GI1 + \\ \beta_5 * MXEA + \beta_6 * MXEU000G + \beta_7 * MXWOLC + \beta_8 * SCRTEM + \\ \beta_9 * SPBDU10T + \beta_{10} * MXWOLC.vs.MXWOSC + \\ + \beta_{11} * USDEUR + \beta_{12} * GB1M \end{aligned} \tag{30}$$

This model gives an acceptable fit with a correlation on 89.108%, but with a Sharpe ratio difference from the hedge fund of 16.874%-point, the clone does not generate as good an excess return given the taken risk as the hedge fund. The rest of the performance measurements are considered good, which means that the forecasted returns are relatively close to the hedge fund, and have somewhat the same consistency, but are throughout the whole period underperforming the hedge fund, which is both seen in the graph in appendix 8.6 and a negative information ratio.

Post-crisis period

The post-crisis period gives very similar results as in pre-crisis period. The difference the two periods between are the over-performance of the clone in the post-crisis period, with a mean return on 0.41% versus 0.289% for the hedge fund index, as opposed to the clone for the pre-

crisis period.[60] This leads to a positive information ratio. The performance measurements are, for the forecasted returns and consistency, considered acceptable, but this period does have a high negative Sharpe ratio, which in general is not a good characteristic for an investment portfolio. But the hedge fund index also has a high negative Sharpe ratio, so the clone is replicating that characteristic, but it is in general not a good performance.
The adjusted R^2 on 75.52%, and the correlation on 71.067%, are also showing an insufficient clone for the hedge fund index. The model for this replication has 15 of a total of 25 factors within the model, and in general for all three clones the number of significant factors are quite high, which can be too many factors to handle if we were to execute these replications in "real" life. Also, from previous studies in replicating hedge fund indices show that, many factors often lead to a less good fit.

The model is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * DHY + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \beta_4 * EEM + \\ \beta_5 * GI1 + \beta_6 * MXEA + \beta_7 * MXEF000G + \beta_8 * MXWOSC + \\ \beta_9 * SBBMGLU + \beta_{10} * SPBDU10T + \beta_{11} * MXEULC.vs..MCUSLC + \\ \beta_{12} * USDEUR + \beta_{13} * USDGBP + \beta_{14} * USDJPY + \beta_{15} * GB1M \end{aligned} \tag{31}$$

Macro strategy replication

The overall replication performance for this strategy is not very good, and both the final clones and the statistics analysis before and after are showing the same result: that this strategy is very difficult to replicate. This can be due to the chosen 25 factors or to many factors within the clone or other unknown factors that might not be possible to replicate. However, this study shows signs of that the Dow Jones US Growth and Value indices are present with short positions, and the S&P US treasury 10-year bond is most likely present with a long position. The main markets of the hedge fund index for macro are found to be the US, but the replication also leaves us with a huge share for markets which cannot be identified.

[60] Appendix 8.6

5.4.4 Multi-strategy - EHFI254

Before replication is the models for each period with multi-strategy at an acceptable adjusted R^2 level. For the pre-crisis period, the adjusted R^2 is given at 86.1%, for the crisis period the adjusted R^2 is 85.7% and for the post-crisis period the adjusted R^2 is 89.15%.

Pre-crisis period

The model for the pre-crisis period is found by linear regression, just as in the rest of the analysis, and it is given with the following equation:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \\ \beta_4 * EEM + \beta_5 * GI1 + \beta_6 * LUACTRUU + \beta_7 * MXWOLCG + \\ \beta_8 * SCRTEM + \beta_9 * MXWOLC.vs.MXWOSC + \beta_{10} * USDEUR + \\ \beta_{11} * USDJPY + \beta_{12} * GB1M \end{aligned} \tag{32}$$

The performance measurements, RMSE, tracking error and Theil's inequality coefficient are showing that performance in the continuous development of the clone compared to the hedge fund is acceptable. But, looking at the clone's ability to generate excess return relative to both the hedge fund and the risk-free rate, the clone does not perform. The clone is under-performing the hedge fund and the information ratio is therefore negative.

The Sharpe ratio is -1.011% for the hedge fund index and -24.592% for the clone. This clone is having the same trouble as earlier clone's from the pre-crisis period, where the financial crisis large decrease over short time are having huge effect. This effects that the clone are ending at index 102, where the hedge fund are ending at 118 for the out-sample period. All this is are shown in figure 12 below.

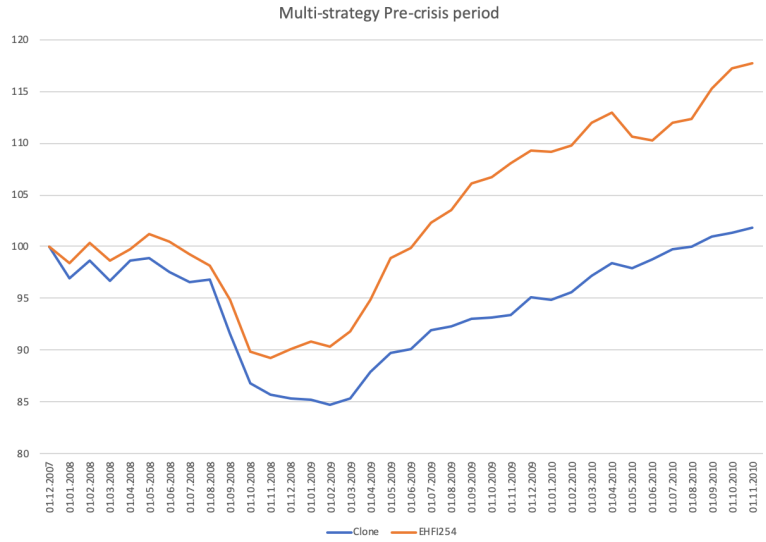


Figure 12: Multi-strategy pre-crisis period

Crisis period

The model for the backwards elimination regarding the multi-strategy hedge fund in the crisis period is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSVA + \beta_3 * EEM + \beta_4 * GI1 + \\ \beta_5 * LUACTRUU + \beta_6 * MXEF000G + \beta_7 * MXEU000G + \beta_8 * MXWOLC + \\ \beta_9 * SCRTEM + \beta_{10} * SPBDU10T + \beta_{11} * MXWOLC.vs.MXWOSC + \\ \beta_{12} * USDEUR + \beta_{13} * USDJPY + \beta_{14} * GB1M \end{aligned} \tag{33}$$

The number of factors are huge, which might be because of the strategy, which most likely contain multiple hedge fund strategies. The large numbers of factors could lead to a bad replication, but when looking at the results, the clone has a more than acceptable performance. The RMSE and the tracking error are both showing a good performance. Even though the forecasted returns are close and consistent with the hedge funds returns, it is still not performing as well as the hedge fund and is therefore under-performing. The clone does deliver a positive mean return and Sharpe ratio but negative information ratio. This is also seen graphically, where the distance between the clone and the hedge fund is increasing over the out-sample period.[61]

[61] Appendix 8.7

This result is showing possibilities of replicating the multi-strategy hedge fund, but still the result is not uniquely.

Post-crisis period

The post-crisis period has the same trouble as in the long short equity strategy, because the model contains the spread between Morgan Stanley EAFE index and Morgan Stanley US Large Cap index, which makes a huge decrease in the clone.
In the long short equity strategy, this was overcome by using the extended rolling window method. This could be done here again, for changing the outcome. We have chosen not to use the extended method, because of the low correlation (30%) in the first half of the out-sample period. See appendix 8.7

The model for this replication is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * DJUSVA + \beta_2 * EEM + \beta_3 * LUACTRUU + \\ \beta_4 * MXEA + \beta_5 * MXEF000G + \beta_6 * SBBMGLU + \\ \beta_7 * DJUSGR.vs.DJUSVA + \beta_8 * MXUS000G.vs..MXUS000V + \\ \beta_9 * MXEA.vs..MXUSLC + \beta_{10} * MXWOLC.vs.MXWOSC + \\ \beta_{11} * USDGBP + \beta_{12} * USDJPY + \beta_{13} * GB1M \end{aligned} \tag{34}$$

Multi-strategy replication

This strategy shows opportunities for replication in the crisis period, but is, given the chosen factor and limit for what to include in the model for this study, not performing in the pre-crisis period and the post-crisis period. In general, multi-strategy replication includes too many factors in the final models, which can be due the fact that multi-strategy works across all strategies to combine the most profitable investment. This is also shown when investigating the results of the replications; we see that the clones contain different strategies with a very equal weight, and the different markets are also represented one way or the other. The only factor which seems to be within the index is Morgan Stanley World Large Cap, but this is only in the pre-crisis and crisis replication periods; in the post-crisis period, it seems to be dropped

by Eurekahedge Hedge Fund Managers for the S&P 500 Global Broad Market index instead. These two factors are a little alike, with 23 and 25 developed markets within the indices; this change of factors could also be because of an estimation error or similar.

5.4.5 Relative value - EHFI255

The models for the pre-crisis period and the crisis period have an acceptable adjusted R^2 at respectively 77.2% and 78.87% and are given by:

Pre-crisis period

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * GI1 + \beta_3 * LUACTRUU + \beta_4 * MXWOLC + \beta_5 * MXWOLC.vs.MXWOSC + \beta_6 * USDEUR + \beta_7 * GB1M \quad (35)$$

Crisis period

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \beta_4 * LUACTRUU + \beta_5 * MXEU000G + \beta_6 * SCRTEM + \beta_7 * MXWOLC.vs.MXWOSC + \beta_8 * USDEUR + \beta_9 * GB1M \quad (36)$$

Pre-crisis period

The replication for the pre-crisis period is giving some acceptable results. The standard deviation is almost equal, and with a RMSE and Tracking Error less than 1 the clone is performing well on those parameters.

The correlation is at 87%, but like earlier hedge fund strategies, the clone is performing very well the first 18 months of the out-sample period, but in the last 18 months the distance between the clone and the hedge fund is increasing.[62] The mean return for the hedge fund is 0.697% and only at 0.23% for the clone; the Sharpe ratio is also negative for the clone compared to a positive value for the hedge fund, meaning that the clone is not really performing relative to the risk taken.

[62] Appendix 8.8

This strategy within the pre-crisis period also has the most negative information ratio, meaning that this clone is generating the least return compared to the respective hedge fund index.

Crisis period

The out-sample for the crisis period gives a better clone result than the pre-crisis period. The mean return and the share ratio for the clone and hedge fund are more alike and with a correlation index of 98.635%, the clone and the hedge fund have a very similar development, which is also shown in acceptable performance measurements, which refer to development.

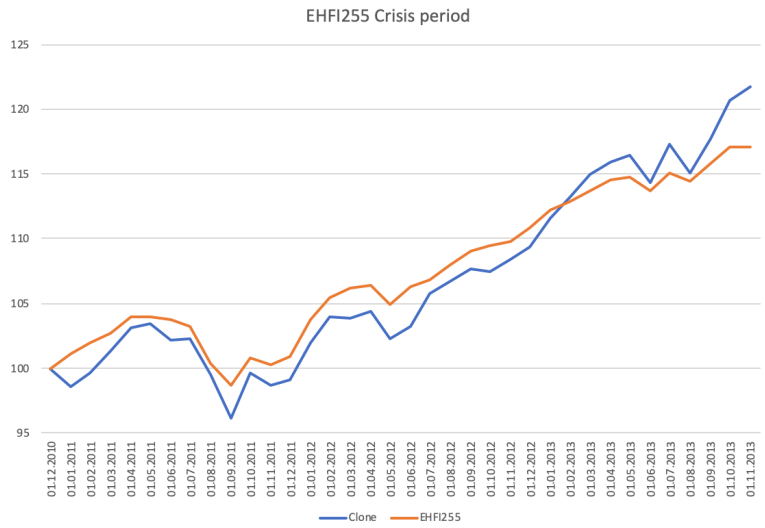


Figure 13: Relative value strategy the crisis period

As shown in figure 13, the clone and the hedge fund have very similar development through the out-sample period. But, unlike earlier replications, the clone is over-performing the hedge fund. The reason for the over-performance is given within the long/short positions.

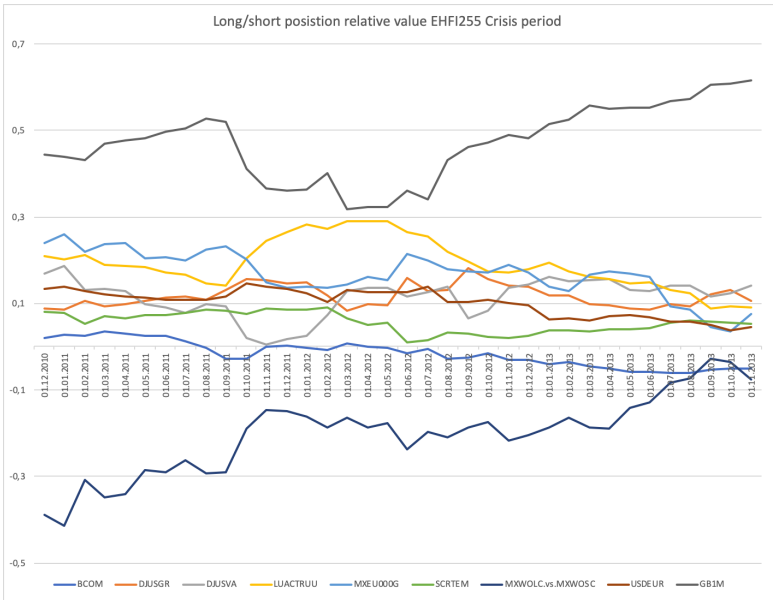


Figure 14: Long/short position - relative value strategy the crisis period

In figure 14 all the long or short positions for the different factors for the entire out-sample period are shown, and as we see the long position on the Barclays US government bond is decreasing in the second half of the period, and instead the spread between Morgan Stanley World Large Cap and Small Cap are decreasing the short position and the risk-free asset are increasing the long position.

Post-crisis period

For the post-crisis period the replication for the relative value hedge fund index is for the past 36 months from August 2019 and back. The model for the replication has an acceptable adjusted R^2 on 68.84%, and the function for the model is given with:

$$\begin{aligned} R_t^{Clone} = \beta_1 * DHY + \beta_2 * LUACTRUU + \beta_3 * MXWOLC + \\ \beta_4 * MXWOSC + \beta_5 * SPBDU10T + \\ \beta_6 * MXWOLC.vs.MXWOSC + \beta_7 * GB1M \end{aligned} \tag{37}$$

Looking at the performance measurements of the clone versus the hedge fund, the correlation is good at almost 89%, and the standard deviation are almost equal to each other. Despite some good measurements, the Sharpe ratio is on -46% for the clone but -122% for the hedge fund. This is because of the risk-free rate, which is increasing a lot in the out-sample period.

The risk-free rate development in the out-sample period has huge effect on the clone, and with a mean return of 1.042% instead of the 0.29% for the hedge fund; the clone is ending the out-sample period on index 145 in opposite to the hedge fund, which ended on 110, meaning that the clone is over-performing the hedge fund in the post-crisis period.

To overcome this and hopefully to replicate a better clone, we are changing the length of the window from 36 to 12. In doing this, we hope to establish the volatility in the hedge fund, but instead the model got more significant with the risk-free rate. This increased the mean to 1.172% and the Sharpe ratio increased to -25.821%. The result for the short window did not establish a better result or a better clone.

We still want to overcome the risk-free rate significantly and deliver a better replication for the hedge fund. To do this we changed the length of the in-sample period to see if it would change the model through the backwards elimination.
The in-sample period was changed from function (37) to:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * MXEU000V + \beta_3 * SBBMGLU + \\ \beta_4 * MXWOLC.vs.MXWOSC + \beta_5 * USDGBP + \\ \beta_6 * USDJPY + \beta_7 * GB1M \end{aligned} \tag{38}$$

The statistics for the third model are showing improvement in the clone. The impact from the risk-free asset is still making the return higher for the clone than the hedge fund, but the new clone has a better fit. The Sharpe ratio is decreasing to -46% from the original clone to -59.9%. Furthermore, the correlation is going from almost 89% to almost 95%. The RMSE, tracking error, information ratio and Theil's Inequality Coefficient are all improved compared to the original clone.

The result for the clone with the short in-sample period shows that changing the estimation period has a huge impact on the final clone; it is not only the coefficient estimations. Therefore, it is important the have the "right" factors for achieving a good fit for the replication clone.

Relative value strategy replication

This strategy also shows opportunities for replication in the crisis period, but is more challenging in the pre-crisis period and the post-crisis period, because of the financial crises in the pre-crisis period out-sample and the increase in the risk-free rate in the post-crisis period. The change with risk-free rate is tried to be overcome by changing different properties within the rolling window method. It is overcome to some extent but not enough to call the replication in the post-crisis period for a good clone.

In the Relative Value replications, it is seen that there is a good chance that Barclays US Corporate Bond Index is present with long positions in the hedge fund index, or perhaps a similar corporate bond index. Otherwise, it is not possible to conclude any market from our replication clones because of changing in every period, but the asset classes are most likely equities at around 40% of the total index.

5.4.6 Arbitrage - EHFI285

For the replication of the arbitrage strategy, the result was acceptable before the replication. The adjusted R^2 for the three periods are between 74.2% and 76.71%, and all three models are not showing any sign of non-normality in the residuals. The test for homoscedasticy cannot be rejected for the models regarding the pre-crisis period and the post-crisis period, but the null hypothesis can be rejected for the crisis period, meaning that OLS estimate is not biased, which might not then give the best results for the model.

The model for the three periods is given by the following equations:

Pre-crisis period

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * GI1 + \beta_3 * LUACTRUU + \\ \beta_4 * MXEU000G + \beta_5 * MXWOLC + \beta_6 * SPBDU10T + \\ \beta_7 * MXEULC.vs..MXUSLC + \beta_8 * USDEUR + \beta_9 * GB1M \end{aligned} \tag{39}$$

Crisis period

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DHY + \beta_3 * LUACTRUU + \beta_4 * MXEU000G + \beta_5 * USDEUR + \beta_6 * USDJPY + \beta_7 * GB1M \quad (40)$$

Pre-crisis period

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * EEM + \beta_3 * GI1 + \beta_4 * MXWOLC + \beta_5 * SBBMGLU + \beta_6 * SPBDU10T + \beta_7 * GB1M \quad (41)$$

Performance pre-crisis, crisis and post-crisis period

The replications for the pre-crisis period are performing on an acceptable level. The performance measurements are all showing an acceptable fit for the clone.[63] The distance between mean and standard deviation are not very big, and the RMSE, tracking error, information ratio and Theil's Inequality coefficient are at an acceptable level. But the clone is not able to generate excess return compared to the risk taken, which the hedge fund index is. When comparing the clone and the hedge fund graphically, the two indices have very similar development through the out-sample period; see figure 15.

The crisis period clone is performing at a good level. All the statistics are performing despite the Sharpe ratio, where the difference is 30%-point, where the clone is at 30.172%, so still a good value, but nowhere near as good as the hedge fund. The RMSE and the Tracking Error are some of the best performing clones within the crisis period, with values of 0.00523 and 0.468% respectively. The information ratio for this clone is more negative when looking only at the crisis period, despite the relative good performance. This is mainly because of the extremely good performance the hedge fund has in this period and the fact that the clone does not perform in the last 12 months of the crisis period.

For the post-crisis period, the performance is very similar to the pre-crisis period and the crisis period, only influenced by the risk-free rate that is coursing the mean return to be 3

[63] See appendix 8.3

times as high for the clone versus the hedge fund. But despite this, the clone has the same pattern, which is also seen in the correlation of almost 94%.

Graphically, the three periods are showing the same pattern as the performance measurements. The clone and the corresponding hedge fund are seen in figure 15.

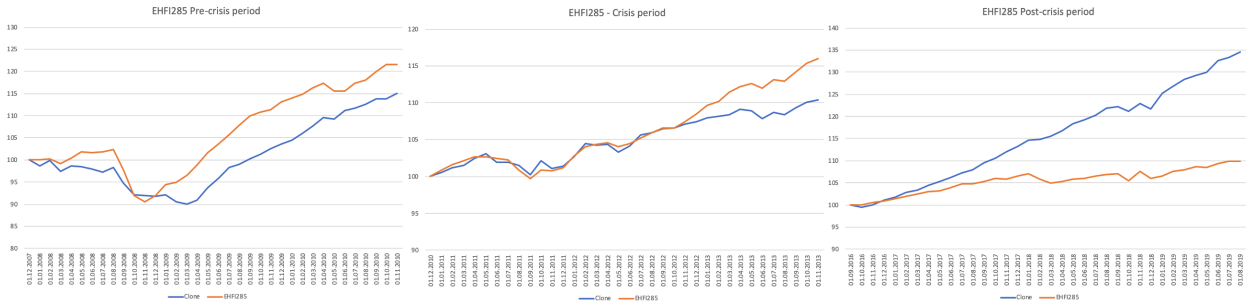


Figure 15: Arbitrage RW clone

The correlations are at a high level, all over 90%, but we want to see if we can make the replication even better by getting more similar returns. Therefore, we are using the rolling window extended on all three periods.

Performance pre-crisis, crisis and post-crisis period - Extended RW

As in the strategy for the long short equity, we are again trying to optimize the replication clone and the results by using the extended rolling window method. When repeatedly making the linear regression and backwards elimination for estimating each coefficients for each period, the results are changing.

The results for the two different methods are compared in figure 16

	Pre-crisis period			Crisis period			Post-crisis period		
	EHFI285	RW clone	Ex RW clone	EHFI285	RW clone	Ex RW clone	EHFI285	RW clone	Ex RW clone
Mean	0,565%	0,408%	0,370%	0,442%	0,269%	0,300%	0,276%	0,837%	0,658%
Standard Deviation	1,766%	1,463%	1,458%	0,654%	0,751%	0,823%	0,642%	0,704%	0,852%
Sharpe Ratio	3,190%	-6,883%	-9,520%	61,163%	30,172%	31,340%	-180,883%	-85,272%	-91,453%
Correlation index		94,560%	97,495%		97,061%	92,948%		93,958%	95,664%
RMSE		1,238%	0,997%		0,523%	0,639%		0,787%	0,756%
Tracking Error		1,262%	0,991%		0,468%	0,630%		0,548%	0,659%
Information ratio		-12,438%	-19,685%		-37,095%	-22,563%		102,417%	57,973%
Theils Inequality Coefficient		9,098%	7,326%		5,887%	7,185%		9,388%	9,029%

Figure 16: Results RW vs. Extended RW

For the pre-crisis period, the clone has improved the correlation and the performance measurements for the clone comparison are improving. This dot not imply to the mean, standard deviation and Sharpe ratio, which are not improving compared to the first clone.

For the crisis period, the extended rolling window is making the clone over-perform in the middle of the out-sample period, which is leading to worse replication. The decrease in the performance measurements are small, but significant.

The post-crisis period is performing at a high level with the rolling window method, but after analyzing with the extended rolling window method, the clone is performing at an even higher level. The clone is still over-performing the hedge fund, but the clone is approaching the hedge fund compared to the rolling window clone. The rolling window clone is ending at a level of 135, and the extended rolling window clone are ending at 126 index. Compared the hedge fund are ending at 110 index for the out-sample period.

Arbitrage strategy replication

Within the arbitrage strategy there are opportunities for replications. The clones does still suffer for not delivering as good a return, when they are not too heavily influenced by the risk-free rate. This can be overcome to some extent by monthly change the factors and there associated sensitivities, to act more like the hedge fund managers, but this method does not perform in the pre-crisis period and the crisis period, and are therefore not overall seen as a good solution to this strategy.

For this strategy the replications are not showing any direct signs that we have found factors within the hedge fund index, and when looking at the asset classes we see that both equity, fixed income, cash & cash equivalents and commodities are present - although, the markets seem most likely to be mainly emerging markets, which also are seen in historical findings. For instance Kooli, M. & Sharma, S. (2012) also find that Arbitrage strategy contains mainly emerging markets, which they discuss in the article *"Should we give hedge funds clones a chance?"*

5.4.7 Distressed debt - EHFI287

Pre-crisis period

Compared to many of the other strategies, the distressed debt clones are not as impressionable to the financial crisis decrease in returns as the hedge fund for distressed debt. The clones are therefore over-performing the hedge fund for almost the entire out-sample period, but end just below the hedge fund, and the information ratio is therefore relative small negative.

The model for the clone is given by:[64]

$$R_t^{Clone} = \beta_1 * LUACTRUU + \beta_2 * MXEF000G + \beta_3 * MXWOSC + \\ \beta_4 * SCRTEM + \beta_5 * SPBDU1OT + \beta_6 * USDEUR + \beta_7 * GB1M \tag{42}$$

7 factors are seen as a very reasonable numbers of factors to include in the clone and is therefore setting a good building block for the out-sample period.
The adjusted R^2 for the model is 72.71%, which is acceptable. This is not as good as some of other replication models, and some of the performance measurements are also not showing a good fit for the clone. Although the correlation are on 93.57%, both the RMSE and the Tracking Error are on a bad level, due to the gap between the clone and the hedge fund index and the difference in consistency over the financial crises and in the end of the period.
Looking at the performance measurements, which refers to performance on return, the clone is performing very close to the hedge fund.

[64] Appendix 8.19, Figure 37

Crisis period

The crisis period has an adjusted R^2 on 70.35% and a correlation of 95.75%.

The model for the crisis period is given by:

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * LUACTRUU + \beta_3 * MXEA + \beta_4 * MXEF000G + \beta_5 * MXEU000G + \beta_6 * MXWOSC + \beta_7 * SCRTEM + \beta_8 * USDEUR + \beta_9 * USDGBP + \beta_{10} * GB1M \quad (43)$$

Graphically, the clone seems to have a great fit,[65] but the difference in the pattern is shown in the performance measurements.[66] The clone and the hedge fund are very even, but the returns are opposing each other, which is seen in the performance measurements.

Post-crisis period

The post-crisis period model is given by:

$$R_t^{Clone} = \beta_1 * LUACTRUU + \beta_2 * MXWOLC + \beta_3 * MXWOSC + \beta_4 * SPBDU10T + \beta_5 * MXWOLC.vs.MXWOSC + \beta_6 * GB1M \quad (44)$$

This model has an adjusted R^2 on 58.28%, which is not not acceptable. Furthermore, the performance measurements for the clone are not acceptable. The not acceptable fit is again due to the increase in the risk-free rate, and the clone is therefore over-performing the hedge fund, and delivering way better performance on the return than the hedge fund, but is not a good fit.

Distressed debt strategy replication

Distressed debt strategy has shown difficulties when trying to replicate the hedge fund index. The pre-crisis period delivers good performance on returns, but is not a good fit to analyze transparency of the hedge fund index. The crisis period and the post-crisis period are not performing, which gives the signal that this strategy might be difficult to replicate as the crisis period has been performing acceptably on most of the other strategies.

[65] Appendix 8.10

[66] Appendix 8.3

In order to understand the index, the clones are showing signs that the Barclays US Corporate Bond Index and the Morgan Stanley World Small Cap are both present with long positions. Furthermore, the index is constructed with mainly Equities and Fixed income and with emerging market and developed markets as the primary markets. This provides us with a solid foundation for understanding the index, but as mentioned earlier, the clones are not performing at a high level in replicating the hedge fund index, which is why the results also are less credible.

5.4.8 Event driven - EHFI288

The analysis for the event driven strategy is, as in the rest of the analysis, made by linear regression and backwards elimination. The three models for the analysis are found via the statistical software RStudio, as in the rest of the analysis. RStudio is also used for estimating the coefficients for each factor in each period.
For the event driven strategy, the models are found to be:

Pre-crisis period

$$\begin{aligned} R_t^{Clone} = \beta_1 * LUACTRUU + \beta_2 * MXEF000G + \beta_3 * MXEU000G + \\ \beta_4 * MXWOLC + \beta_5 * SCRTEM + \beta_6 * SPBDU10T + \\ \beta_7 * MXEULC.vs..MXUSLC + \beta_8 * MXWOLC.vs.MXWOSC + \\ \beta_9 * USDEUR + \beta_{10} * USDJPY + \beta_{11} * GB1M \end{aligned} \tag{45}$$

Crisis period

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * LUACTRUU + \beta_3 * MXEF000G + \\ \beta_4 * MXEU000G + \beta_5 * MXWOLC + \beta_6 * SCRTEM + \beta_7 * SPBDU10T + \\ \beta_8 * MXWOLC.vs.MXWOSC + \beta_9 * USDEUR + \\ \beta_{10} * USDJPY + \beta_{11} * GB1M \end{aligned} \tag{46}$$

Post-crisis period

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSVA + \beta_3 * EEM + \beta_4 * GI1 + \\ \beta_5 * LUACTRUU + \beta_6 * MXEA + \beta_7 * MXWOLC + \beta_8 * MXWOSC + \\ \beta_9 * SBBMGLU + \beta_{10} * MXUS000G.vs..MXUS000V + \\ \beta_{11} * MXWOLC.vs.MXWOSC + \beta_{12} * GB1M \end{aligned} \tag{47}$$

Pre-crisis period

The adjusted R^2 for the clone is on 87.88%, and the results for the clone are very similar to other pre-crisis period clones. The first half of the out-sample period the clone has a good fit to the hedge fund, but after the financial crisis the clone and the hedge fund are getting more and more unequal; see appendix 8.11. Despite the correlation on 92% and acceptable standard deviation, RMSE and tracking error, are the performance measurements concerning ability to generate return near the hedge fund, because of the development after the financial crises. The normality test for the residuals are showing sign on outliers, which can have an effect on the model in a negative way. The QQ-plot for the normality is shown in appendix 8.17.

Crisis period

For the crisis period the model is performing at an acceptable level. Before the rolling window, the adjusted R^2 is on 89.53%. The final clone as a correlation on 98.685% with the hedge fund index and the return performance measurements are better in relation to the hedge fund than for the pre-crisis period. The clone is not generating excess return relative to the hedge fund at any point in the out-sample and is not performing as well as the hedge fund compared to the risk taken. But overall, as seen before, the crisis period is performing better than the pre-crisis period, but we are aware that the fit possible can be better because we can reject the null hypothesis for homoscedasticy.

In figure 17 the clone and the hedge fund index are graphically shown over the out-sample period. The high correlation is easily shown along with the clones lack of ability to create excess return.

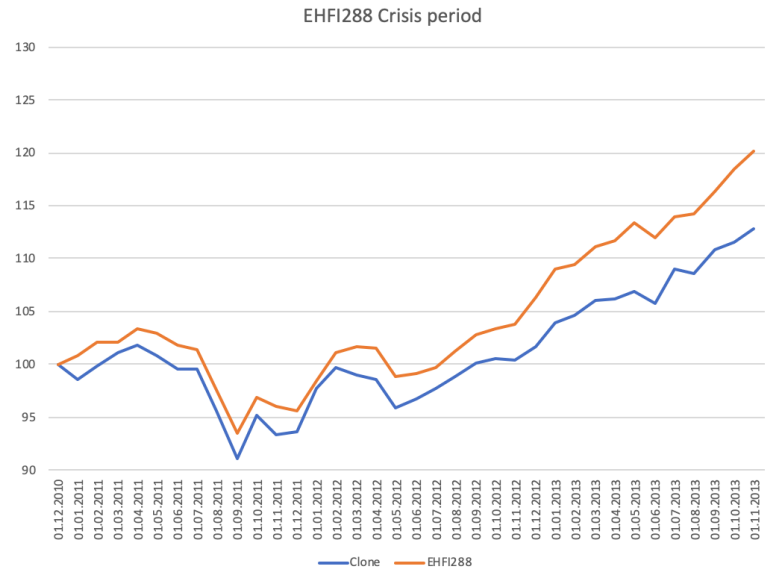


Figure 17: Event driven clone - crisis period

Post-crisis period

The clone in the post-crisis period is, compared to the other strategy clones in the post-crisis period, performing relatively well. But, just looking graphically, the performance of the clone is not at the same level as the hedge fund; especially until November 2018, the clone is having trouble replicating the hedge fund.[67] But looking just at the performance measurements, most of them indicate a good replication, but the clone is missing out when comparing the Sharpe ratio.

Compared to other clones in out-sample for the post-crisis period, the event driven strategy is not as influenced as the others to the risk-free asset. Through the analysis we have seen a lot of post-crisis clones, which has a very high mean return over the period, because of the risk-free asset, but in this strategy the clone are performing on a good level compared to the hedge fund index we are trying to replicate.

Event driven strategy replication

Event driven strategy replication is showing some of the same patterns as this analysis has shown previously. The financial crises in the out-sample period has a negative effect on the

[67] Appendix 8.11

clone and the replication is not over-performing the hedge fund. The crisis period is similar to most other clones, and could indicate possibilities for replication. The post-crisis period is also showing replication possibilities when looking at the performance measurements, and is not as influenced by the huge increase in the risk-free rate as other strategies, which can indicate good replication possibilities within this strategy.

The strategy is also showing a lot of interesting knowledge for the hedge fund index. In all 3 periods, the replications are showing that the index possibly contains long positions in both Barclays US Corporate Bond Index and Morgan Stanley World Large Cap. Furthermore, the clones are showing signs that the share of the index is about 30% equities. It could be more, because of unknown factors. Finally, we see that the index mainly contains developed and emerging markets, also shown in broad markets. All this information, gives us a knowledge that could possibly lead to an even better replication, than already established. For instance, by extending the model with factors within the asset class of equities and within the emerging or developed markets, we could find factors with greater potential to be within the index, and thereby gain even bigger knowledge of the index.

5.4.9 Fixed income - EHFI289

Pre-crisis period

For the fixed income strategy, the pre-crisis period model is given by 11 factors, and corresponding long/short positions to establish the clone. The model is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSVA + \beta_3 * GI1 + \beta_4 * LUACTRUU + \\ \beta_5 * MXEF000G + \beta_6 * MXEU000G + \beta_7 * SCRTEM + \beta_8 * SPBDU10T + \\ \beta_9 * MXEULC.vs..MXUSLC + \beta_{10} * USDEUR + \beta_{11} * GB1M \end{aligned} \tag{48}$$

Despite a low adjusted R^2 of 74.74%, and relative many factor in the model is the fixed income clone performing better compared to most other strategies in the out-sample for the pre-crisis period, which is including the financial crises.[68] This is coursing most other strategies troubles.

[68] Appendix 8.12

The correlation is on 94% and the ability to generate overall similar return to hedge fund is acceptable. The clone is, however, still not able to deliver compared to the risk taken, which the hedge fund is doing in the pre-crisis period.

Crisis period

For the crisis period, the adjusted R^2 is on 75.52%, which is acceptable.

The model for the crisis period is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * LUACTRUU + \beta_3 * MXEF000G + \\ \beta_4 * MXEU000G + \beta_5 * MXWOLC + \beta_6 * SCRTEM + \\ \beta_7 * MXWOLC.vs.MXWOSC + \beta_8 * USDEUR + \beta_9 * GB1M \end{aligned} \tag{49}$$

The performance for the model is good, with a correlation on 96.512% and RMSE on 0.00635 and tracking error around 0.6%.[69] But when looking at the pattern, it is clear that the high correlation is because of the high relationship between the development in the returns, but the small increase in returns in the hedge fund are constant over the period, which gives a Sharpe ratio that is 61% compared to 29% and a mean return for the hedge fund on 0.6% compared to 0.372%, and here it is clear that the replications are not delivering the wished return, which perhaps could be because of homoscedasticy that the fit are not the possible best fit.

Post-crisis period

The model for the post-crisis period has an adjusted R^2 on 82.1%, and the model is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * GI1 + \beta_3 * LUACTRUU + \\ \beta_4 * MXWOLC + \beta_5 * MXWOSC + \beta_6 * MXWOLC.vs.MXWOSC + \\ \beta_7 * USDJPY + \beta_8 * GB1M \end{aligned} \tag{50}$$

This model is highly influenced by the development in the risk-free rate, which makes the clone out-perform the hedge fund index in the out-sample period, which is both seen in appendix 8.12 and the performance measurements. Even though the clone over-perform the hedge fund

[69] Appendix 8.12

by a lot, the replication is not a good match. As earlier, this could be change if we changed the method to extended rolling window but this is not tried for this strategy, as we will expected the same outcome as previously.

Event driven strategy replication

Overall, this strategy replication is performing at an acceptable level, but it does not give us any result for a final conclusion whether replication of the hedge fund is possible or not, as this strategy is struggling in some areas such as generating a similar return, not under-performing or over-performing by to much, which is the case for this strategy.

Like under the event driven strategy, our replications are showing signs that the Barclays US Corporate Bond Index is within the hedge fund index with a long position. The long position is around 30% of the portfolio, and this factor alone is affecting the different asset class, such that the fixed income is highly represented. Furthermore the equity and cash & cash equivalents are represented within the index, but on a smaller level than the fixed income.

5.4.10 Equity market neutral - EHFI751

The three models for the equity market neutral strategy have a lower goodness of fit. The adjusted R^2 for the three periods are respectively given by 60.71%, 57.07% and 51.82%, which is not acceptable. Furthermore the statistics test not all delivering to the assumptions under OLS, which is also not giving the best replication requirements. So before rolling out the window over the out-sample periods is the models not set to perform, and other more suitable models can be found, than the linear regressions with the provides factors for this study, but this is just to keep in mind when concluding on the ability to replicate this strategy, and we will proceed with those models.

The model for the three periods is given by:

Pre-crisis period

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * DJUSVA + \beta_4 * EEM + \beta_5 * LUACTRUU + \beta_6 * MXEF000G + \beta_7 * MXEU000G + \beta_8 * SBBMGLU + \beta_9 * USDEUR + \beta_{10} * GB1M \quad (51)$$

Crisis period

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSVA + \beta_3 * EEM + \beta_4 * MXEF000G + \beta_5 * MXEU000G + \beta_6 * SCRTEM + \beta_7 * SPBDU10T + \beta_8 * USDEUR + \beta_9 * GB1M \quad (52)$$

Post-crisis period

$$R_t^{Clone} = \beta_1 * DJUSGR + \beta_2 * DJUSVA + \beta_3 * MXEA + \beta_4 * MXWOLC + \beta_5 * DJUSGR.vs.DJUSVA + \beta_6 * GB1M \quad (53)$$

Performance

In general, the performance for the equity market neutral can be divided in two. In the pre- and post-crisis periods, the clone is not performing at a high level. The correlations are respectively 75% and 73%, and the difference in the mean return compared to the hedge fund is quite large.[70] The RMSE and the tracking error are actually on a good level at around 0.01 and 0.6%, respectively, but looking at the other statistics, the clones are not performing at a high level, but as mentioned earlier, the model analyses did show that the goodness of the fit was not very high.

The crisis period, on the other hand, is performing acceptably. The clone are constantly under-performing compared to the hedge fund, but looking at the statistics the clone is at a good level. The correlation is 93% and with mean returns on the same level and a RMSE

[70] Appendix 8.13

under 0.01 and tracking error under 1%, is it only the Sharpe ratio and standard deviation which shows lack of performance in the clone compared to the hedge fund.[71] The fit could possibly be better, but because of homoscedasticy we cannot determine if it is the best possible fit.

Equity market neutral strategy replication

Overall, the equity market neutral strategy is not very easy to replicate using the given factors and periods, but perhaps another method or other factors would change this fact, as historical findings has shown the ability to replicate and over-perform this strategy.
Like the lack of performance in the clone, the possibility of getting any knowledge of the hedge fund index is very small. We are not able to subtract any knowledge of which factors could be present within the hedge fund, or which asset class or markets are most likely to be within the index and their weights. This could be because of the lack of performance in the replication or because of wrong factors chosen for this analysis.

5.4.11 Equity short bias - EHFI807

The last strategy we are replicating is the equity short bias strategy. The index does not contain multiple funds, and it is the only Eurekahedge index which has a negative performance since the establishment in December 99. Currently the index is 81, and especially since October 2012, the index has decreased (from index 152). This is solving some of the earlier mentioned problems - among others, the risk-free asset problem in the post-crisis period.

Pre-crisis period

In the pre-crisis period, the model has an adjusted R^2 on 64.48%, which is not acceptable, and the number of factors are at the limit of how many is good for replication.

The model for the pre-crisis period is given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * EEM + \beta_3 * LUACTRUU + \beta_4 * MXEA + \\ \beta_5 * MXEF000G + \beta_6 * MXEU000G + \beta_7 * MXWOLC + \\ \beta_8 * MXWOLC.vs.MXWOSC + \beta_9 * USDEUR + \beta_{10} * GB1M \end{aligned} \tag{54}$$

[71] Appendix 8.13

The performance for this clone is bad. The correlation is on 70% and the tracking error is over 3.3% and RMSE are both over 0.03, meaning that the clone is not as close to the actual values as wished; furthermore, the Sharpe ratio is very different from the hedge fund. Looking at the index and the corresponding returns graphically, we see many observations where the hedge fund is having a negative return and the clone are having a positive return. This is shown in figure 18; this is leading to the over-performance most of the time in the pre-crisis period.

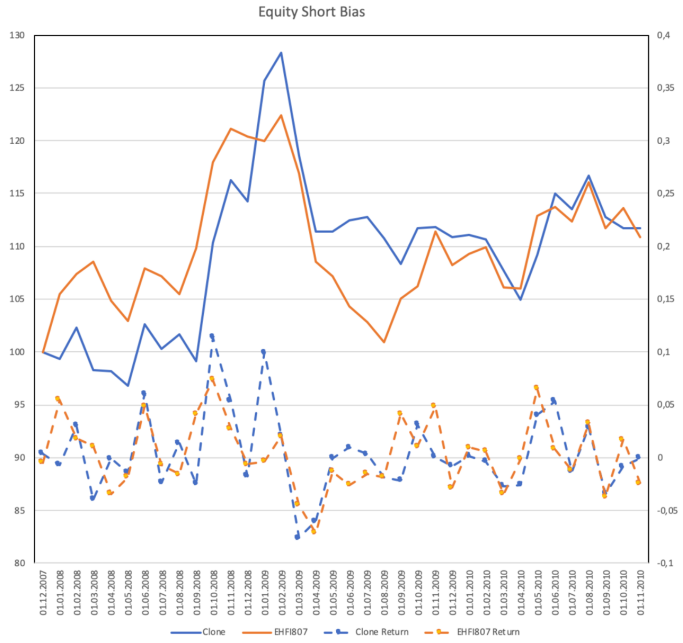


Figure 18: Equity short bias pre-crisis period

Crisis period

For the crisis period the model has an adjusted R^2 on 63.43%. This is, like in the pre-crisis period, not an acceptable goodness of fit, and it is also shown in the results. The model for the crisis period is given by:

$$R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DJUSGR + \beta_3 * MXEU000G + \beta_4 * SBBMGLU + \beta_5 * SPBDU10T + \beta_6 * MXWOLC.vs.MXWOSC + \beta_7 * USDEUR + \beta_8 * USDGBP + \beta_9 * GB1M \quad (55)$$

The result are not replication the strategy well, the clone are performing on a higher level than the hedge fund, but are not close to the strategy at all. The correlation is down at 11.885%, and with a mean return for the hedge fund at -0.623% compared to 0.323% for the clone, this is clearly a bad replication of the Equity short bias strategy and is not creating any transparency with this strategy. When looking at the entire in-sample period and the out-sample period for the hedge fund index and comparing with the clone, the clone is following the pattern which is a steady increase. This is shown in figure 19.

The huge decrease in the hedge fund at the end of the out-sample period is not possible

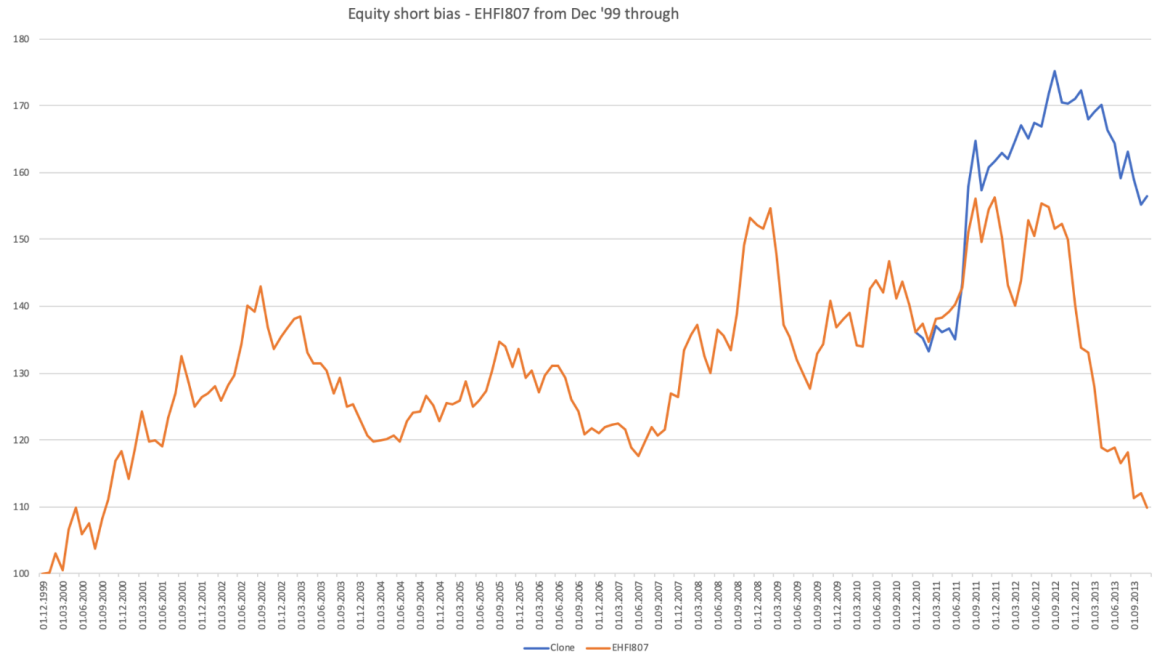


Figure 19: Equity short bias from '99 to 2013

for the clone to replicate, and therefore the statistics are showing bad results. As this huge drop has not been seen before in this strategy, this will therefore not be captured in the model generated over the in-sample period, and will therefore have difficulty to be created, based on the factors, if none of the factors does not either have this drop in the out-sample. Which is not the case, and the factor for the clone are therefore not shown to be factors included in this strategy.

Post-crisis period

The post-crisis period for the equity short bias, the model are given by:

$$\begin{aligned} R_t^{Clone} = \beta_1 * BCOM + \beta_2 * DHY + \beta_3 * GI1 + \beta_4 * MXEA + \\ \beta_5 * MXEF000G + \beta_6 * MXEU000G + \beta_7 * MXEU000V + \beta_8 * MXWOSC + \\ \beta_9 * SBBMGLU + \beta_{10} * SPBDU10T + \beta_{11} * MXEA.vs..MXUSLC + \\ + \beta_{12} * MXWOLC.vs.MXWOSC + \beta_{13} * USDJPY + \beta_{14} * GB1M \end{aligned} \tag{56}$$

The model are performing on a medium level, with an adjusted R^2 on 83.15%.

Looking at the clone result, the mean return, standard deviation and Sharpe ratio are implying a good fit, but looking at the correlation on 50.9% and a RMSE around 0.04 and tracking error on approximately 3.8%, the clone shows a bad fit to the hedge fund. The whole period the clone is under-performing compared to the hedge fund, even though the risk-free rate is increasing a lot in the same period, but because of the negative return from the hedge fund, short positioning in the risk-free asset is making the clone under-perform.[72] This does not have the same high influence as in the other strategies, mainly because of the negative return.

Equity short bias strategy replication

This strategy does not imply possibility for replication; this could due to the fact that none of the chosen factors does develop negatively as this strategy, and the chosen method does not capture the negative development and set the associated sensitivities to match the short selling strategy. We see a pattern from the replication that the index contains mainly emerging markets and equities short selling but because of our difficulties in replicating this short selling strategy, the conclusion for the markets and asset class should be taken lightly. Perhaps if the study had more factors with a negative return over the out-sample period, it could give us some other answers or confirm the light conclusion.

[72] Appendix 8.14

6 Discussion

In the previous section, the analysis, the results for the different strategies are shown for each period, and the general findings within the strategy are explained under the given strategy.
To get a deeper understanding of the findings in this study, this section will look at the results within the three different periods and the general overall findings done through hedge fund replication by the use of two rolling window methods, Rolling Window and Extended Rolling Window.
This section will also look at this study's results and how the replications are performing compared to historical findings.

Overall, the replications from this study struggle to over-perform the hedge fund indices when the clones are not influenced by the massive positive development in the risk-free rate.
Especially in the pre-crisis period, the clones have difficulties after the financial crisis in 2008. The clones are not able to quickly generate the same level of return as the hedge funds after the financial crises. This is due to the drop under the crises that are effecting the estimations afterwards. The hedge funds are quicker to recover, the clones are first at a later state obtaining this increase, and are therefore struggling with the mean return over the entire period, compared to the hedge fund indices.
This implies weakness in the hedge fund replication. As it could indicate that the hedge fund managers might be able to recover quickly after a financial drop in the market due to their skills/experience (alpha). But this missing ability to replicate the recovery can also be due to other unknown impacts, such as missed opportunities when selecting factors.
This study choose the factors based on historical findings and limited knowledge about the strategies. Even though this study is, compared to previous research, working with relative many factors could there most likely be some missed opportunities when selection factors for 11 hedge fund strategies.
But even though hedge funds might seem attractive just after the crises, they also suffer under the financial crisis and due to their lack of liquidity, some investors did therefore prefer hedge fund clones as they provided them with more liquidity, lower fee and more transparency.[73] To

[73] Jawadi, F & Khanniche, S. (2011)

make hedge fund replication clones attractive product after a financial crisis, it could be a possibility to either change the estimation periods or the replication method, which maybe could lead to a quicker turnaround, and thereby a greater return. By changing the estimation periods to contain more observations, a drop could have less influence on the forecast. This could either be done by extending the window size under rolling window or by using fixed weight replication, which only have a long in-sample period. But by doing so, perhaps some investors start to doubt the credibility of the clone because of too old data do not represent the present, most likely. However, it is also important to keep a decent size on the estimation period as too short periods will be influenced a lot by small and huge changes in the market, because of few observations. Furthermore, this could lead to unreal significant levels, which could create a bad fit compared to a longer period replication.
It could also be a possibility to use a method that takes into account the adjustment in the investment portfolio over time, such as the described Kalman filter method by Kooli, M. & Sharma, S. (2012) or the extended rolling window in this study. Those methods do better reflect the manager's work over time, and might quickly reflect the changes during drops and increases as well, as both the factors and their associated sensitivities change over time.

When looking at which strategies in the pre-crisis period that is showing signs of possibility for replication, the strategy *distressed debt* is one of the better clones when focusing on the ability to replicate similar return. See appendix 8.10 for graph development and appendix 8.3 for performance measurements. The ability to generate a similar return over the total out-sample period is in this case due to the minor drop during the financial crises. The clone is still struggling to contain high returns on the same level as the hedge fund after the crises, but overall the mean return is showing signs of a good fit. The minor drop is, on the other hand, not generating the wished transparency within the strategy, as the development is not following closely the hedge fund index.
Another strategy that in this study has shown possibilities for replication is *arbitrage*, as this strategy better maintain the gap between the clone and the hedge fund. But also here is the clone not able to quickly recover after the crises and does therefore over the whole pre-crisis period deliver a lower mean return and Sharpe ratio. But this strategy is showing a better performing clone compared to the other strategies because they are not recovering as quickly

and thereby maintaining the gap to the hedge fund.
Where this clone and most other hedge fund clones in the pre-crisis period are also struggling, is to generate excess return relative to the risk-free rate. Most hedge fund indices are generating excess return compared to the risk-free rate. But when analyzing the clone's Sharpe ratio, the negative values are not appealing for choosing hedge fund replication over hedge funds.

When looking at the previous studies that are mentioned earlier in section 2.2.2, there seems to exist better replication results over the financial crises than found in this study.
Kooli, M. & Sharma, S. (2012) found for the period 2004 to 2009, 7 out of 14 strategies did over-perform the hedge fund index by the using the Kalman filter replication method. They also found that 6 out of 14 strategies did over-perform the hedge fund by the use of the fixed-weight replication method. This could imply that hedge fund replication is possible for some strategies, but what still is doubtful, is the recovery of the clones after the financial crisis. Kooli, M. & Sharma, S. (2012) replication ends just after the upswing after the financial crises (2008), and does therefore not show if the clones are able to increase at the same speed as the hedge fund indices. By the small increase shown in their selected period it seems like the clones might have the same struggles after a huge drop, as found in this study.

When comparing the performance for the strategies in the pre-crisis period, with previously findings, the strategy distressed debt are showing equally performance. This could indicate that the clone for this strategy is not as influenced by huge drops in the market, which can be appalling for investing in hedge fund replication under this strategy.
The arbitrage strategy in Kooli, M. & Sharma, S. (2012) is not suitable for replication by the use of fixed-weight, but by the use of a method that better eliminates look-ahead bias and capture time changes, does the clone performs close to the hedge fund. This might imply that for replication of arbitrage strategy is it most likely to get the best result if the replication in done by a method that adjust the portfolio over time, such as rolling window or Kalman filter.

One of the strategies this study is struggling to replicate in the pre-crisis period is Marco. This strategy has, however, shown better performance in previous studies. This could indicate that the use of a different replication method might be a better tool, to replicate this strategy.

As Hasanhodzic, J. & W. Lo, A. (2007) fixed weight replication performed better on replicating this strategy.
Strategies that have shown difficulties when trying to replicate in this study in the pre-crisis period and the previous studies included are Event Driven and Relative Value, which could indicate true alpha or just an unexplained variable because one or more factors were left out, due to the limited knowledge about hedge funds.

The crisis period is the period that has the best results when replicating the different strategies, which can be due to the fact that no huge external events affect this period. The financial crisis is only in the first part of the estimation periods, where the recovery after the crisis is also included and has a larger effect on the linear regression estimates and the risk-free rate is on a steady development. Overall, a period with less fluctuation in the market as the market need to be stabilized after the crisis.
Even though this is the best performing period in this study, the replication clones still struggle to generate excess return relative to the hedge funds indices. Only two strategies are over-performing with a higher mean return compared to the hedge fund and positive information ratio. One is Relative Value, which was difficult to replicate under the financial crisis. This can indicate that under huge changes in the market, it is difficult to identify and exploit price discrepancies when only using predetermined factors and statistical calculations, where this is easier under financial times where the economy is stabilizing.
The other strategy that is over-performing is Equity Short Bias, which is mainly performing because of the huge drop in the end of the crisis-period, which the rolling window method is not capturing. The replication does not explain much about the strategy as the clone and hedge fund index are not close to be identical.

Distressed debt strategy is also, in the crisis period, showing a slight possibility for replication, graphically.[74] At the end of the period the indices for the clone and the hedge fund are almost equal, and this gives signs that a good replication is possible. However, looking at the performance measurements, the clone is not performing at an acceptable level, which also is seen looking into the returns. A better result can maybe be obtained by the use of

[74] Appendix 8.10

a different method or by selection more factors corresponding to the factors that have been shown to be involved in the strategy under replication. The opportunity for replication of this strategy compared to the other strategies can be due to the fact that this strategy clone is more stable and not as susceptible to drastic changes as other strategies.

The crisis period is still struggling hard in some strategies, such as Long Short Equity and Fixed Income. The two strategies are not replicating on an acceptable level in this period, but have shown better results in Kooli, M. & Sharma, S. (2012) research - although they are struggling in Hasanhodzic, J. & W. Lo, A. (2007) research, which both indicate possibilities for replications due of missed opportunities or difficulties when replicating those strategies as they are not showing any pattern in replication that can lead to a possible replication of the strategy.

In the post-crisis period most clones are extremely affected by the development in the risk-free rate. After the financial crisis, the Federal Reserve aimed to stimulate the economy again, by dropping the rate to almost zero, where it remained the following six years. This was done to encourage consumers to spend more and thereby make America recover after the crisis.
In 2017 when the economy began to see recovery, the development in risk-free followed and started to increase along with the appreciating of the US dollar.
This massive changes in the risk-free rate are affecting the replication method for this study in the post-crisis period and the replication clones are not good enough to capture it. This affects seven out of eleven strategies, so the result is not good replications. The clones are over-performing the hedge fund indices because of the change in the risk-free rate and we are therefore not able to conclude on the performance relative to the hedge fund or to create a deeper understanding or transparency of the strategy in the post-crisis period due to the massive influence the risk-free rate has on the replication in this study.

By trying to change elements in rolling window, such as length on in-sample period where the factors are set, length on rolling window, establishing gross returns instead of regular returns or by changing the factors over time, it still has not showed any trustworthy results, which indicates missed opportunities to replicate hedge fund strategies doing massive changes in the economy.

Even on the strategies that seems to replicate on an acceptable level in the post-crisis period, such an Event Driven, the clone is still struggling to fit over the entire out-sample period.

The US risk-free rate is directly connected to the asset US Treasury Bond (T-bill), which is therefore considered a risk-free asset where any other investment must offer a better rate to keep being attractive. Along this and the fact that the risk-free rate in the post-crisis period was expected to increase did many investors place their money in America and this lead to huge pressure on the emerging markets. This can be shown to some extent in this study as all strategies is either take no position in the emerging market or short positions, but as the replication results in the post-crisis period are not performing on a level to create trustworthy knowledge is this only a light conclusion.

Another issue that appeared in the post-crisis period is the changes in the spread between Morgan Stanley Europe, Australasia and Far East Index and MXUS Largge Cap, described under strategy Long Short Equity in the analysis. The drastic changes are due to fluctuations in the US large cap in the period September 2018 through November 2018. This change does have a huge effect on the clones where this factors is included. To overcome this, a method that adjusts the portfolio over time can be used, like the extended rolling window or Kalman filter. These methods adjust to time changes and reflect the manager's work within a hedge fund. In this study and Kooli, M. & Sharma, S. (2012) research, these methods have shown better replication results than more passive replication methods and can therefore be more appealing to look into when aiming to replicate hedge fund return.

The last topic this discussion will look into is the different possibilities for not having the best possible fit to the model. At first, more factors can give a better fit in the in-sample period but worse fit over the out-sample period. As the explanatory power of the model will decrease with too many factors in the model, and lead to a potential over-fit. This study did set some predefined settings for the factors to include in the clone, because of previous results that have shown that models with to many factors are seen to deliver less good clones than predicted.
In this study, 15 of 33 clons are containing more than 8 factors (excluding the risk-free rate),

which is seen as many factors in a replication. These 15 models are performing on various levels, but perhaps the performance would increase, with fewer factors.

Furthermore, the statistical test for autocorrelation, normality and heteroscedasticy are reveling difficulties to get the best possible fit. Through the entire analysis, autocorrelation is shown in the residuals, and in crisis period, several of the models are showing signs of homoscedasticy and non-normality in the factors. This could lead to doubt in the good results, but this is not possible to conclude.

7 Conclusion

To be able to create hedge fund clones that can deliver a similar performance to hedge funds is an attractive field to investigate and find answers to, as this will eliminate some of the downsides of hedge funds, but still deliver the same returns and performance. Analyzing some of the Eurekahedge hedge fund indices by the use of the factor-based approach, rolling window with a 36-month estimation window, has some hedge fund replication shown some possibility for delivering a similar performance within different hedge fund strategies. But by breaking down the out-sample period in to smaller sample sizes to analyze different financial periods, pre-crisis, crisis and post-crisis, this study found that external events influenced the replication in a negative way. This could be things such as the huge economic changes under the financial crises in 2008 and the following near zero risk-free rate.

This can, to some extent, be overcome by extending the replication method, rolling window, to operate more like a manager in a hedge fund, by changing the factors over time when they are no longer significant for the clone, or adding factors to the model when they are significant on a predetermined significant level. But no strong results from this method have been shown and this study therefore does not overcome the huge financial changes in the market, which makes hedge fund replication less trustworthy.

This addition to the rolling window method has, however, overcome drastic changes in the factors for the clone and does overall deliver a more similar performance to the hedge fund compared to rolling window with fixed factors. This finding was also found in a previous study by Kooli, M. & Sharma, S. (2012).

Within the pre-crisis, crisis, and post-crisis periods all three periods did have their own difficulties. The pre-crisis period clones suffered from the drop in the market in 2008. The crisis period clones had the most statistical fluctuations and deficiencies, which can lead to less trustworthy result, and unbiased models. The post-crises period clones where affected by the development in the risk-free rate which made conclusion on transparency and performance difficult.

However, some of the strategies that this study has had most success replicating are both strategies that have seen possibilities within previous studies and strategies that have not

shown possibilities for replication before. This does, therefore, strengthen the replication opportunities within some strategies and maybe shows possibilities within new strategies, which perhaps can lead to a deeper understanding of hedge funds in the future.

Overall, the clones for this study are struggling to over performing the hedge fund indices, and therefore not creating the wanted transparency or performance. This could be due to the factors or method chosen, or it could imply that the managers do contribute with knowledge/alpha that hedge fund replication methods cannot create. It could, therefore, be interesting to further investigate within the field of hedge fund replication to look more deeply into the approach Payoff distribution replication, as this method shows some stronger results on replication where the factor-based approach is still not showing any strong results.

References

Articles

[1] Amenc, N. & Martiellini, L. & Mayfredi J-C. & Ziemann, V. (2009). Performance of Passive Hedge Fund Replication Strategies. Nice, France: EDHEC Risk and Asset Management Research Centre.

[2] EDHEC Risk & Asset Management Research Centre publication. (2008). The pros and Cons of Passive Hedge Fund Replication. Nice, France: EDHEC Risk and Asset Management Research Centre.

[3] Hasanhodzic, J. & W. Lo, A. (2007) Can Hedge-Fund Returns be Replicated?: The Linear Case. Journal of investment management, Vol 5, No. 2, 5-45.

[4] Jawadi, F & Khanniche, S. (2011). Are Hedge fund clones attractive financial products for investors?. Applied Economics Letters, 19:8, 739-743. Doi: 10.1080/13504851.2011.599783

[5] Kat, H.M. & Palaro, H.P. (2005) Who Needs Hedge Funds? A Copula-Based Approach to Hedge Fund Return Replication. Alternative investment research centre working paper series, 0027, 4-52.

[6] Kat, H.M. & Palaro, H.P. (2006). Replication and Evaluation of Fund of Hedge Funds Returns. Alternative investment research centre working paper series, 0028, 3-24.

[7] Kat, H.M. & Palaro, H.P. (2006). Superstars or average joes? Alternative investment research centre working paper series, 0030, 3-30.

[8] Kooli, M. & Sharma, S. (2012). Should we give hedge funds clones a chance?, Vol 38, No. 1, 44-66

[9] 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, Vol 7, No. 1, 15. Doi: 10.3390/ijfs7010015

[10] Osborne, J. & Waters, E. (2002). Four Assumptions of multiple regression that researchers should always test. Practical Assessment, Research & Evaluation, Vol 8, No. 2, 1-5.

[11] Takahashi, A. & Yamamoto, K. (2013) Generating a target payoff distribution with the cheapest dynamic portfolio: an application to hedge fund replication. Quantitative Finance, 13:10, 1559-1573, Doi: 10.1080/14697688.2013.779014

[12] Yobero, C. (2016). Methods for Detecting and Resolving Heteroskedasticity. Rpubs.com

Books

[13] Baker, H. & Filbeck, G. (2017). Hedge Funds: Structure, Strategies, and Performance. Oxford, England: Oxford University Press.

[14] Dalgaard P. (2008). Introductory Statistics with R (2nd). Copenhagen, Danmark: Springer.

[15] Damodar, G & Porter, D. (2009). Basic Econometrics. London, Englang: McGraw-Hill.

[16] Gregoriou, G. & Kooli, M. (2012). Hedge Fund Replication. Hampshire, England: Palgrave Macmillan

[17] McCann, B. (2014) Tacitcal Portfolios: Strategies and Tactics for Investing in Hedge Funds and Liquid Alternatives. Hoboken, NJ, USA: John Wiley & Sons, Inc.

[18] Newbold, P. & Carlson, W. & Thorne, B. (2013). Statistics for Business and Economics (8. edition). London, England: Pearson Education UK

[19] Rosefielde, S. & Kuboniwa, M. & Mizobata, S. (2011). Two Asias: The Emerging Postcrisis Divide. United Kingdom: World Scientific Pub Co Pte.

[20] Wackerly, D. & Mendenhall III, W. & Scheaffer, R. (2008) Mathematical Statistics with Applications (7.). Belmont, USA: Thomson Learning, Inc.

Websites

[21] https://www.bloomberg.com/

[22] https://cran.r-project.org/

[23] http://www.eurekahedge.com/

[24] https://finansdanmark.dk/

[25] https://www.investopedia.com.

[26] https://investor.vanguard.com/

[27] https://www.ishares.com/

[28] https://www.msci.com/

[29] https://www.prudential.co.za/

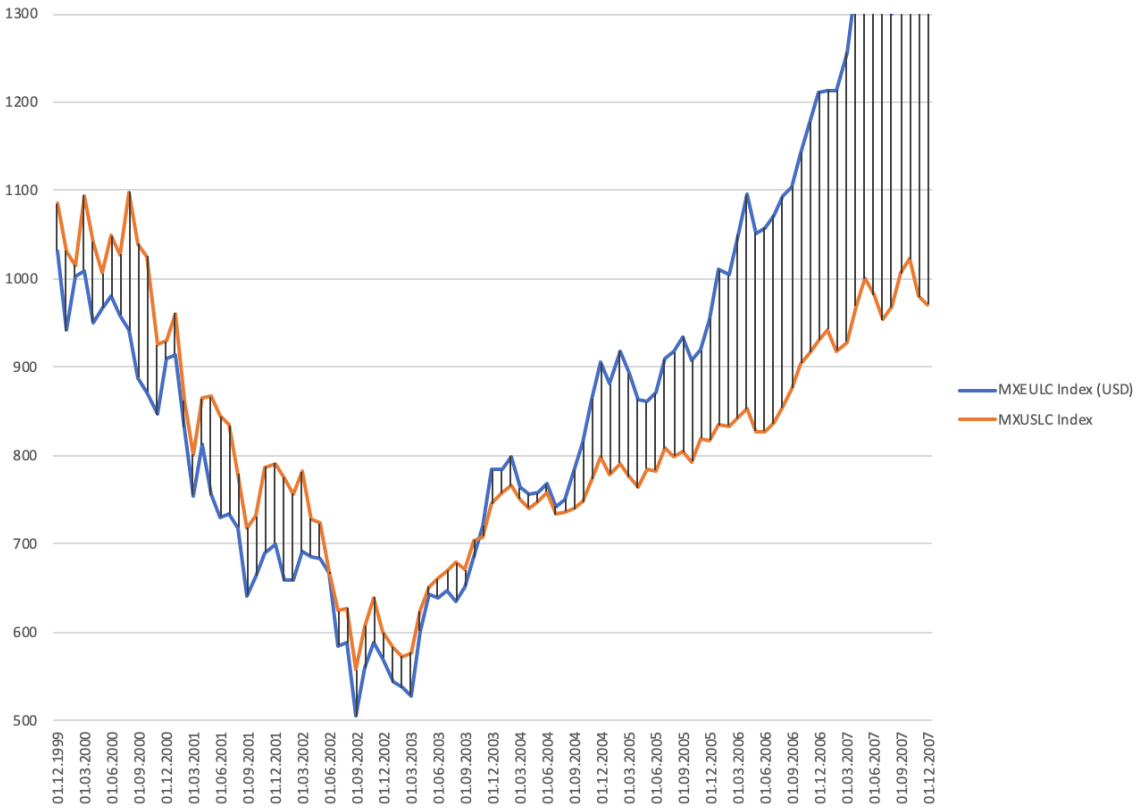
[30] https://us.spindices.com/

Other sources

[31] Bloomberg Terminals (data source)

8 Appendix

8.1 Spread between Morgen Stanley Large Cap respectively Europe and US



MXEULC vs. MXUSLC

8.2 Jarque Bera

```
> col_jarquebera(Insample1)
                      obs     skewness  kurtosis df    statistic       pvalue
BCOM                   96  0.036254812  2.514354  2 9.644385e-01 6.174117e-01
DHY                    96  0.385058090  6.143305  2 4.189378e+01 7.996154e-10
DJUSGR                 96 -0.485248786  3.729309  2 5.895028e+00 5.246998e-02
DJUSVA                 96 -0.186947251  4.374760  2 8.119053e+00 1.725719e-02
EEM                    96  0.194676939  4.277851  2 7.137997e+00 2.818406e-02
GI1                    96 -0.123134509  2.775449  2 4.442863e-01 8.008007e-01
LUACTRUU               96 -0.519506931  4.312503  2 1.120886e+01 3.681524e-03
MXEA                   96 -0.506369041  3.097531  2 4.140603e+00 1.261478e-01
MXEF000G               96 -0.466119910  2.677001  2 3.893598e+00 1.427303e-01
MXEU000G               96 -0.306202942  3.288088  2 1.832142e+00 4.000879e-01
MXEU000V               96 -0.525389682  4.494558  2 1.335136e+01 1.261212e-03
MXWOLC                 96 -0.482241237  3.156938  2 3.819423e+00 1.481231e-01
MXWOSC                 96 -0.255219957  3.208547  2 1.216164e+00 5.443941e-01
SBBMGLU                96 -0.538512038  3.007941  2 4.640176e+00 9.826495e-02
SCRTEM                 96 -0.503885792  3.041735  2 4.069381e+00 1.307209e-01
SPBDU10T               96 -0.581753260  3.932963  2 8.896671e+00 1.169802e-02
DJUSGR.vs.DJUSVA       96 -3.228164000 29.525868  2 2.981223e+03 0.000000e+00
MXUS000G.vs..MXUS000V  96  0.356357682  4.526766  2 1.135592e+01 3.420536e-03
MXEA.vs..MXUSLC        96 -0.304584647  2.778239  2 1.681061e+00 4.314816e-01
MXEULC.vs..MXUSLC      96  8.417868664 77.668552  2 2.343534e+04 0.000000e+00
MXWOLC.vs.MXWOSC       96 -0.434503482  3.540482  2 4.189175e+00 1.231210e-01
USDEUR                 96 -0.105086878  2.938051  2 1.920426e-01 9.084447e-01
USDGBP                 96 -0.001041029  2.740452  2 2.694790e-01 8.739435e-01
USDJPY                 96  0.170690798  3.219127  2 6.582327e-01 7.195593e-01
GB1M                   96  0.109190589  1.566604  2 8.409263e+00 1.492629e-02
```

(a) Pre-crisis period

```
> col_jarquebera(Insample2)
                      obs     skewness   kurtosis df    statistic       pvalue
BCOM                  132 -0.634864772   5.094349  2 3.299180e+01 6.853642e-08
DHY                   132 -0.215295437   5.601048  2 3.822971e+01 4.994863e-09
DJUSGR                132 -0.546686520   3.519617  2 8.060068e+00 1.777373e-02
DJUSVA                132 -0.643791601   4.452387  2 2.072014e+01 3.167231e-05
EEM                   132 -0.419805761   5.353692  2 3.434646e+01 3.481448e-08
GI1                   132 -0.552053690   4.592009  2 2.064450e+01 3.289297e-05
LUACTRUU              132 -0.923803790   8.424898  2 1.806374e+02 0.000000e+00
MXEA                  132 -0.682211144   4.343161  2 2.016151e+01 4.187776e-05
MXEF000G              132 -0.659865082   3.965138  2 1.470248e+01 6.417955e-04
MXEU000G              132 -0.533359266   3.914959  2 1.086271e+01 4.377155e-03
MXEU000V              132 -0.446235833   4.305572  2 1.375562e+01 1.030396e-03
MXWOLC                132 -0.621266628   3.931029  2 1.325888e+01 1.320905e-03
MXWOSC                132 -0.606853173   4.587361  2 2.196038e+01 1.703583e-05
SBBMGLU               132 -0.733253896   4.542160  2 2.490897e+01 3.900201e-06
SCRTEM                132 -0.657155632   4.558063  2 2.285235e+01 1.090623e-05
SPBDU10T              132  0.009017577   4.598526  2 1.405585e+01 8.867680e-04
DJUSGR.vs.DJUSVA      132 -3.783845868  40.108447  2 7.888687e+03 0.000000e+00
MXUS000G.vs..MXUS000V 132  0.305144160   4.566763  2 1.554958e+01 4.201949e-04
MXEA.vs..MXUSLC       132 -0.438400150   3.444788  2 5.316385e+00 7.007476e-02
MXEULC.vs..MXUSLC     132  9.856616849 106.442827  2 6.098967e+04 0.000000e+00
MXWOLC.vs.MXWOSC      132 -0.528644099   3.660790  2 8.549757e+00 1.391374e-02
USDEUR                132  0.296761950   4.118801  2 8.821927e+00 1.214347e-02
USDGBP                132  0.518866999   5.337813  2 3.598243e+01 1.536437e-08
USDJPY                132  0.377060125   3.334851  2 3.744524e+00 1.537754e-01
GB1M                  132  0.399048424   1.773172  2 1.178136e+01 2.765097e-03
EHFI251               132 -0.338775565   4.283897  2 1.159107e+01 3.041110e-03
EHFI252               132 -0.285810483   4.535431  2 1.476365e+01 6.224642e-04
EHFI253               132  0.546668930   3.831485  2 1.037715e+01 5.579958e-03
EHFI254               132 -0.895384581   5.924253  2 6.466960e+01 9.103829e-15
EHFI255               132 -0.462833823   4.714908  2 2.088774e+01 2.912625e-05
EHFI285               132 -2.741932346  17.917155  2 1.389269e+03 0.000000e+00
EHFI287               132 -1.043161279   7.498760  2 1.352537e+02 0.000000e+00
EHFI288               132 -1.190954833   7.550582  2 1.450971e+02 0.000000e+00
EHFI289               132 -1.923024330  13.783894  2 7.209645e+02 0.000000e+00
EHFI751               132  0.131547428   7.430529  2 1.083434e+02 0.000000e+00
EHFI807               132  0.186337162   2.848283  2 8.904738e-01 6.406725e-01
```

(b) Crisis period

```
> col_jarquebera(Insample3)
                      obs     skewness  kurtosis df    statistic       pvalue
BCOM                   81 -0.210216306  3.519097  2 1.506009e+00 0.4709494287
DHY                    81  0.054404915  2.920491  2 6.129424e-02 0.9698177429
DJUSGR                 81 -0.005710795  3.032027  2 3.902039e-03 0.9980508825
DJUSVA                 81 -0.128373835  3.009383  2 2.227750e-01 0.8945920320
EEM                    81  0.092894183  3.872657  2 2.686663e+00 0.2609747974
GI1                    81 -0.290542872  2.880208  2 1.188036e+00 0.5521043588
LUACTRUU               81 -0.203308959  2.951542  2 5.659413e-01 0.7535418949
MXEA                   81 -0.226765818  3.080043  2 7.158302e-01 0.6991324327
MXEF000G               81 -0.092390642  3.582037  2 1.258577e+00 0.5329708535
MXEU000G               81 -0.189475966  3.400859  2 1.026988e+00 0.5984011898
MXEU000V               81 -0.030686358  2.757033  2 2.119479e-01 0.8994480905
MXWOLC                 81 -0.161261880  3.193581  2 4.775463e-01 0.7875935393
MXWOSC                 81 -0.179409978  3.426067  2 1.047212e+00 0.5923805999
SBBMGLU                81 -0.163749057  3.417078  2 9.490800e-01 0.6221711957
SCRTEM                 81 -0.032254656  3.338367  2 4.004553e-01 0.8185444015
SPBDU10T               81  0.225194166  2.793904  2 8.279730e-01 0.6610098729
DJUSGR.vs.DJUSVA       81  1.924436774  9.228645  2 1.809332e+02 0.0000000000
MXUS000G.vs..MXUS000V  81  0.311103534  3.509137  2 2.181474e+00 0.3359688750
MXEA.vs..MXUSLC        81 -0.202411889  2.490978  2 1.427576e+00 0.4897853234
MXEULC.vs..MXUSLC      81 -6.186399704 53.845719  2 9.242010e+03 0.0000000000
MXWOLC.vs.MXWOSC       81 -0.146373919  3.047798  2 2.969527e-01 0.8620204029
USDEUR                 81  0.381991420  3.522776  2 2.892255e+00 0.2354804731
USDGBP                 81  0.517561815  3.527671  2 4.555970e+00 0.1024904968
USDJPY                 81 -0.022914021  3.103283  2 4.309049e-02 0.9786851952
GB1M                   81  1.114440303  3.350823  2 1.718208e+01 0.0001857632
EHFI251                81 -0.377571087  3.203797  2 2.064733e+00 0.3561630545
EHFI252                81 -0.594625791  3.535660  2 5.741720e+00 0.0566501732
EHFI253                81 -0.152697357  2.745853  2 5.327655e-01 0.7661458102
EHFI254                81 -0.502965033  3.186206  2 3.532166e+00 0.1710014584
EHFI255                81 -0.496315294  3.713526  2 5.043718e+00 0.0803101758
EHFI285                81 -0.557638415  3.295716  2 4.493105e+00 0.1057631910
EHFI287                81  0.130342675  4.820564  2 1.141563e+01 0.0033199211
EHFI288                81 -0.559893563  3.210302  2 4.381257e+00 0.1118464106
EHFI289                81 -0.268080190  2.984679  2 9.709966e-01 0.6153904779
EHFI751                81 -0.904468821  3.998287  2 1.440731e+01 0.0007438611
EHFI807                81  0.260703091  3.709714  2 2.617508e+00 0.2701564701
```

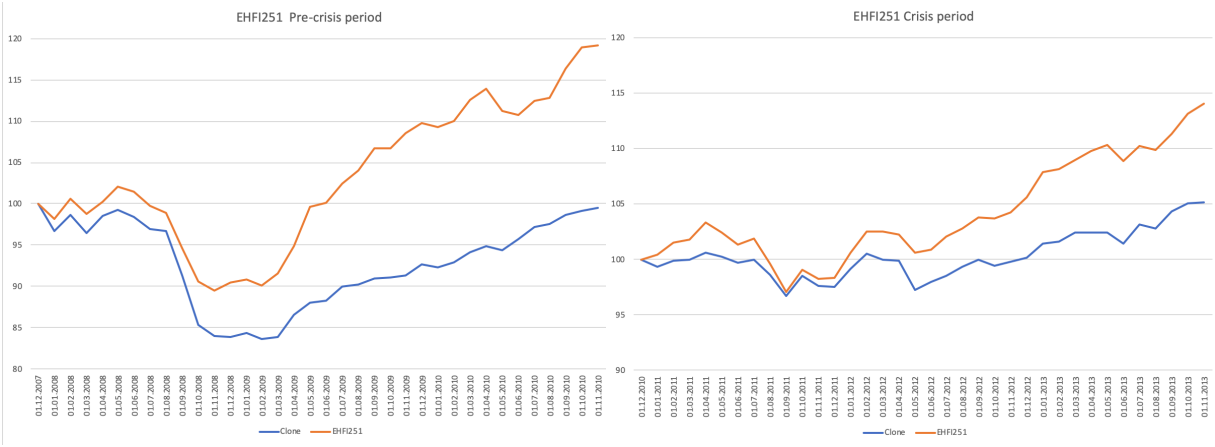
(c) Post-crisis period

Figure 20: Jarque Bera test of all factors and Eurekahedge

8.3 Clone Performance Measurements

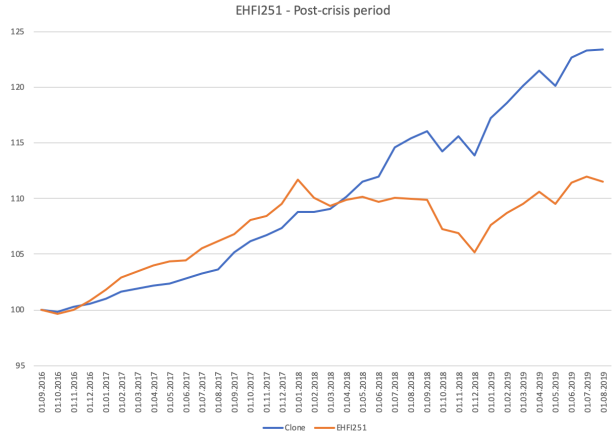
	Period	31/12/2007-30/11/2010										
Hedge Fund	Strategy	Mean Hedge Fund	Mean	St. Dev. Hedge Fund	Standard Deviation	Sharpe Ratio Hedge Fund	Sharpe Ratio	Correlation index	RMSE	Tracking Error	Information ratio	Theils Inequality Coefficient
EHFI251	Overall	0,535%	0,005%	2,016%	1,947%	1,320%	-25,871%	64,347%	0,01149	1,046%	-0,50692	0,07938
EHFI252	Long Short Equity	0,315%	-0,162%	2,903%	2,515%	-6,672%	-26,686%	70,575%	0,01305	1,248%	-0,38262	0,07610
EHFI253	Marco	0,680%	-0,101%	1,171%	1,715%	14,588%	-35,565%	-18,994%	0,01682	1,532%	-0,50948	0,14379
EHFI254	Multi Strategy	0,490%	0,072%	1,860%	1,778%	-1,011%	-24,592%	78,072%	0,00969	0,899%	-0,46561	0,06974
EHFI255	Relative Value	0,697%	0,230%	1,680%	1,671%	11,211%	-16,713%	86,988%	0,00947	0,847%	-0,55210	0,07008
EHFI285	Arbitrage	0,565%	0,408%	1,766%	1,463%	3,190%	-6,883%	94,560%	0,01238	1,262%	-0,12438	0,09098
EHFI287	Distressed Debt	0,573%	0,552%	3,299%	2,872%	1,941%	1,501%	93,570%	0,02418	2,485%	-0,00842	0,13140
EHFI288	Event Driven	0,659%	0,296%	3,216%	3,031%	4,671%	-7,032%	92,060%	0,01320	1,306%	-0,27823	0,07236
EHFI289	Fixed Income	0,629%	0,400%	1,858%	1,809%	6,459%	-6,027%	94,326%	0,01204	1,215%	-0,18857	0,08576
EHFI751	Equity Market Neutral	0,333%	0,147%	0,800%	1,127%	-22,013%	-32,123%	75,240%	0,01059	1,070%	-0,17386	0,11361
EHFI807	Equity Short Bias	0,327%	0,391%	3,299%	3,897%	-5,523%	-3,030%	69,631%	0,03374	3,468%	0,01850	0,18247
	Period	31/12/2010-29/11/2013										
Hedge Fund	Strategy	Mean Hedge Fund	Mean	St. Dev. Hedge Fund	Standard Deviation	Sharpe Ratio Hedge Fund	Sharpe Ratio	Correlation index	RMSE	Tracking Error	Information ratio	Theils Inequality Coefficient
EHFI251	Overall	0,448%	0,176%	1,228%	0,986%	33,021%	13,557%	94,532%	0,00617	0,519%	-0,52362	0,05406
EHFI252	Long Short Equity	0,544%	0,274%	1,979%	1,215%	25,375%	19,075%	98,764%	0,00941	0,900%	-0,30057	0,06586
EHFI253	Marco	0,296%	0,124%	0,787%	0,529%	32,250%	15,376%	89,108%	0,00689	0,645%	-0,26736	0,07548
EHFI254	Multi Strategy	0,426%	0,188%	1,194%	0,927%	32,127%	15,727%	96,393%	0,00537	0,492%	-0,48411	0,04786
EHFI255	Relative Value	0,498%	0,627%	1,078%	1,675%	42,264%	34,945%	98,635%	0,00943	0,959%	0,13507	0,08589
EHFI285	Arbitrage	0,442%	0,269%	0,654%	0,751%	61,163%	30,172%	97,061%	0,00523	0,468%	-0,37095	0,05887
EHFI287	Distressed Debt	0,867%	0,819%	1,477%	2,171%	55,813%	35,788%	95,755%	0,01496	1,539%	-0,03087	0,11320
EHFI288	Event Driven	0,621%	0,456%	1,767%	1,989%	32,765%	20,823%	98,685%	0,00764	0,763%	-0,21597	0,05554
EHFI289	Fixed Income	0,602%	0,372%	0,917%	1,144%	61,063%	28,849%	96,512%	0,00635	0,595%	-0,38569	0,06053
EHFI751	Equity Market Neutral	0,371%	0,348%	0,537%	1,176%	61,278%	25,985%	93,302%	0,00958	0,982%	-0,02372	0,11780
EHFI807	Equity Short Bias	-0,623%	0,323%	3,231%	2,813%	-20,603%	9,984%	11,885%	0,03392	3,339%	0,28348	0,18602
	Period	30/09/2016-30/08/2019										
Hedge Fund	Strategy	Mean Hedge Fund	Mean	St. Dev. Hedge Fund	Standard Deviation	Sharpe Ratio Hedge Fund	Sharpe Ratio	Correlation index	RMSE	Tracking Error	Information ratio	Theils Inequality Coefficient
EHFI251	Overall	0,323%	0,592%	0,957%	0,908%	-116,406%	-93,031%	79,409%	0,00693	0,645%	0,41691	0,06901
EHFI252	Long Short Equity	0,385%	-1,353%	1,423%	7,170%	-73,935%	-38,911%	-14,659%	0,06791	6,754%	-0,25731	0,48346
EHFI253	Marco	0,289%	0,410%	0,816%	0,686%	-140,624%	-149,824%	71,067%	0,00886	0,902%	0,13419	0,09543
EHFI254	Multi Strategy	0,326%	0,022%	0,908%	1,762%	-122,362%	-80,305%	30,160%	0,01529	1,542%	-0,19708	0,15425
EHFI255	Relative Value	0,290%	1,042%	0,940%	0,854%	-122,094%	-46,314%	88,851%	0,01173	0,900%	0,83578	0,11798
EHFI285	Arbitrage	0,276%	0,837%	0,642%	0,704%	-180,883%	-85,272%	93,958%	0,00787	0,548%	1,02417	0,09388
EHFI287	Distressed Debt	0,468%	1,483%	0,952%	1,390%	-101,825%	3,314%	85,355%	0,01727	1,395%	0,72763	0,16550
EHFI288	Event Driven	0,410%	0,475%	1,264%	0,774%	-81,202%	-124,426%	90,874%	0,00823	0,826%	0,07765	0,07162
EHFI289	Fixed Income	0,374%	0,769%	0,510%	0,675%	-208,578%	-98,889%	95,512%	0,00648	0,513%	0,77098	0,08123
EHFI751	Equity Market Neutral	0,110%	0,560%	0,484%	0,446%	-273,955%	-196,798%	72,869%	0,00698	0,550%	0,81811	0,09922
EHFI807	Equity Short Bias	-0,453%	-0,186%	3,734%	3,559%	-50,614%	-45,610%	50,900%	0,03777	3,837%	0,06947	0,19292

8.4 Main Hedge Fund Index



(a) Pre-crisis period

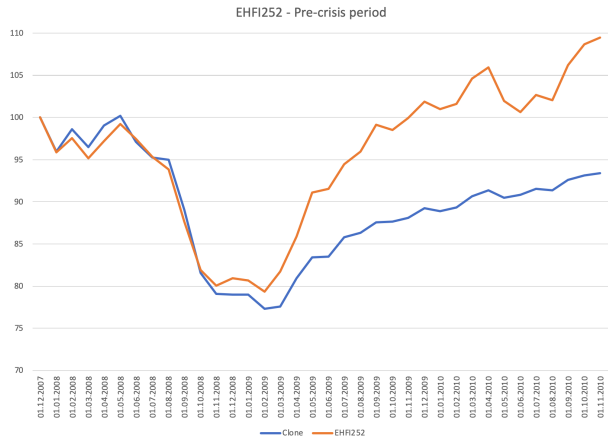
(b) Crisis period



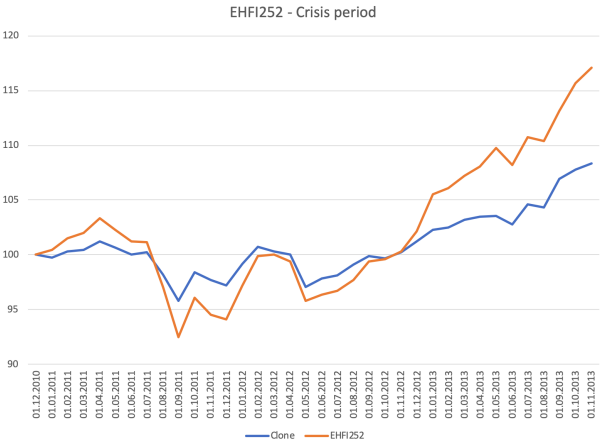
(c) Post-crisis period

Figure 21: Main Index

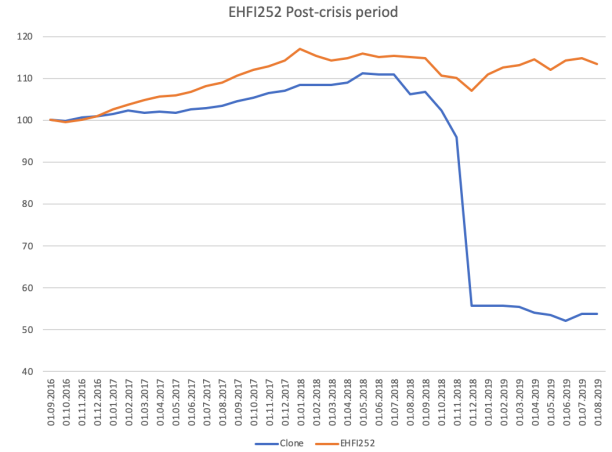
8.5 Long Short Equity Index

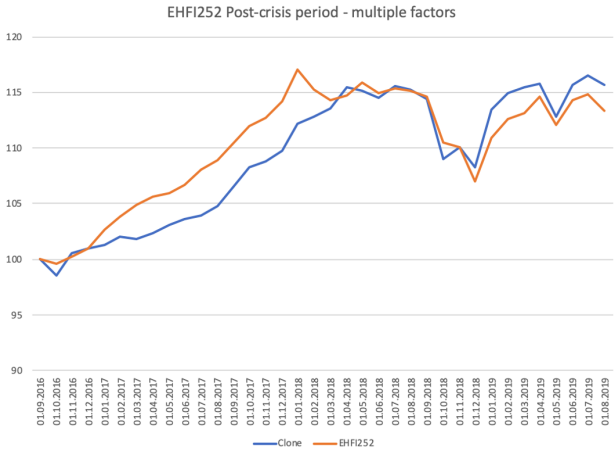


(a) Pre-crisis period



(b) Crisis period

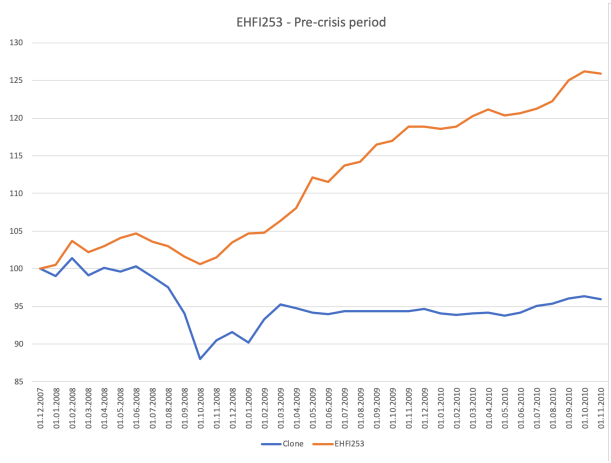




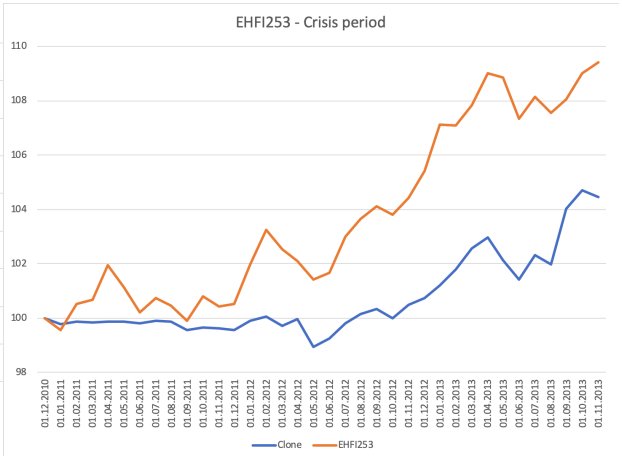
(c) Post-crisis period RW vs. Extended RW

Figure 22: Long Short Equity Strategy

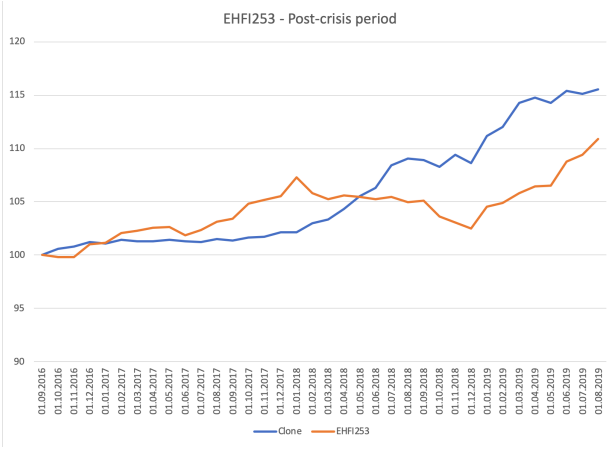
8.6 Marco Index



(a) Pre-crisis period



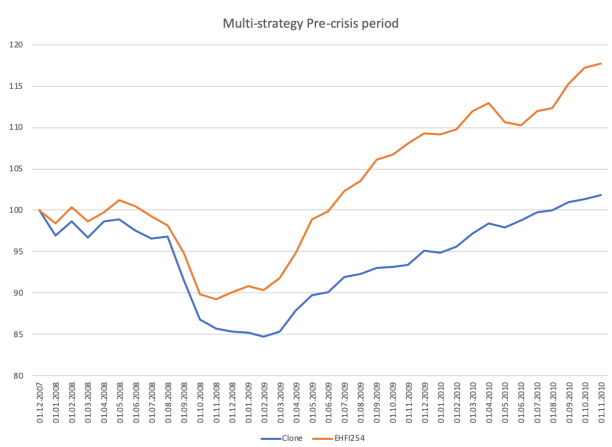
(b) Crisis period



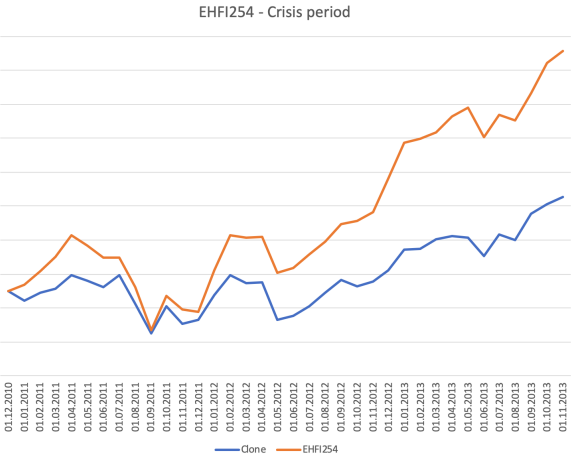
(c) Post-crisis period

Figure 23: Marco strategy

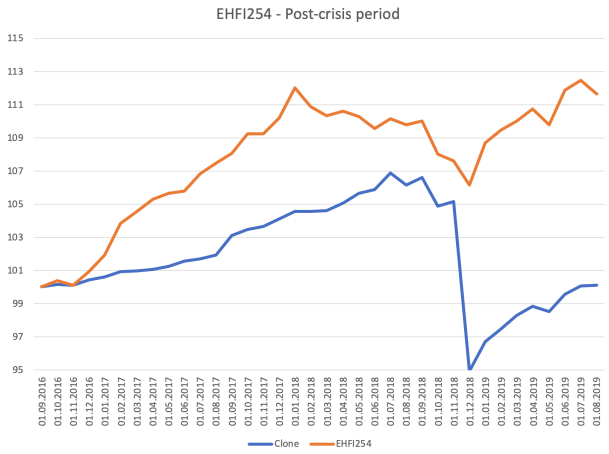
8.7 Multi-Strategy Index



(a) Pre-crisis period



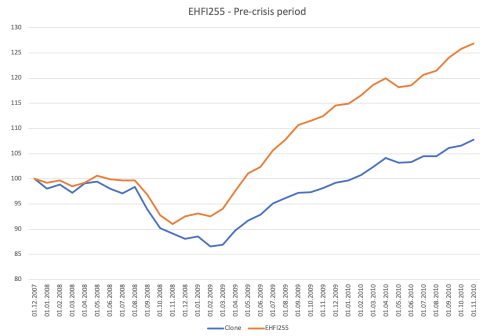
(b) Crisis period



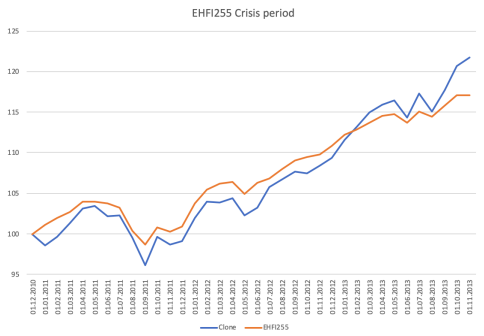
(c) Post-crisis period

Figure 24: Multi-Strategy

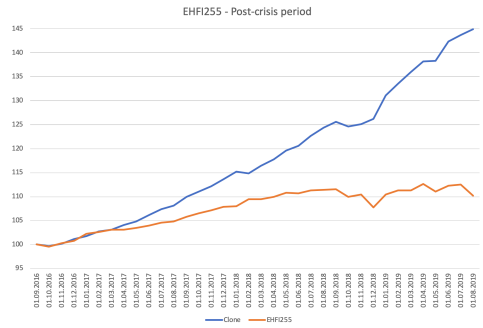
8.8 Relative Value Index



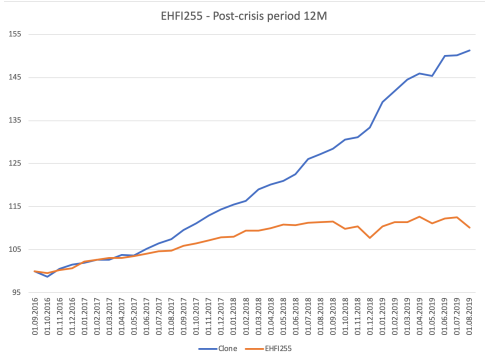
(a) Pre-crisis period



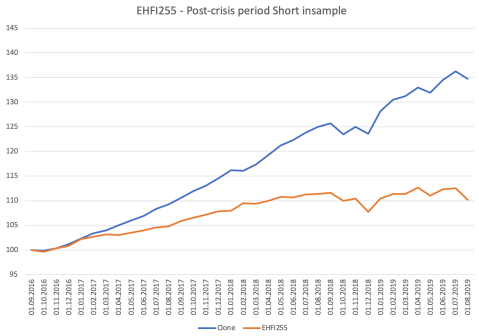
(b) Crisis period



(c) Post-crisis period RW - 36M



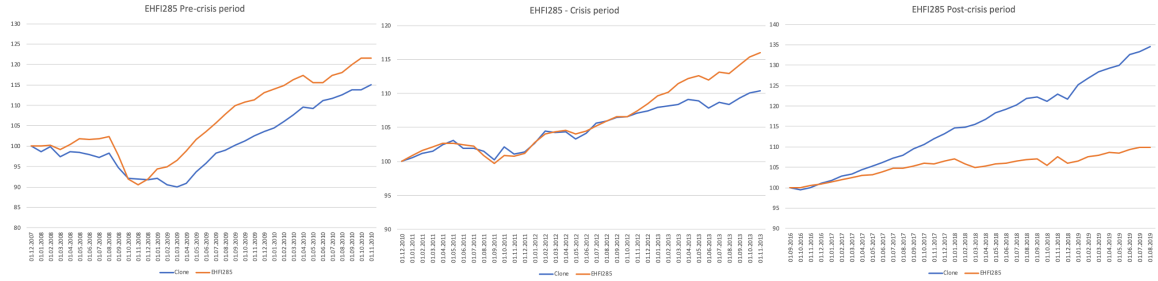
(d) Post-crisis period RW - 12M



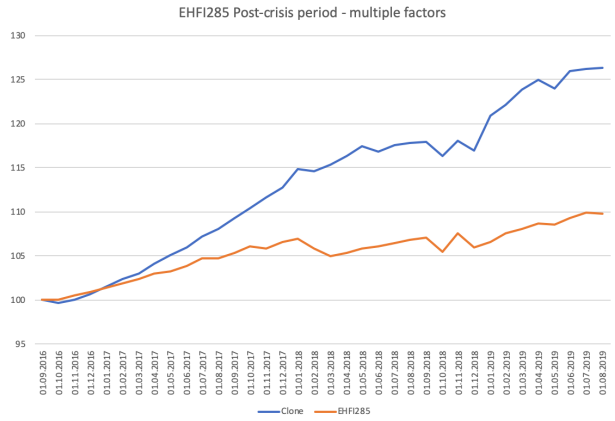
(e) Post-crisis period - RW short in-sample

Figure 25: Relative Value strategy

8.9 Arbitrage Index



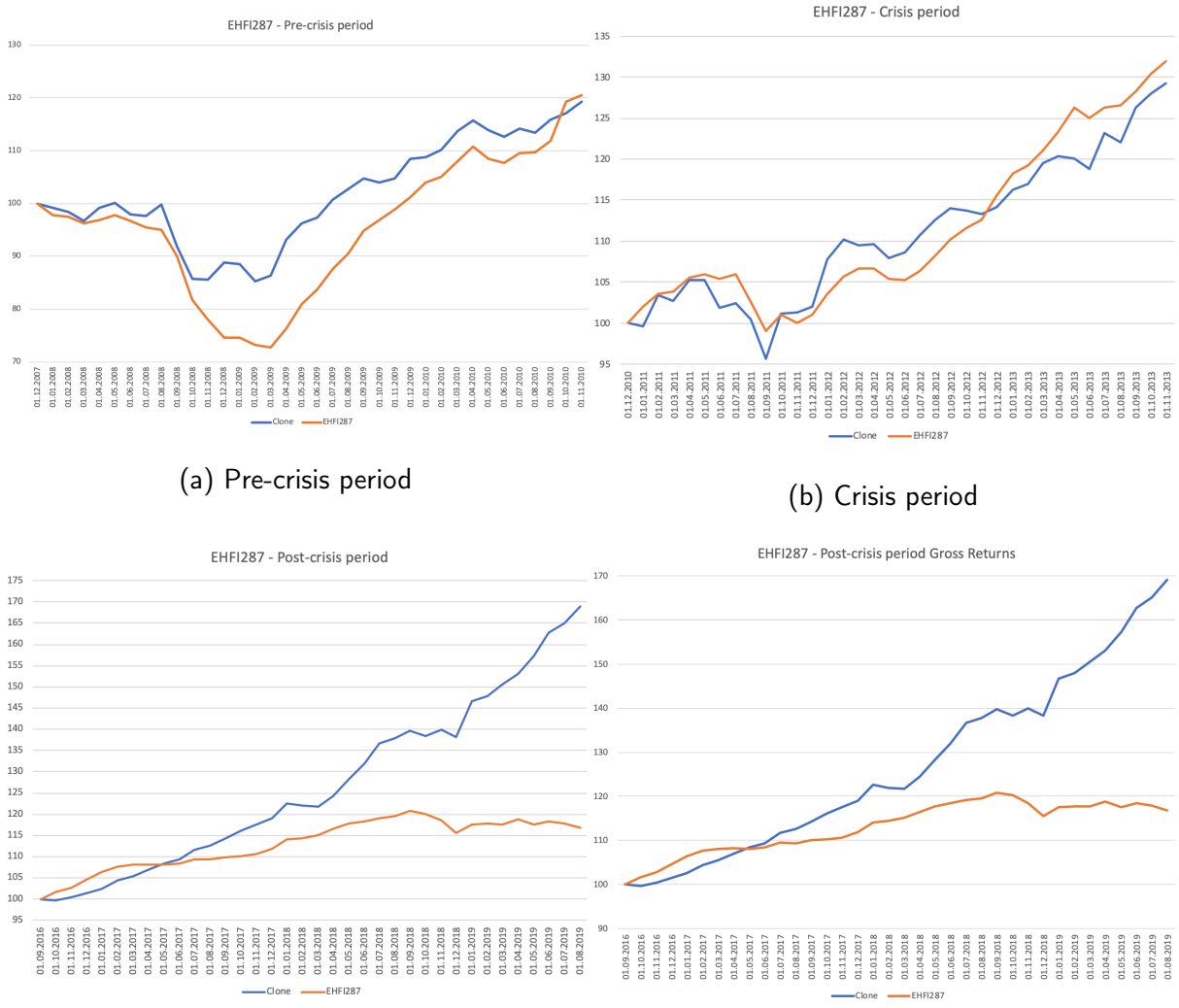
(a) Rolling Window



(b) Post-crisis period - Extended Rolling Window

Figure 26: Arbitrage strategy

8.10 Distressed Debt Index



(a) Pre-crisis period

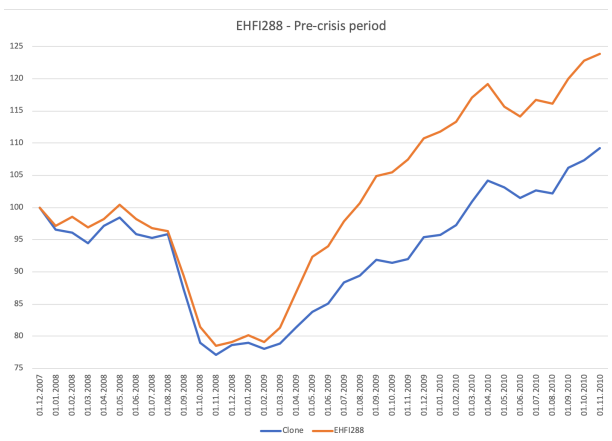
(b) Crisis period

(c) Post-crisis period

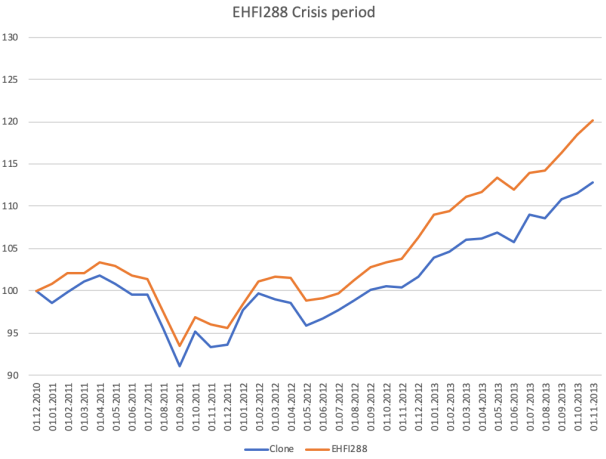
(d) Post-crisis period - GR

Figure 27: Distressed Debt strategy

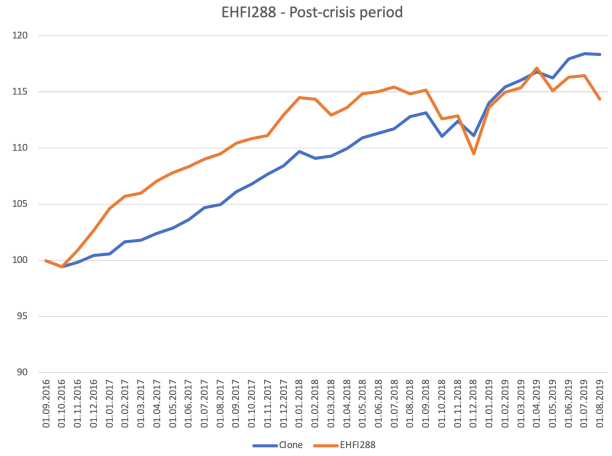
8.11 Event Driven Index



(a) Pre-crisis period



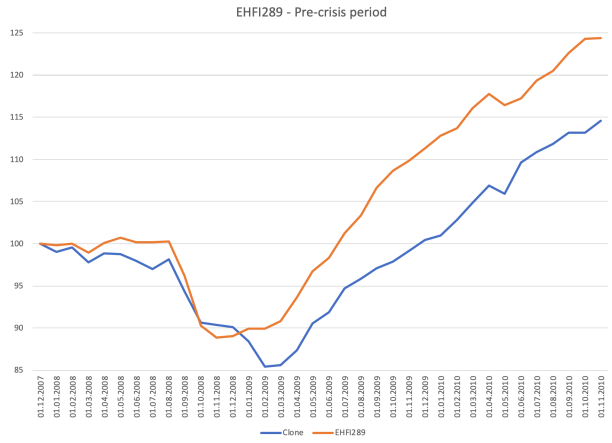
(b) Crisis period



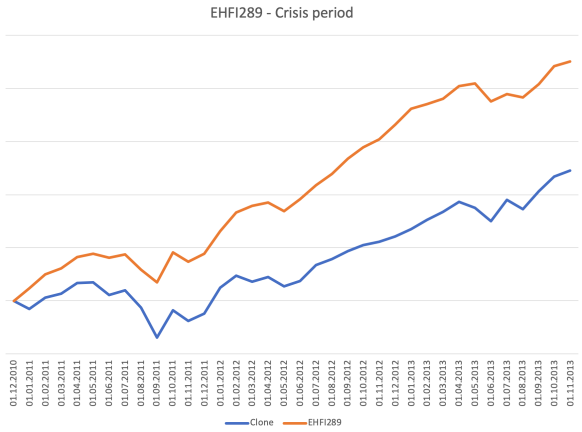
(c) Post-crisis period

Figure 28: Event Driven strategy

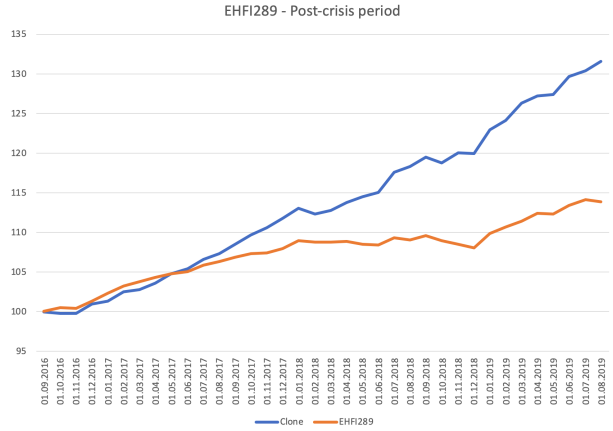
8.12 Fixed Income Index



(a) Pre-crisis period



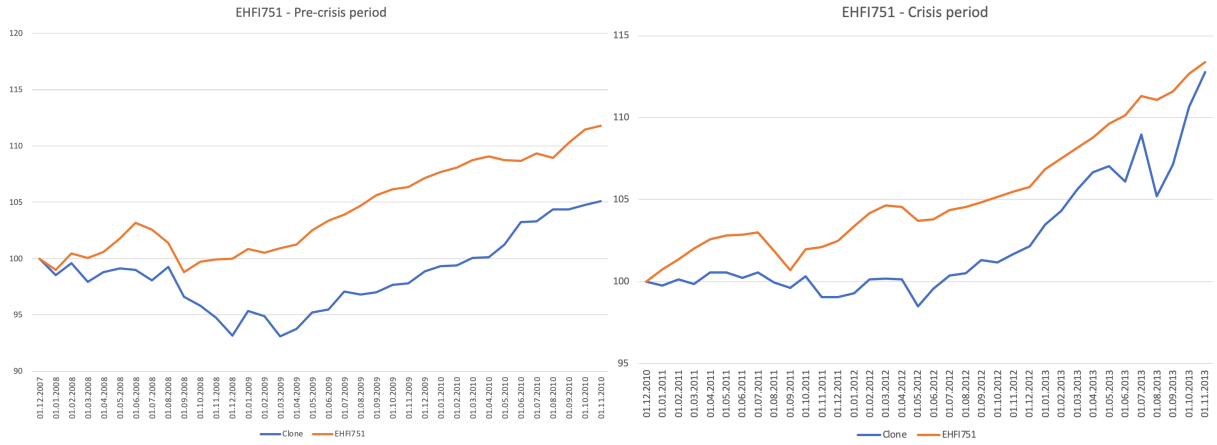
(b) Crisis period



(c) Post-crisis period

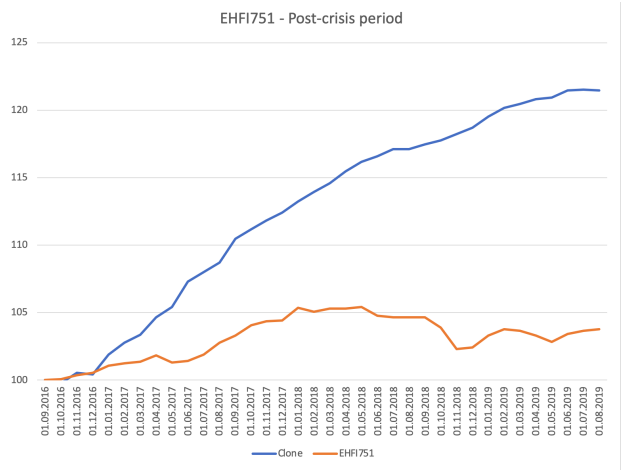
Figure 29: Fixed Income strategy

8.13 Equity Market Neutral Index



(a) Pre-crisis period

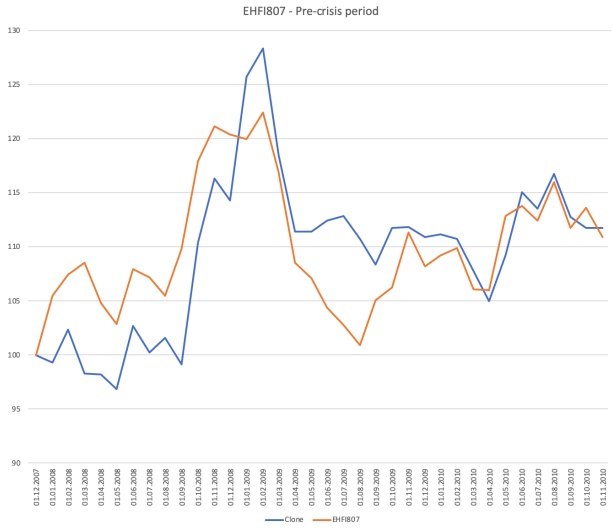
(b) Crisis period



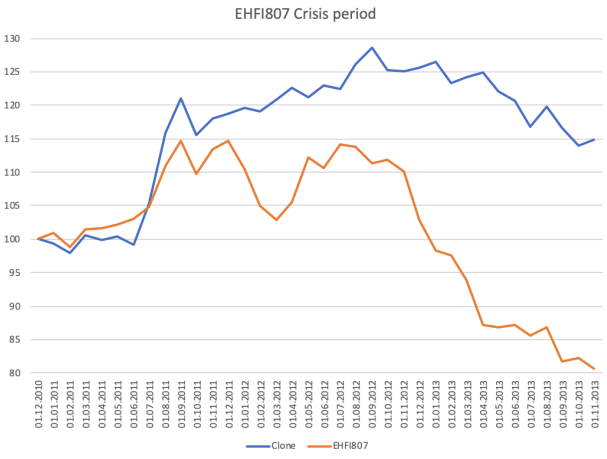
(c) Post-crisis period

Figure 30: Equity Market Neutral strategy

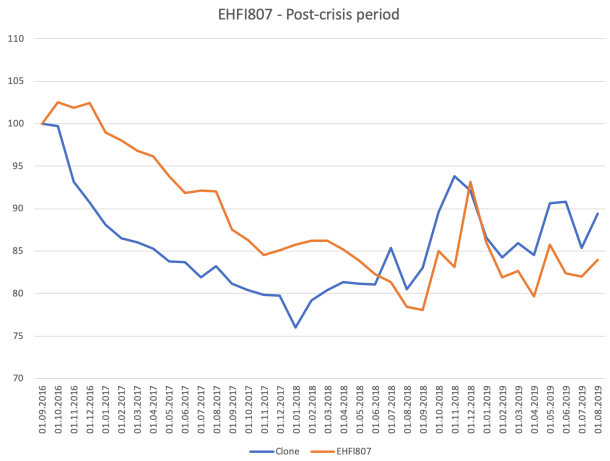
8.14 Equity Short Bias Index



(a) Pre-crisis period



(b) Crisis period



(c) Post-crisis period

Figure 31: Equity Short Bias strategy

8.15 Long Short Equity Crisis period QQ-plot

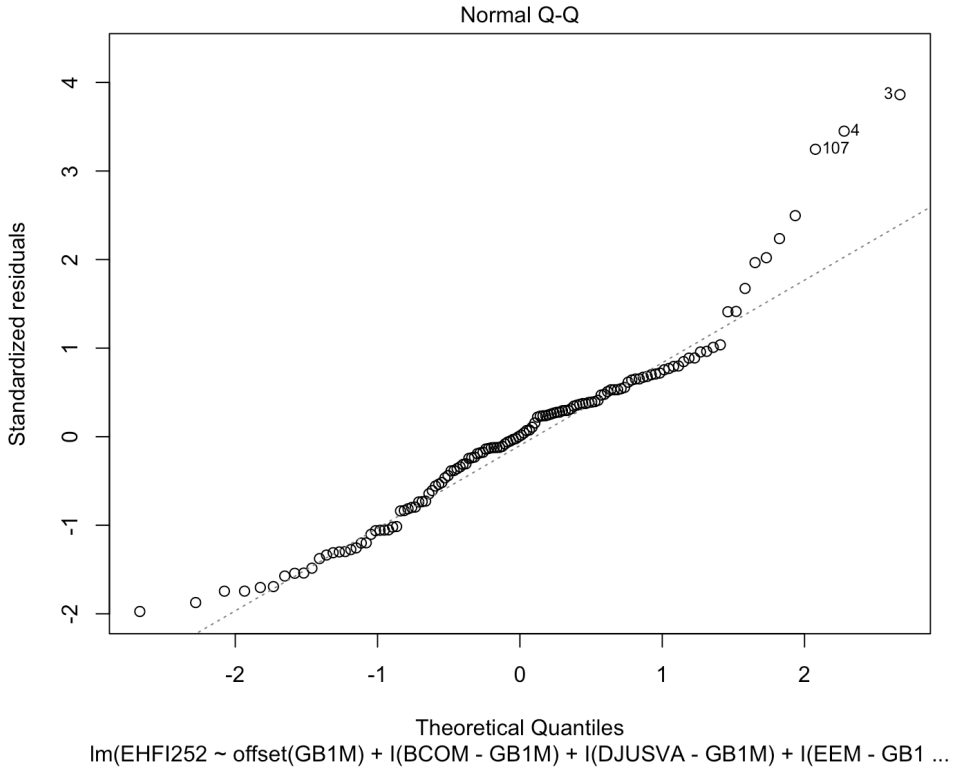
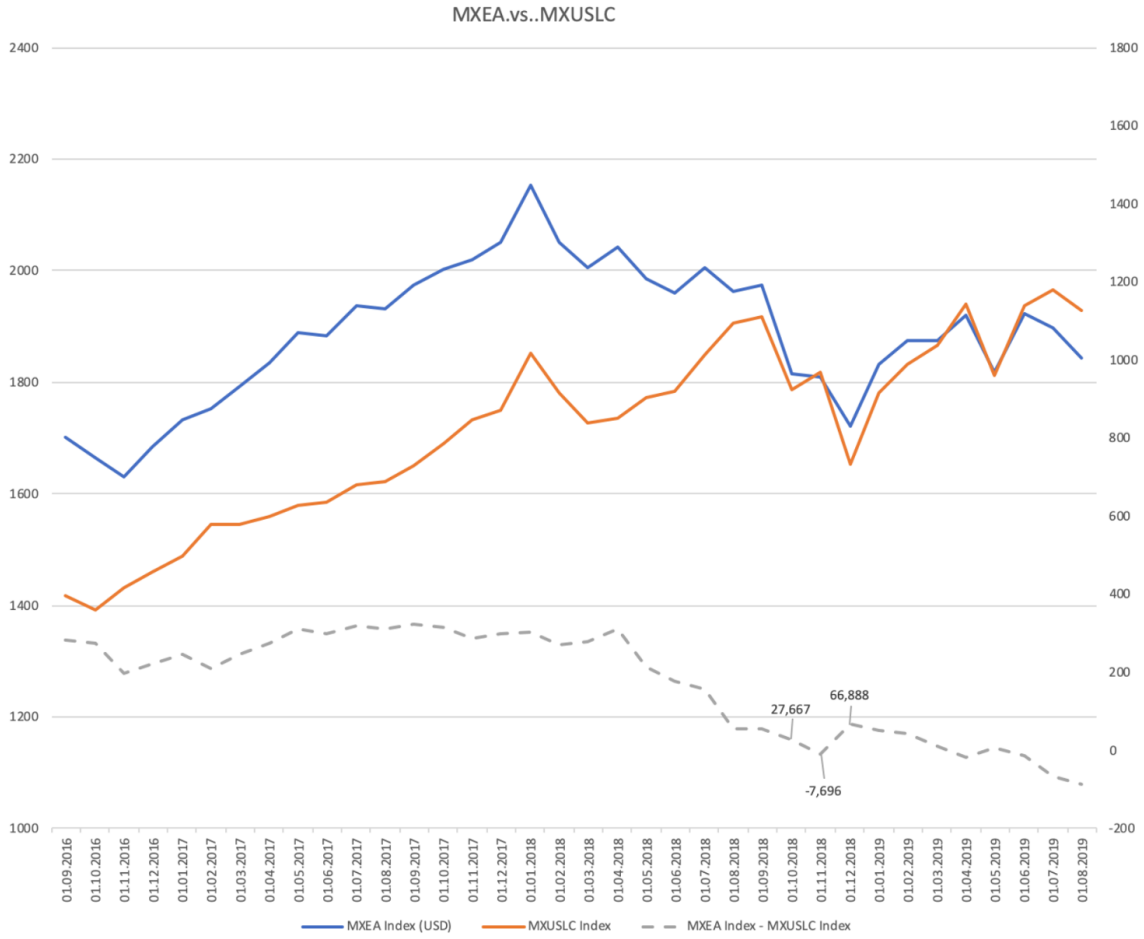


Figure 32: QQ-Plot Crisis period EHFI252

8.16 Spread between MXEA and MXUSLC



8.17 Event driven Pre-crisis period QQ-plot

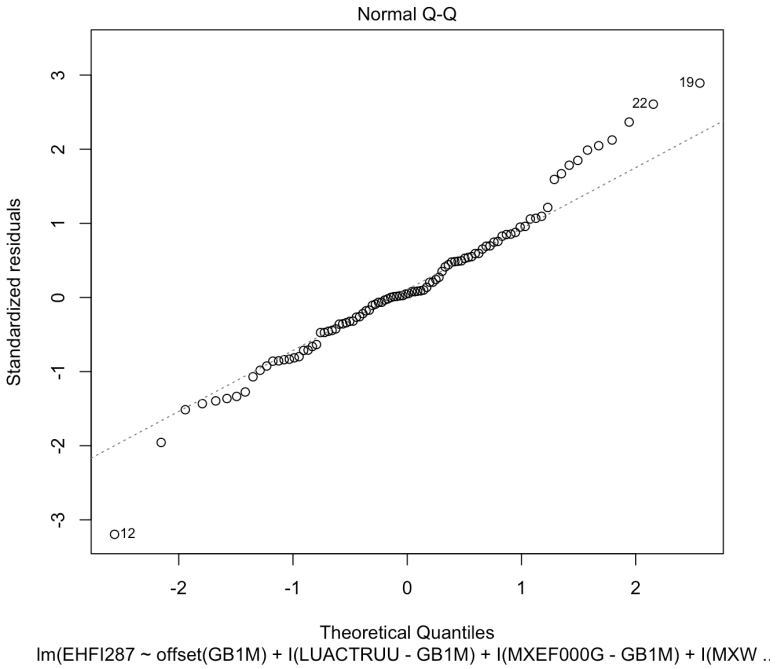


Figure 33: QQ-plot Pre-crisis period

8.18 Data Characteristics

Name	Period	Number of observations	Missing observations	Minimum	Maximum	Mean	Median	Standard Deviation	Average excess return	Sharpe Ratio
		Pre-crisis period Dec/99 to Nov/07								
BCOM Index (USD) - Last Price	BCOM	96	0	-0,0764	0,1007	0,0076	0,0058	0,0398	-0,0228	-0,5735
DHY US Equity (USD) - Last Price	DHY	96	0	-0,2018	0,2691	-0,0050	-0,0055	0,0652	-0,0354	-0,5427
DJUSGR Index (USD) - Last Price	DJUSGR	96	0	-0,1800	0,1441	-0,0029	0,0007	0,0592	-0,0333	-0,5620
DJUSVA Index (USD) - Last Price	DJUSVA	96	0	-0,1060	0,1217	0,0050	0,0093	0,0359	-0,0255	-0,7095
EEM US Equity (USD) - Last Price	EEM	56	40	-0,1114	0,1420	0,0282	0,0332	0,0536	-0,0023	-0,0420
GI1 Index (USD) - Last Price	GI1	96	0	-0,1602	0,1571	0,0132	0,0121	0,0617	-0,0172	-0,2785
LUACTRUU Index (USD) (R1)	LUACTRUU	96	0	-0,0434	0,0358	0,0055	0,0075	0,0132	-0,0250	-1,8917
MXEA Index (USD) - Last Price	MXEA	96	0	-0,1088	0,0938	0,0036	0,0073	0,0403	-0,0268	-0,6653
MXEF000G Index (USD) - Last Price	MXEF000G	96	0	-0,1653	0,1369	0,0105	0,0134	0,0620	-0,0199	-0,3212
MXEU000G Index (USD) - Last Price	MXEU000G	96	0	-0,1217	0,1162	0,0025	0,0005	0,0445	-0,0279	-0,6268
MXEU000V Index (USD) - Last Price	MXEU000V	96	0	-0,1647	0,1442	0,0066	0,0106	0,0478	-0,0238	-0,4984
MXWOLC	MXWOLC	96	0	-0,1115	0,0852	0,0011	0,0068	0,0389	-0,0294	-0,7548
MXWOSC	MXWOSC	96	0	-0,1329	0,1209	0,0085	0,0132	0,0466	-0,0219	-0,4704
SBBMGLU Index (USD) - Last Price	SBBMGLU	96	0	-0,1053	0,0838	0,0040	0,0097	0,0395	-0,0265	-0,6703
SCRTEM Index (USD) - Last Price	SCRTEM	96	0	-0,1670	0,1503	0,0112	0,0149	0,0595	-0,0193	-0,3238
SPBDU10T Index (USD) - Last Price	SPBDU10T	96	0	-0,0556	0,0406	0,0058	0,0062	0,0172	-0,0246	-1,4269
DJUSGR Index (USD) - DJUSVA Index (USD)	DJUSGR vs.DJUSVA	96	0	-5,1836	3,3554	-0,0609	0,0030	0,7754	-0,0913	-0,1178
MXUS000G Index - MXUS000V Index	MXUS000G vs. MXUS000V	96	0	-0,6825	0,6764	0,0098	-0,0020	0,2209	-0,0207	-0,0935
MXEA Index - MXUSLC Index	MXEA vs. MXUSLC	96	0	-0,1493	0,1557	0,0090	0,0132	0,0622	-0,0214	-0,3446
MXEULC Index - MXUSLC Index	MXEULC vs. MXUSLC	96	0	-1,6024	23,2433	0,3760	0,0242	2,4752	0,3456	0,1396
MXWOLC Index - MXWOSC Index	MXWOLC vs.MXWOSC	96	0	-0,1191	0,0930	-0,0004	0,0056	0,0393	-0,0309	-0,7851
USDEUR Curncy (USD) - Last Price	USDEUR	96	0	-0,0740	0,0535	-0,0036	-0,0037	0,0260	-0,0340	-1,3060
USDGBP Curncy (USD) - Last Price	USDGBP	96	0	-0,0509	0,0504	-0,0023	-0,0029	0,0214	-0,0327	-1,5313
USDJPY Curncy (USD) - Last Price	USDJPY	96	0	-0,0684	0,0764	0,0012	0,0030	0,0264	-0,0292	-1,1078
GB1M	GB1M	96	0	0,0000	0,0596	0,0304	0,0313	0,0171	0,0000	0,0000
EHFI251 (Eurekahedge hedge fund index)	EHFI251	96	0	-0,0139	0,0528	0,0104	0,0109	0,0130	-0,0200	-1,5420
EHFI252 (Eurekahedge Long Short Equity)	EHFI252	96	0	-0,0299	0,0895	0,0107	0,0118	0,0195	-0,0197	-1,0133
EHFI253 (Eurekahedge Marco)	EHFI253	96	0	-0,0182	0,0540	0,0097	0,0089	0,0129	-0,0207	-1,6003
EHFI254 (Eurekahedge Multi Strategy)	EHFI254	96	0	-0,0100	0,0400	0,0115	0,0119	0,0109	-0,0189	-1,7350
EHFI255 (Eurekahedge Relative Value)	EHFI255	96	0	-0,0182	0,0407	0,0101	0,0096	0,0109	-0,0203	-1,8607
EHFI285 (Eurekahedge Arbitrage)	EHFI285	96	0	-0,0093	0,0260	0,0075	0,0073	0,0062	-0,0229	-3,7073
EHFI287 (Eurekahedge Distressed Debt)	EHFI287	96	0	-0,0285	0,0493	0,0122	0,0126	0,0142	-0,0182	-1,2775
EHFI288 (Eurekahedge Event Driven)	EHFI288	96	0	-0,0287	0,0415	0,0102	0,0131	0,0146	-0,0202	-1,3829
EHFI289 (Eurekahedge Fixed Income)	EHFI289	96	0	-0,0069	0,0349	0,0080	0,0082	0,0073	-0,0225	-3,0777
EHFI751 (Eurekahedge Equity Market Neutral)	EHFI751	96	0	-0,0091	0,0398	0,0057	0,0057	0,0072	-0,0247	-3,4434
EHFI807 (Eurekahedge Equity Short Bias)	EHFI807	96	0	-0,0426	0,0609	0,0028	0,0031	0,0235	-0,0277	-1,1776

Figure 34: Pre-Crisis period

Period		Crisis period Dec/99 to Nov/10								
Name		Number of observations	Missing observations	Minimum	Maximum	Mean	Median	Standard Deviation	Average excess return	Sharpe Ratio
BCOM Index (USD) - Last Price	BCOM	132	0	-0,2134	0,1299	0,0048	0,0066	0,0497	-0,0187	-0,3773
DHY US Equity (USD) - Last Price	DHY	132	0	-0,2944	0,2691	-0,0030	-0,0021	0,0774	-0,0265	-0,3423
DJUSGR Index (USD) - Last Price	DJUSGR	132	0	-0,1850	0,1441	-0,0024	0,0042	0,0608	-0,0259	-0,4263
DJUSVA Index (USD) - Last Price	DJUSVA	132	0	-0,1638	0,1217	0,0020	0,0089	0,0451	-0,0215	-0,4760
EEM US Equity (USD) - Last Price	EEM	92	40	-0,2558	0,1686	0,0174	0,0225	0,0733	-0,0061	-0,0835
GI1 Index (USD) - Last Price	GI1	132	0	-0,2793	0,2116	0,0107	0,0133	0,0702	-0,0128	-0,1822
LUACTRUU Index (USD) (R1)	LUACTRUU	132	0	-0,0777	0,0680	0,0057	0,0071	0,0177	-0,0178	-1,0076
MXEA Index (USD) - Last Price	MXEA	132	0	-0,2024	0,1227	0,0003	0,0058	0,0520	-0,0232	-0,4455
MXEF000G Index (USD) - Last Price	MXEF000G	132	0	-0,2775	0,1594	0,0075	0,0091	0,0726	-0,0161	-0,2211
MXEU000G Index (USD) - Last Price	MXEU000G	132	0	-0,1985	0,1162	-0,0001	0,0005	0,0536	-0,0236	-0,4402
MXEU000V Index (USD) - Last Price	MXEU000V	132	0	-0,2250	0,1891	0,0014	0,0072	0,0633	-0,0221	-0,3495
MXWOLC	MXWOLC	132	0	-0,1826	0,1035	-0,0010	0,0046	0,0479	-0,0245	-0,5118
MXWOSC	MXWOSC	132	0	-0,2271	0,1630	0,0063	0,0115	0,0568	-0,0172	-0,3035
SBBMGLU Index (USD) - Last Price	SBBMGLU	132	0	-0,2061	0,1222	0,0017	0,0071	0,0503	-0,0218	-0,4346
SCRTEM Index (USD) - Last Price	SCRTEM	132	0	-0,2754	0,1954	0,0086	0,0116	0,0703	-0,0149	-0,2118
SPBDU10T Index (USD) - Last Price	SPBDU10T	132	0	-0,0556	0,0817	0,0061	0,0062	0,0194	-0,0174	-0,8956
DJUSGR Index (USD) - DJUSVA Index (USD)	DJUSGR vs.DJUSVA	132	0	-5,1836	3,3554	-0,0494	0,0001	0,6634	-0,0729	-0,1099
MXUS000G Index - MXUS000V Index	MXUS000G vs. MXUS000V	132	0	-0,6825	0,6764	0,0171	0,0191	0,2109	-0,0065	-0,0306
MXEA Index - MXUSLC Index	MXEA vs. MXUSLC	132	0	-0,2422	0,1655	0,0036	0,0085	0,0727	-0,0199	-0,2736
MXEULC Index - MXUSLC Index	MXEULC vs. MXUSLC	132	0	-1,6024	23,2433	0,2714	0,0184	2,1193	0,2479	0,1170
MXWOLC Index - MXWOSC Index	MXWOLC vs.MXWOSC	132	0	-0,1684	0,0930	-0,0027	0,0031	0,0472	-0,0262	-0,5565
USDEUR Curncy (USD) - Last Price	USDEUR	132	0	-0,0918	0,1070	-0,0014	-0,0012	0,0322	-0,0249	-0,7755
USDGBP Curncy (USD) - Last Price	USDGBP	132	0	-0,0864	0,1077	0,0006	0,0001	0,0265	-0,0229	-0,8635
USDJPY Curncy (USD) - Last Price	USDJPY	132	0	-0,0720	0,0852	-0,0011	0,0000	0,0290	-0,0246	-0,8495
GB1M	GB1M	132	0	0,0000	0,0596	0,0235	0,0171	0,0188	0,0000	0,0000
EHFI251 (Eurekahedge hedge fund index)	EHFI251	132	0	-0,0439	0,0528	0,0090	0,0106	0,0154	-0,0145	-0,9436
EHFI252 (Eurekahedge Long Short Equity)	EHFI252	132	0	-0,0662	0,0895	0,0086	0,0112	0,0226	-0,0149	-0,6578
EHFI253 (Eurekahedge Marco)	EHFI253	132	0	-0,0182	0,0540	0,0089	0,0086	0,0126	-0,0146	-1,1547
EHFI254 (Eurekahedge Multi Strategy)	EHFI254	132	0	-0,0530	0,0426	0,0097	0,0112	0,0137	-0,0138	-1,0100
EHFI255 (Eurekahedge Relative Value)	EHFI255	132	0	-0,0425	0,0407	0,0093	0,0091	0,0128	-0,0143	-1,1146
EHFI285 (Eurekahedge Arbitrage)	EHFI285	132	0	-0,0603	0,0297	0,0070	0,0075	0,0106	-0,0165	-1,5631
EHFI287 (Eurekahedge Distressed Debt)	EHFI287	132	0	-0,0943	0,0663	0,0105	0,0117	0,0211	-0,0131	-0,6181
EHFI288 (Eurekahedge Event Driven)	EHFI288	132	0	-0,0895	0,0665	0,0092	0,0132	0,0208	-0,0143	-0,6855
EHFI289 (Eurekahedge Fixed Income)	EHFI289	132	0	-0,0611	0,0349	0,0075	0,0083	0,0115	-0,0160	-1,3968
EHFI751 (Eurekahedge Equity Market Neutral)	EHFI751	132	0	-0,0254	0,0398	0,0051	0,0054	0,0075	-0,0184	-2,4737
EHFI807 (Eurekahedge Equity Short Bias)	EHFI807	132	0	-0,0714	0,0738	0,0029	0,0024	0,0263	-0,0206	-0,7842

Figure 35: Crisis period

Name	Period	Post-crisis period Dec/09 to Nov/16								
		Number of observations	Missing observations	Minimum	Maximum	Mean	Median	Standard Deviation	Average excess return	Sharpe Ratio
BCOM Index (USD) - Last Price	BCOM	81	0	-0,1474	0,1067	-0,0054	-0,0043	0,0448	-0,0060	-0,1343
DHY US Equity (USD) - Last Price	DHY	81	0	-0,0789	0,0922	-0,0003	0,0000	0,0350	-0,0009	-0,0256
DJUSGR Index (USD) - Last Price	DJUSGR	81	0	-0,0800	0,1190	0,0099	0,0081	0,0408	0,0092	0,2268
DJUSVA Index (USD) - Last Price	DJUSVA	81	0	-0,0823	0,1000	0,0083	0,0120	0,0360	0,0077	0,2134
EEM US Equity (USD) - Last Price	EEM	81	0	-0,1791	0,1630	0,0000	-0,0025	0,0571	-0,0006	-0,0102
GI1 Index (USD) - Last Price	GI1	81	0	-0,1431	0,1214	-0,0034	0,0051	0,0582	-0,0040	-0,0684
LUACTRUU Index (USD) (R1)	LUACTRUU	81	0	-0,0276	0,0303	0,0050	0,0036	0,0120	0,0044	0,3664
MXEA Index (USD) - Last Price	MXEA	81	0	-0,1209	0,0959	0,0019	-0,0021	0,0469	0,0013	0,0271
MXEF000G Index (USD) - Last Price	MXEF000G	81	0	-0,1556	0,1345	0,0017	-0,0028	0,0523	0,0011	0,0213
MXEU000G Index (USD) - Last Price	MXEU000G	81	0	-0,1234	0,1184	0,0039	0,0022	0,0479	0,0033	0,0691
MXEU000V Index (USD) - Last Price	MXEU000V	81	0	-0,1401	0,1414	-0,0008	-0,0067	0,0578	-0,0014	-0,0239
MXWOLC	MXWOLC	81	0	-0,0998	0,1010	0,0053	0,0047	0,0401	0,0047	0,1166
MXWOSC	MXWOSC	81	0	-0,1143	0,1175	0,0081	0,0086	0,0450	0,0075	0,1654
SBBMGLU Index (USD) - Last Price	SBBMGLU	81	0	-0,1011	0,1077	0,0053	0,0037	0,0415	0,0047	0,1128
SCRTEM Index (USD) - Last Price	SCRTEM	81	0	-0,1519	0,1226	0,0004	-0,0017	0,0530	-0,0003	-0,0048
SPBDU10T Index (USD) - Last Price	SPBDU10T	81	0	-0,0327	0,0476	0,0049	0,0004	0,0165	0,0043	0,2619
DJUSGR Index (USD) - DJUSVA Index (USD)	DJUSGR vs.DJUSVA	81	0	-0,5081	1,0433	0,0189	-0,0158	0,2279	0,0183	0,0802
MXUS000G Index - MXUS000V Index	MXUS000G vs. MXUS000V	81	0	-0,1325	0,2304	0,0170	0,0146	0,0652	0,0164	0,2514
MXEA Index - MXUSLC Index	MXEA vs. MXUSLC	81	0	-0,2127	0,1434	-0,0106	-0,0103	0,0820	-0,0113	-0,1372
MXEULC Index - MXUSLC Index	MXEULC vs. MXUSLC	81	0	-21,1349	7,9780	-0,2275	0,0244	2,6061	-0,2281	-0,0875
MXWOLC Index - MXWOSC Index	MXWOLC vs.MXWOSC	81	0	-0,1007	0,0944	0,0042	0,0043	0,0396	0,0036	0,0911
USDEUR Curncy (USD) - Last Price	USDEUR	81	0	-0,0701	0,0804	0,0035	0,0017	0,0296	0,0029	0,0980
USDGBP Curncy (USD) - Last Price	USDGBP	81	0	-0,0475	0,0879	0,0029	0,0038	0,0249	0,0022	0,0905
USDJPY Curncy (USD) - Last Price	USDJPY	81	0	-0,0698	0,0640	0,0017	0,0009	0,0277	0,0011	0,0386
GB1M	GB1M	81	0	-0,0002	0,0026	0,0006	0,0004	0,0007	0,0000	0,0000
EHFI251 (Eurekahedge hedge fund index)	EHFI251	81	0	-0,0253	0,0313	0,0045	0,0047	0,0114	0,0039	0,3392
EHFI252 (Eurekahedge Long Short Equity)	EHFI252	81	0	-0,0471	0,0412	0,0047	0,0064	0,0181	0,0041	0,2263
EHFI253 (Eurekahedge Marco)	EHFI253	81	0	-0,0163	0,0226	0,0033	0,0046	0,0075	0,0027	0,3563
EHFI254 (Eurekahedge Multi Strategy)	EHFI254	81	0	-0,0256	0,0254	0,0044	0,0045	0,0109	0,0038	0,3493
EHFI255 (Eurekahedge Relative Value)	EHFI255	81	0	-0,0269	0,0279	0,0051	0,0059	0,0098	0,0044	0,4512
EHFI285 (Eurekahedge Arbitrage)	EHFI285	81	0	-0,0147	0,0170	0,0044	0,0054	0,0067	0,0038	0,5631
EHFI287 (Eurekahedge Distressed Debt)	EHFI287	81	0	-0,0348	0,0663	0,0068	0,0090	0,0156	0,0062	0,3981
EHFI288 (Eurekahedge Event Driven)	EHFI288	81	0	-0,0412	0,0361	0,0055	0,0070	0,0168	0,0049	0,2893
EHFI289 (Eurekahedge Fixed Income)	EHFI289	81	0	-0,0143	0,0275	0,0054	0,0059	0,0084	0,0048	0,5702
EHFI751 (Eurekahedge Equity Market Neutral)	EHFI751	81	0	-0,0114	0,0124	0,0034	0,0043	0,0051	0,0028	0,5440
EHFI807 (Eurekahedge Equity Short Bias)	EHFI807	81	0	-0,0742	0,1047	-0,0026	-0,0013	0,0325	-0,0032	-0,0987

Figure 36: Post-Crisis period

8.19 R-studio code sample for Rolling Window & Extended Rolling Window

```
# Rolling Window
# Estmating beta's
betas<-data.frame(matrix(nrow=36,ncol=6))
colnames(betas)<-c("LUACTRUU","MXEF000G","MXWOSC","SCRTEM","SPBDU10T","USDEUR")
i=1
for (i in 1:36) {
  newdataset<-`RWIIn&Outsample1287`[i:(i+35),]
  print(i)
  fit<-lm(EHFI287~offset(GB1M)+I(LUACTRUU-GB1M)+I(MXEF000G-GB1M)+I(MXWOSC-GB1M)+I(SCRTEM-GB1M)+I(SPBDU10T-GB1M)+I(USDEUR-GB1M)-1,newdataset)
  print(summary(fit))

  j=1
  for (j in 1:6){
    a<-summary(fit)$coefficients[j,1]
    print(a)
    betas[i,j]<-a
  }
  betas
}
```

Figure 37: Rolling Window

```
m=30 # Update after every single run
RW2ndNew252<-RWInsample3252[m:(m+35),]
fit23<-step(lm(EHFI252~offset(GB1M)+I(BCOM-GB1M)+I(DHY-GB1M)+
                 I(DJUSGR-GB1M)+I(DJUSVA-GB1M)+I(EEM-GB1M)+I(GI1-GB1M)+
                 I(LUACTRUU-GB1M)+I(MXEA-GB1M)+I(MXEF000G-GB1M)+
                 I(MXEU000G-GB1M)+I(MXEU000V-GB1M)+I(MXWOLC-GB1M)+
                 I(MXWOSC-GB1M)+I(SBBMGLU-GB1M)+I(SCRTEM-GB1M)+
                 I(SPBDU10T-GB1M)+I(DJUSGR.vs.DJUSVA-GB1M)+
                 I(MXUS000G.vs..MXUS000V-GB1M)+I(MXEA.vs..MXUSLC-GB1M)+
                 I(MXEULC.vs..MXUSLC-GB1M)+I(MXWOLC.vs.MXWOSC-GB1M)+
                 I(USDEUR-GB1M)+I(USDGBP-GB1M)+I(USDJPY-GB1M)-1,
               data=RW2ndNew252), direction="backward",k=3.841459)
print(summary(fit23))
#Removing +5% non significant factors step by step
summary(lm(EHFI252~offset(GB1M)+I(BCOM-GB1M)+I(DHY-GB1M)+I(EEM-GB1M)+
             I(GI1-GB1M)+I(MXEA-GB1M)+I(MXEU000G-GB1M)+I(MXEU000V-GB1M)+
             I(MXWOLC-GB1M)+I(MXWOSC-GB1M)+I(SBBMGLU-GB1M)+I(SCRTEM-GB1M)+
             I(SPBDU10T-GB1M)+I(MXUS000G.vs..MXUS000V-GB1M)+
             I(MXEA.vs..MXUSLC-GB1M)+I(MXEULC.vs..MXUSLC-GB1M)+
             I(USDEUR-GB1M)+I(USDGBP-GB1M)-1,data=RW2ndNew253))
#Test residual for autocorrelation
Box.test(resid(fit23),type="Ljung",lag=1,fitdf=1)
```

Figure 38: Sample of Extended Rolling Window