CIRJE-F-592

Hedge Fund Replication

Akihiko Takahashi
University of Tokyo

Kyo Yamamoto
Graduate School of Economics, University of Tokyo

September 2008; Revised in November 2008

Hedge Fund Replication*

Akihiko Takahashi and Kyo Yamamoto†
University of Tokyo

November 28, 2008

JEL Classification: G11, G20, G23

Keywords: Hedge Funds, Alternative Investments, Hedge Fund Clones

Forthcoming in The Recent Trend of Hedge Fund Strategies

Abstract

This chapter provides a comprehensive explanation of hedge fund replication. This chapter first reviews the characteristics of hedge fund returns. Then, the emergence of hedge fund replication products is discussed. Hedge fund replication methods are classified into three categories: Rule-based, Factor-based, and Distribution replicating approaches. These approaches attempt to capture different aspects of hedge fund returns. This chapter explains the three methods.

1 Introduction

The established cycles of the manufacturing industry has eventually made its way to the hedge fund industry. In the manufacturing industry, many companies compete to develop new products protected by patent. This is often a very expensive process involving large amounts of money invested to develop sophisticated proprietary technology. Rival companies release similar products and the prices of these products fall gradually. Eventually, the previously sophisticated technology becomes common, allowing for the products to be produced at low cost. This in turn leads to a further drop of prices. Finally, the products are mass-produced. At this stage, the companies are unable to earn sufficient profits, and look to develop new sophisticated products. This results in an endless cycle. However, some of these mass-produced products also survive due to low costs.

Hedge funds provide investment strategy as a product. Fund managers are money management professionals, and receive rewards by providing the strategy to investors.

*We thank Mr. Hideki Yamauchi, Mr. Takahiko Suenaga, Mr. Tetsuya Aoki, and Mr. Takeshi Hakamada in GCI Group for their constant support. We are greatful to Mr. Nandakumar Lokanathan and Mr. Allen Zheng in GCI Group for extensive help in improving English in this version.

†E-mail address: ee077031@mail.ecc.u-tokyo.ac.jp

They make efforts to sophisticate their own investment strategies. Their strategies are not protected by patent, but the prevailing industry structure enables their strategies to be kept confidential. The hedge fund industry has become severely competitive with new funds launched one after another, while many others collapse. In spite of this, the fees of investing into hedge funds are still at high levels and have not fallen down. For a typical case, investors need to pay 2% as management fee and 20% as incentive fee. This seems peculiar, when compared with the manufacturing industry. The reason for this lies in the difficulty of evaluating hedge fund performances. All fund managers believe that the high fees charged are justifiable as they are capable of earning high returns through their strategy. Investors in return invest in the fund because they believe that the fund manager will provide returns that will justify the high cost. To evaluate a hedge fund's performance, at leaset a few years of performance track record is needed. However, even if the track record is outstanding, there is no consensus on the persistence of a fund manager's skill. It is this difficulty in evaluating the performance of hedge funds that has kept fees unusually high despite the highly competitive nature of the industry.

In recent years, investment banks and investment companies have released hedge fund replication products, which provide investors access to hedge fund returns at lower costs. In addition, these products avoid some shortcomings of hedge funds that will be discussed later. Some replication products mimic a simple trading strategy of hedge funds while others attempt to infer the actual investment positions of hedge funds and take similar positions. Given the highly competitive nature of the hedge funds industry, the emergence of hedge fund clones can be considered a natural progression that has been long overdue when compared to the business cycle of the manufacturing industry.

The development of such techniques has proven to be a challenging task. Currently, the biggest banks such as Goldman Sachs, Merrill Lynch, and JP Morgan, and some large investment companies such as, Partners Group, have launched such products. (See, for example, [15].) Some of these institutions developed cloning technique collaborating with the pioneers in hedge fund research such as William Fung, David Hsieh, and Narayan Naik. (See also [15].) They and other researchers have proposed various methods, but these techniques are still work-in-progress. This chapter explores the methods of hedge fund replication in detail.

In the following section, we review the empirical characteristics of hedge fund returns, and reconfirm the properties of replication targets. Section 3 describes how and why hedge fund clones emerged and classifies replication methodologies into three distinct approaches. Sections 4 till 6 are devoted to explaining in detail the three approaches. Section 7 concludes the chapter.

2 Empirical characteristics of hedge fund returns

This section reviews the empirical characteristics of hedge fund returns before the aim and methodologies of hedge fund replication products are discussed. Originally, hedge fund managers seek absolute returns, i.e. they try to earn profits under any market circumstances. These managers make use of their unique insights into financial markets to find investment opportunities and extract returns. Since managers' abilities and insights are unique, employed financial instruments and investment strategies exhibit a wide variety. Some managers invest in stocks, bonds, or currencies, and others invest in complex structured products. Common investment strategies include long/short equities, event driven, global macro, relative value arbitrage, and so on. Due to the variety of investment tools and strategies, hedge funds exhibit a spectrum of risk-return profiles. Large investors, such as university endowments and pension funds, determine asset allocations to various asset classes. Recently, hedge funds have been included as one of the asset classes. Therefore, those investors also need to understand the return characteristics of hedge funds in broad terms.

Although investment instruments and strategies that hedge funds employ are wide-ranging, investment styles that capture absolute returns can be classified into two categories. The first is dynamically changing exposure to markets. Some managers predict market directions and take long or short positions accordingly. They realize profits when the markets move in line with their predictions. These managers are referred to as directional traders.

The second is exploiting market inefficiencies, i.e. mispriced securities. There are mainly two approaches that exploit market inefficiency: These are relative value and security selection. Relative value is finding two or more securities or portfolios that have the same value but are priced differently in markets. Relative value traders create positions when there is mispricing and close positions when the two market values converge. Security selection is finding undervalued or overvalued securities, taking long or short positions respectively and closing these positions when the manager judges the prices to have reached fair value.

There is much literature that has tried to define the characteristics of hedge fund returns. Their challenge has been extracting unique properties of hedge fund returns from cross-sectional and time-series data. These works can be classified into two main approaches that help to explain hedge fund returns. The first approach is through analyzing distribution and time-series properties of hedge fund returns. The second approach is through identifying factors and understanding how these factors drive hedge fund returns. This approach is also commonly referred to as style analysis.

The most outstanding find from returns distribution analysis is the non-normality characteristic of hedge fund returns. This is documented by [23], in which the authors calculated skew and kurtosis of hedge fund returns by strategy and tested for normality of returns. The result showed most strategies to exhibit negative skew and high kurtosis, and hence, their hypotheses of normality were rejected. The sign of skew is also dependent on data sample. [23] used TASS database, which covers mainly US and European hedge funds and reported negative skew of returns. [17] analyzed the return of hedge funds in Eurekahedge Asia-Pacific database and showed positive skew of returns. While there exists inconsistency in the signs of skew, it is recognized that the returns of hedge funds exhibit skew and fat-tails, leading to non-normality. Time-series analysis has been carried out by [16]. Their research showed the returns of hedge funds that hold illiquid assets to exhibit positive autocorrelation.

For further analysis, we compared the performance of Asian hedge funds against stock indices. AsiaHedge Indices (AsiaHedge composite, Asia excluding Japan, Japan Long/Short index) were used for this purpose. The panel A, B, and C of figure 1 graphs asset value growth of the respective indices against relevant stock indices. The reason why stock indices are used is that most hedge funds in Asia-Pacific region employ equity-oriented strategies as noted in [17].

The graphs show that hedge funds returns closely match stock returns in up markets but decline less that stock returns in down markets. This in turn leads to higher returns than stocks in the long run. Figure 2 shows statistics of monthly excess returns of hedge fund and stock indices. Mean excess returns of hedge funds are higher than stock indices with lower standard deviations. They exhibit positive skew, fat-tails and positive correlation with the stock indices (0.73-0.84). Their serial correlations, which are measured by AR (1), is not high.

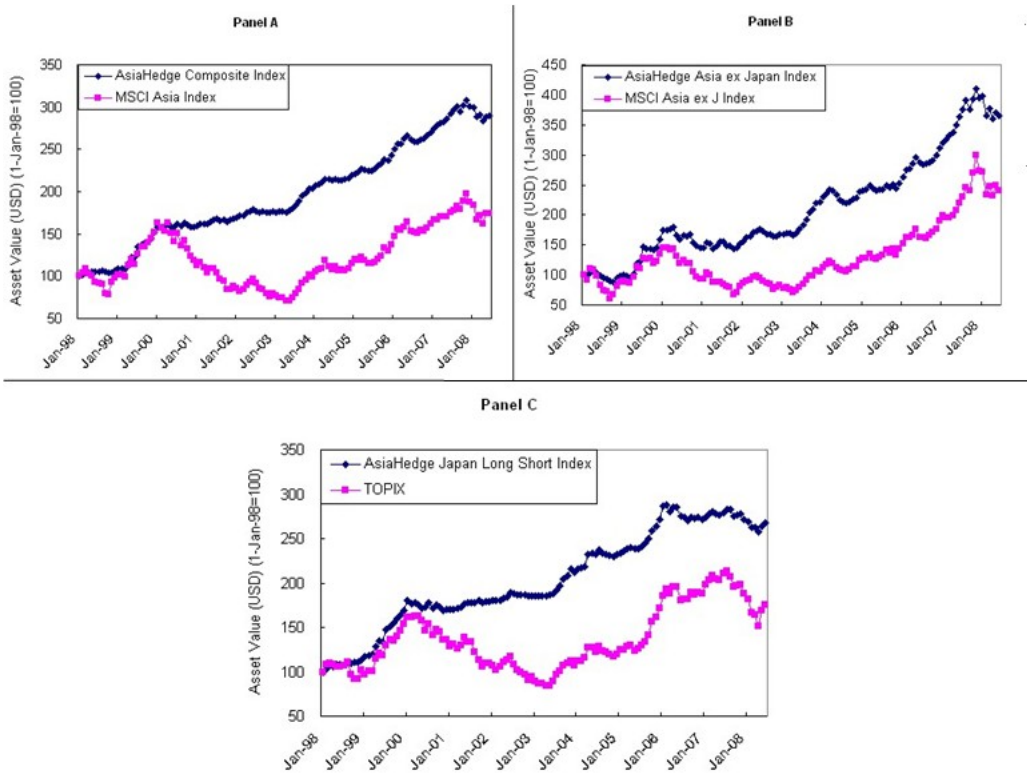

Figure 1. Asset value growth of AsiaHedge indices [Data source: Bloomberg and homepage of AsiaHedge (http://www.asiahedge.net/)]

Style analysis decomposes hedge fund returns into some common risk factors. The most fundamental and heavily-used model is the linear factor model;

$$R_i = \alpha_i + \sum_k \beta_{ik} F_k, \tag{1}$$

where R_i is the return of fund i, F_k is the return of factor k, β_{ik} is the exposure of fund

	AsiaHedge Composite Index	MSCI Asian Index	AsiaHedge Asia ex-J Index	MSCI Asia ex-J Index	AsiaHedge Japan L/S Index	TOPIX
Mean	0.54 %	0.25 %	0.79 %	0.63 %	0.49 %	0.24 %
Std. dev.	1.73 %	5.19 %	4.02 %	7.10 %	2.09 %	4.84 %
Skew	1.11	0.10	0.94	0.09	1.12	0.06
Kurtosis	8.57	3.13	7.09	3.42	6.47	2.92
Max	9.29 %	17.30 %	18.81 %	21.10 %	9.96 %	13.12 %
Min	-4.11 %	-12.55 %	-8.82 %	-16.15 %	-3.88 %	-12.37 %
AR (1)	0.19	0.13	0.21	0.16	0.18	0.15
Correlation with the stock index	0.76		0.84		0.73	

Figure 2. Monthly statistics of excess returns of hedge fund and stock indices (Jan/98-May/08) [Data source: Bloomberg and homepage of AsiaHedge (http://www.asiahedge.net/)]

i to factor k, and α_i is the rest of the return R_i. In considering the portfolio of hedge funds, the risk is captured by the factors and other components as follows.

$$\sum_{i=1}^{n} w_i R_i = \sum_{i=1}^{n} \alpha_i + \sum_{k} (w_1 \beta_{1k} + \cdots + w_n \beta_{nk}) F_k,$$

where w_i is the weight on fund i.

Before proceeding to explore the empirical results of style analysis for hedge funds, we review the results of such analysis done on mutual funds by [30]. In that research, the returns of mutual funds were decomposed by linear regressions into twelve asset class factors: Bills, Intermediate-term government bonds, Long-term government bonds, Corporate bonds, Mortgage-related securities, Large-capitalization value stocks, Large-capitalization growth stocks, Medium-capitalization stocks, Small-capitalization stocks, Non-US bonds, European stocks, and Japanese stocks. The research showed that a large part of mutual fund returns are explained by the twelve asset class factors. R^2 levels were at 60%-97%.

[7] implemented style analysis for hedge funds for the first time. First, they performed Sharpe's style regression on mutual fund and hedge fund returns. While more than half the mutual funds had R^2s above 75%, nearly half (48%) of the hedge funds had R^2s below 25%. The reasons why Sharpe's style regression does not work for hedge funds are due to their dynamic trading strategies (changing exposures to asset class factors) and spread trading strategies (long/short strategies). In order to find style factors unique to hedge funds, the literature extracted mutually orthogonal principal components, and constructed style factors whose returns were highly correlated to the principal components. It showed that the style factors related non-linearly to the underlying market factors.

Some other research papers reported non-linear relations of specific investment strategies to the market. For example, [9] studied trend-following funds. A trend-follower tries to capture market trends. A trend is a time-series of asset prices that move persistently in one direction over a given time interval, where price changes exhibit positive serial correlation. A trend-follower attempts to identify developing price patterns with this property and trade in the direction of the trend when this occurs. Figure 3 describes the non-linear relationship between trend-following funds and the equity market. Morgan Stanley (MS) World Equity Index is used as a proxy for the world equity market and its returns are sorted into five states. State 1 consists of the worst months, and State 5 the best months. This figure graphs the average monthly returns of an equally weighted portfolio of the six largest trend-following funds, along with that of the world equity markets, for each state. Trend-follower funds earn high returns in the worst state and the best state, and low returns in normal states.

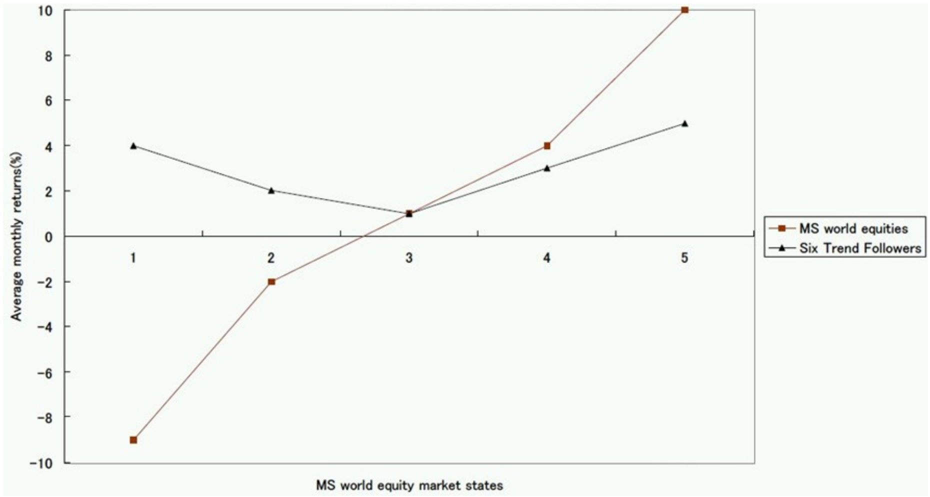

Figure 3. Average monthly returns of six large trend-following funds in five different MS world equity market states [The authors created based on [9].]

Another strategy with non-linear characteristics is active risk arbitrage (merger arbitrage) as reported by [24]. After the announcement of a merger or acquisition, the target company's stock typically trades at a discount to the price offered by the acquiring company. The difference between the target's stock price and the offer price is known as the arbitrage spread. Risk arbitrage refers to an investment strategy that attempts to profit from this spread. If the merger is successful, the arbitrageur captures the arbitrage spread. However, if the merger fails, the arbitrageur incurs a loss, usually much greater than the profits obtained if the deal succeeds. As mentioned in the article, risk arbitrage funds earn profits in both flat and appreciating markets, but suffer large losses in depreciating markets.

As just described, the strategies that employ dynamic trading or bet on corporate events induce the non-linear relationships with the markets. Poor or non-linear relationships with the markets make it difficult to find style factors for hedge funds. Hedge fund researchers navigate this challenging environment to discover style factors by mainly following two methodologies.

The first method is extracting factors from data of hedge fund returns using a statistical method. [7] found style factors by principal component analysis, and [3] by cluster analysis. Using these statistical methods, we can find style factors that explain large parts of hedge fund returns. These style factors can be considered orthogonal or remote with each other.

The second method is constructing style factors based on asset classes like [30]. These factors are referred to as asset based style (ABS) factors. ABS factors can be observed in markets. To be able to find ABS factors that drive hedge fund returns, and to identify the exposures to these factors, allow for the risk of each hedge fund or the portfolio of hedge funds to be effectively monitored. This is useful information for the portfolio construction and rebalancing of hedge funds. More aggressively, this allows for the prediction of fund returns by having forward looking views on these risk factors. This also allows for risk management by controlling exposures to these risk factors when they or securities highly correlated to them are tradable. Moreover, in theory, we can replicate parts of the returns that can be explained by tradable risk factors, or extract excellent alpha by hedging the exposures to these risk factors. For these purposes, ABS factors have been researched more than statistical approaches.

For style analysis of hedge funds, we need the factors that capture arbitrage and dynamic trading strategies. The representative examples of factors that catch spread trading are small-minus-big (SMB), high-minus-low (HML), and credit spread factors. SMB is the difference of returns between small and big cap stocks. Therefore, it represents a strategy that longs small stocks and shorts large stocks. This is a typical long/short strategy that captures the small stock premium. HML is the difference of returns between high book-to-market ratio stocks minus low book-to-market ratio stocks. High and low book-to-market ratio stocks are referred to as value and growth stocks, respectively. HML is a return of the typical strategy that longs value stocks and shorts growth stocks to realize the value premium. Credit spread represents a return of the strategy that buys corporate bond and sells government bonds. It can be inferred that above three factors explain some parts of hedge fund returns. However, there still remains the question of what additional factors can effectively explain arbitrage and dynamic trading strategies non-linear characteristics with the markets.

The terms of non-linearity and dynamic trading strategies remind us of derivative products. Figure 4 shows the profit/loss of contracts of European options at the maturity date. These P/Ls exhibit non-linear relationships to the underlying asset. Moreover, these P/Ls can be replicated using dynamic replication strategies. Figures 3 and 4 show the returns of trend-following funds to resemble the P/Ls of a straddle contract. In addition, a strategy that shorts put options on the stock market earns profits constantly in flat or appreciating markets, and incurs losses in depreciating markets. This P/L characteristic resembles the returns of risk arbitrage funds. From the above examples, we can assume that these derivative products can be used as factors for dynamic trading strategies even if the funds do not trade the products in practice. [1] analyzed risk exposures of equity-oriented hedge funds categorized by investment strategies: Event arbitrage, Restructuring, Event driven, Relative value arbitrage, Convertible arbitrage, Equity hedge, Equity non-hedge, and Short selling. They used SMB, HML, momentum, credit spread, the return of rolling ATM and OTM European call and put options

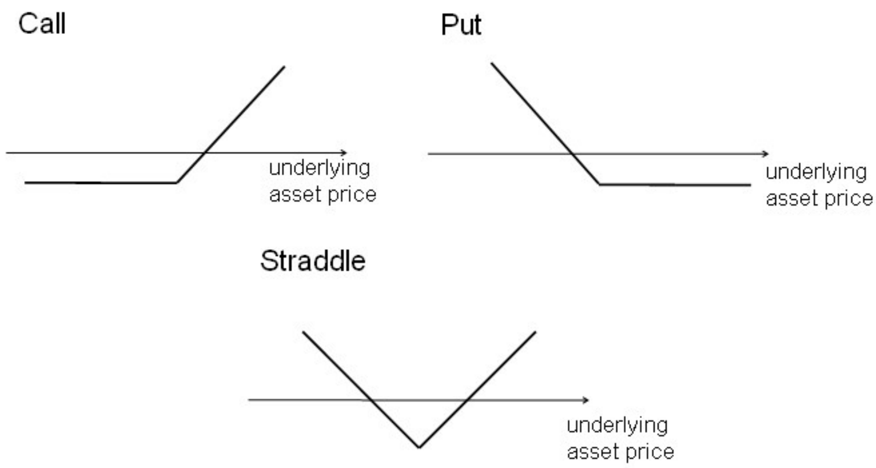

Figure 4. The Profit/Loss of contracts of European options at the maturity date

on S&P 500 composite index in addition to asset class indices as ABS factors. The research showed that some parts of returns for all of the strategies were explained by some spread factors. In addition, option factors were significant for six out of the eight strategies. Some returns of event arbitrage, restructuring, event driven, and relative value arbitrage were explained by a shorting strategy of OTM put options on S&P 500 composite index. Convertible arbitrage and short selling were significantly explained by a shorting strategy of ATM put options and OTM call options, respectively. The R^2s of the regressions ranged between 40-92%, which was higher than the results of [7] using Sharpe's asset class factor model. [9] introduced an interesting ABS factor, a lookback straddle, that captures the returns of trend following funds in addition to a straddle of European options. Lookback straddle is a product that gives the difference between the maximum and minimum prices of an underlying asset within the contract period at maturity date. Figure 5 shows the delta of lookback straddle that replicates the payoff. The parameter setting is same with [9]. Panel A describes the delta in the scenario that the underlying asset price rose to 130, fell down to 115, and rose to 145. Panel B graphs the scenario that the underlying asset price rose to 130, and fell back to 100. In both scenarios, the deltas of lookback straddle tend to increase and decrease in line with asset prices. And since it seems logical to assume that trend followers also change their positions as mentioned above, the rolling strategy of lookback straddle can be a powerful ABS factor in explaining the returns of trend following funds.

As just described, ABS factors, such as options, explain some parts of the return of dynamic trading and arbitrage strategies. Hedge fund exposures to traditional factors such as stocks and bonds are called traditional beta. Those to ABS factors such as long/short and option factors are called alternative beta. The remainder is referred alternative alpha. In many cases, the risk and performance of hedge funds are analyzed through decomposing the returns into the three above mentioned parts. [1], [11]

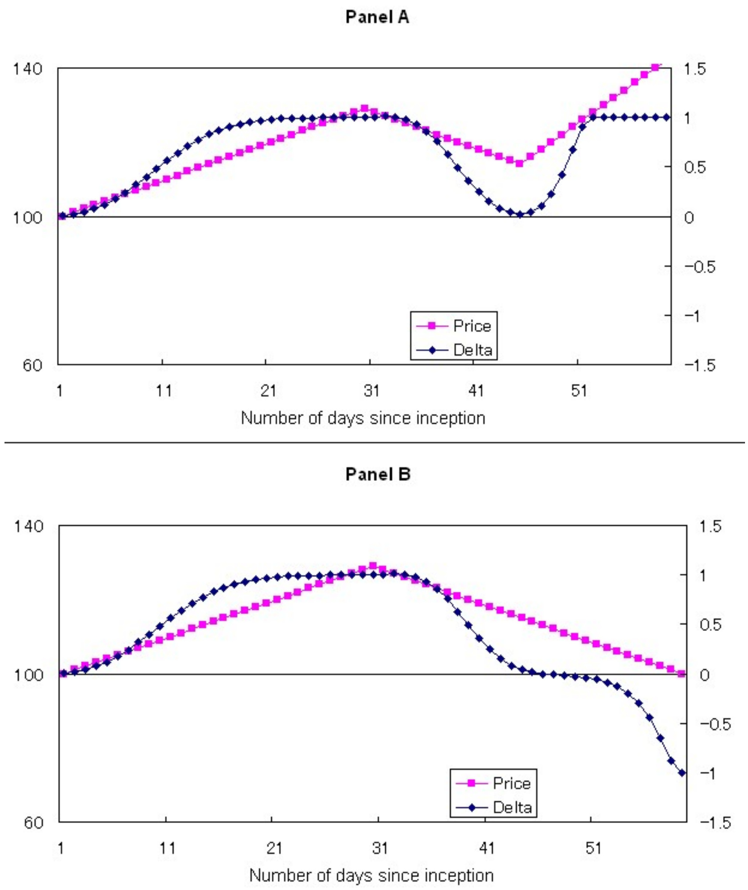

Figure 5. Delta of lookback straddle [The authors created based on [9].]

and [13] show that large parts of hedge fund returns are explained by traditional and alternative beta on an aggregate basis.

3 Emergence of hedge fund replication products and the approaches

Hedge fund investing has been common for large investors, such as university endowments and institutional investors. Although there is no consensus on the exact reason for investing into hedge funds, we have identified two main possibilities: High absolute return and Low correlation to other asset classes. As described in the previous section, hedge fund returns can be decomposed into risk factors and alternative alpha. Therefore, we can consider hedge funds to be efficient portfolios comprised of risk factors and alternative alpha. With this reasoning, we can determine that the actual worth of hedge funds is their accessibility to alternative risk premium, their ability to control exposures to risk factors, and their accessibility to alternative alpha. We define alternative risk premium as risk premium that "buy and hold" asset class strategies cannot access. Combining these three, a skilled fund manager can deliver high absolute returns and low correlation to other asset classes. However, there are some demerits of hedge fund investing. The most common ones are high costs, low transparency, and low liquidity. These are structural problems of hedge funds. Moreover, low transparency adds to manager selection risk. In the worst case, investors end up selecting a fraudulent fund manager. These structural problems make the process for fund selection, due diligence, and risk monitoring very tough. There is also the issue of capacity constraint. Some hedge fund managers set limits on the fund size as they have difficulty managing large amounts of assets. In fact, [14] and [25] clarified that capital inflows attenuated the alpha generating ability of some hedge funds and funds of hedge funds. As a result, these funds do not accept new investments upon reaching their limits. Questions are being raised whether hedge funds deliver returns above and beyond these shortcomings. Moreover, slumps and scandals within the hedge fund industry accelerate flight to quality.

In recent years, asset management firms and investment banks have provided hedge fund replication products. They aim to deliver hedge fund returns without the demerits of hedge funds. They invest in transparent and liquid assets, such as index futures, ETFs, currencies, and offer the products at low costs. They disclose the replication strategy and composition of the portfolio. The use of highly liquid assets also reduces capacity constraints and allows for quick liquidation. Since it is hardly possible to exactly replicate hedge fund returns month-to-month, there are some approaches, which aim to replicate the different characteristics of hedge fund returns instead. As mentioned previously, the worth of hedge funds can be separated into three categories: Accessibility to alternative risk premium, Control over exposures to risk factors, and Alternative alpha.

The first of these approaches is a rule-based approach that aims to replicate hedge fund accessibility to alternative risk premium. Hedge funds can take exposures to risk factors, which a "buy and hold" strategy of traditional asset classes cannot access, by employing long/short, dynamic trading, or relative value strategies. Rule-based hedge fund clones mimic typical trading strategies that hedge funds employ. For example, Merrill Lynch equity volatility arbitrage index mimics the volatility arbitrage strategy based on some simple rules. (See [2].)

The second approach is factor-based replication that seeks to replicate hedge fund accessibility to alternative risk premium and control exposures to risk factors. In other words, they try to deliver both traditional and alternative beta of hedge funds returns. This approach needs three steps. The first step is to find ABS factors that can explain hedge fund returns. They are for example, market indices, spread factors, and option trading and hedging strategies. Moreover, rule-based clones themselves can be considered as ABS factors. The second step is to estimate the factor exposure of the replication target. This procedure is based on statistical techniques, such as regression and filtering. The final step is creating the estimated factor exposures using tradable instruments.

The third approach is distribution replication that aims to replicate the distribution of hedge fund returns. This approach differs from factor-based clones that try to fit their returns to the target hedge fund returns on a month-to-month basis. These distribution replicating clones generate returns sampled from the same distribution as target hedge fund returns. Moreover, the clones replicate correlation structure between hedge funds and investors' portfolio. This is ideal for investors who look to benefit from the low correlation properties of hedge funds.

The above mentioned methodologies are still work-in-progress. The following sections discuss these replication approaches in greater detail.

4 Rule-based approach

As stated in the previous section, the rule based approach mimics typical hedge fund investment strategies, which access alternative risk premium. The recent growth in ETFs allows for us to take exposures to hedge fund style factors that were previously inaccessible. Through these ETFs, we can try to mimic the equity long/short strategies of hedge funds. For example, we can create a SMB factor by longing a small cap stock index and shorting a large cap stock index, and a HML factor by longing a value stock index and shorting a growth stock index. ETFs that track these indices are mainly listed in US and Europe. These are typical equity-oriented strategies that extract small and value stock premiums.

Dynamic trading strategies can be replicated by using listed index options. If an index option is not listed, we can replicate its payoff through a delta-hedging strategy of the underlying asset. For example, simple trend following strategies can be replicated through the delta of a lookback straddle, which is described in Figure 5. Currently, such rule-based clone methodologies used in practice are proprietary. For example, Merrill Lynch provides its equity volatility arbitrage index. The index exploits the market anomaly present in the volatility of S&P 500 index. The high demand for S&P index options has historically led to higher prices, which in turn leads to high implied volatilities. As a result, the implied volatility of S&P index options has tended to be higher than its realized volatility. Merrill Lynch equity volatility arbitrage index extracts returns by systematically shorting variance swaps. According to [15], besides Merrill Lynch, rule based clones are also provided by Deutsche Bank, Rydex Investments, IndexIQ, and so on. There has been little academic research done on rule-based clones. The reason for this is that these clones aim to mimic trading strategies thought to be employed by hedge funds and hence cannot be discussed under a theoretical framework. There is however an academic article that researched the risk/returns of arbitrage strategies. [4] studied five fixed income arbitrage strategies: Swap spread, Yield curve, Mortgage, Fixed income volatility, Capital structure arbitrage. The paper determines trading strategies based on models and rules, implements them on a historical simulation basis, and examines the risk/return characteristics. We can think of these strategies as rule-based clones. The following paragraph is a brief summary of the strategies.

A swap spread arbitrageur enters into a position when the spread between the swap rate and the coupon rate of a Treasury bond is larger than the expected average short term spread. He then closes the position when the spread converges to the expected average value. A yield curve arbitrageur tries to exploit mis-valuations along the yield curve. He assumes a term structure model and fits the model to match exactly certain points along the swap curve for each month. Then, the other swap rates are plotted to identify their distance from the fitted curve. If a swap rate is far from the model value, he takes a position to capture this mispricing until the market swap rate converges to its theoretical value. MBS arbitrage strategy consists of buying MBS passthroughs while hedging their interest rate exposure. In [4], the authors concluded that discount securities, that have weaker negative convexity than premium securities, are more profitable. Fixed income volatility arbitrage strategy is similar to the method employed by Merrill Lynch's equity volatility arbitrage index. For this strategy, the arbitrageur exploits the difference between implied and realized volatility of interest rates, while Merrill Lynch's index does the same for equities. Also, as in the case for equities, the implied volatility for interest rates tends to be higher than the realized volatility. The arbitrageur enters into a sequence of one-month variance swaps that pays him the difference between the initial implied variance of an interest rate caplet and the realized variance for the corresponding Eurodollar futures contract for each month. He earns profits when the realized volatility is less than the implied volatility of interest rate caps. Capital structure arbitrageur exploits mispricing between a company's debt and its listed equity. He computes the theoretical credit default swap (CDS) spread and the size of an equity position needed to hedge changes in the value of the CDS using a model. Following this, the arbitrageur compares the theoretical CDS spread with the level quoted in the market. If the market spread is higher (lower) than the theoretical spread, he shorts (longs) the CDS contract, while simultaneously maintaining the equity hedge. The strategy becomes profitable when the market spread and the theoretical spread converges. Details of these strategies are further explained in the paper.

The main risk and return characteristics of the five above mentioned strategies implemented in the paper are as follows. All five of the strategies on average generate positive excess returns under the 10% volatility levels. Sharpe ratios range from about 0.3 to 0.9, which is comparable with the ratio of 0.72 reported by [31] for a broad set of fixed income arbitrage funds. Most excess returns are positively skewed as even though these strategies produce large negative returns from time to time, they tend to generate even larger offsetting positive returns. After risk adjusting for both equity and bond market factors, the article found significant alpha levels for the yield curve, mortgage, and capital structure arbitrage strategies. These strategies require more "intellectual capital" to implement, whereas swap spread and volatility arbitrage strategies that require less "intellectual capital" resulted in insignificant alpha levels.

As described earlier, one can create rule-based clones by mimicking basic investment strategies thought to be employed by hedge funds. Since trading strategies are implemented based on some rules, there are no style drift and manager's investment misjudging. However, these products have several problems. Firstly, profit-earning opportunities disappear as more people recognize and exploit these opportunities. For example, equity and fixed income volatility arbitrage capture the tendency that implied volatility is higher than realized volatility. As the number of arbitrageurs who short options or variance swaps in order to exploit the market anomaly increases, these investment opportunities will disappear. [4] found that strategies that require more "intellectual capital" to implement resulted in significant alpha, whereas others do not. This problem derives the following problem.

The second problem is that the product provider may not disclose the trading rule sufficiently. High transparency is one of the original strengths of hedge fund replication products. However, the product development process requires the use of portfolio optimization methods or the use of models to find mispricings and calculating hedge ratios. If such methods were made fully transparent and became common, market inefficiencies will disappear. In addition, if many rule-based clones come to market, the providers will seek ways to differentiate themselves by developing proprietary models and optimization programs. Eventually, this process might lead to black box methods that will not allow investors to identify risks adequately. It is difficult to earn profits using common strategies, while strategies using proprietary models often provide low transparency. Investors should select products with a full realization of this trade-off.

5 Factor-based approach

As described in section 3, the factor-based hedge fund clone providers try to replicate hedge fund accessibility to alternative risk premium and control exposures to risk factors. In other words, they attempt to deliver $\sum_k \beta_{ik} F_k$ in equation (1). This is a three step process described below.

Firstly, these providers find risk factors that drive hedge fund returns. Asset class factors such as stock and bond indices, and non-linear factors such as index options and lookback straddles can explain hedge fund returns. Rule-based clones can also be powerful risk factors. It is desirable to find the factors through tradable instruments at this stage. Secondly, they estimate the current exposures of hedge funds to the factors. For this, they need to employ statistical techniques. For example, [22] tried to capture dynamically changing exposures by rolling regression using 24 monthly data. We can assume that proprietary methods have been developed for this purpose. Finally, they create and manage the factor exposures using tradable instruments. If these providers were able to fund factors through tradable instruments during the first step, then this final step is a straightforward process. These steps are used to replicate month-to-month hedge fund returns.

Several firms have launched factor-based clones, which try to replicate aggregate hedge fund returns. Figure 6 graphs the performance of these clones. Although the replication targets are not disclosed exactly, we compare them with HFRI fund weighted composite and HFRI FoF indices for example. Most of these clones successfully approximate the returns of HFRI indices, while Clone 2 and Clone 5 do not. Clone 2 outperforms hedge fund indices. In contrast, Clone 5 underperforms the indices due to large drawdowns, although its correlation with HFRI fund weighted composite index is 0.90.

As just described, the factor-based clones can fail to track hedge fund returns. Furthermore, the potential of these replication methods cannot be evaluated sufficiently due to their short track records. To find a reason for this failure, we considered the misestimation of the current exposures to factors. For example, if we estimate the current

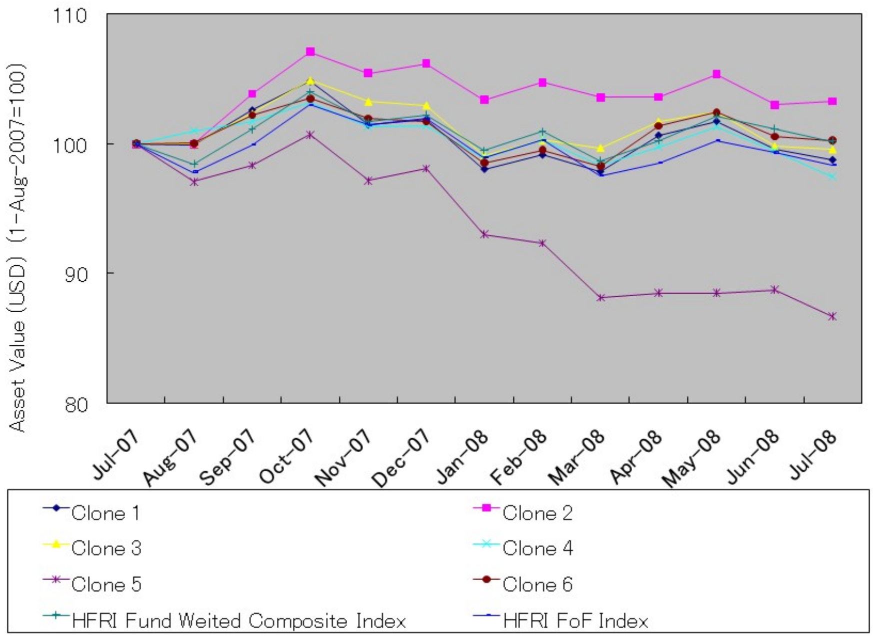

Figure 6. Performance of factor-based clones (Data source: Bloomberg)

exposure by regression using the most recent 24 monthly returns, the estimated exposures are reflective of the average exposures in the previous two years. Hence in this case, we are not able to capture intermediate exposure changes during this period in a timely manner. However, even though data for a shorter time period can estimate more recent exposure changes, such data is prone to large estimation errors. It is therefore a challenging task to estimate the current exposure levels with high accuracy.

In this section, we report our attempt to replicate AsiaHedge Asia excluding Japan and Japan Long/Short Indices. Firstly, we searched for investable risk factors for AsiaHedge Asia excluding Japan Index. Since most of hedge funds in Asia are equity-oriented, we can infer that Asian stock indices explain a substantial part of the hedge fund index return. Figure 7 plots the excess return of the hedge fund index against MSCI Asia Pacific excluding Japan index. There exists a strong linear relation between both indices. This result is consistent with [17]. The MSCI Asia Pacific excluding Japan index is tradable through ETFs. Figure 8 describes the tracking pattern of Lyxor ETF MSCI AC Asia Pacific to the index. Historically, the ETF exhibits minimal tracking error and this allows us to conclude that the Lyxor ETF has almost the same movement as that of the benchmark MSCI index. This in turn allows us to use the historical returns of the MSCI index for performance analysis as the index has a longer track record than the Lyxor ETF. To further improve the explanatory factor, we ran a linear regression adding an additional bond index factor that is represented by iBoxx ABF Pan Asia unhedged total return index. However, as this did not improve the explanatory power, we decided to use MSCI Asia Pacific excluding Japan index as the single factor to replicate AsiaHedge Asia excluding Japan Index.

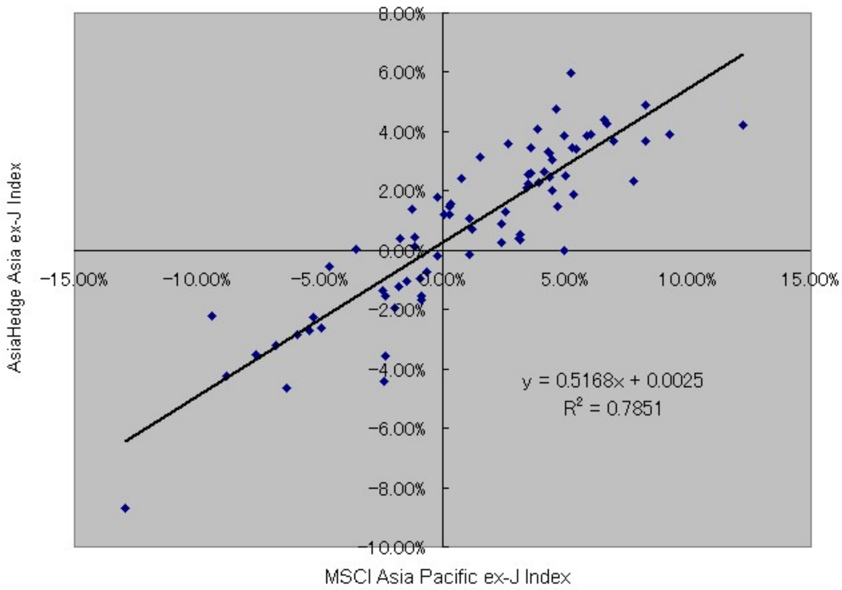

Figure 7. Monthly excess return of AsiaHedge Asia ex Japan Index against MCSI Asia ex-J Index (Jan/2002-May/2008)

Panel A of figure 9 graphs the excess return of the AsiaHedge Japan Long/Short Index against the TOPIX, which is tradable by futures or ETFs. The result showed the explanatory power to be insufficient and to improve it, we added the SMB factor, which is represented by the return of the Russell/Nomura small cap index minus the Nikkei 225 Index. Both the Russell/Normura and Nikkei indices are also tradable by ETFs. We were not able to add HML factors as value/growth style indices in Japan are not tradable. Figure 10 describes the tracking pattern of ETFs to these indices. As shown, tracking errors for these cases are very small. The excess return of the hedge fund index is regressed to TOPIX and SMB factors using the data from January 2002 to May 2008 period. The resulting regression coefficients are 0.34 and 0.25 respectively. Panel B of Figure 9 plots the excess return of the hedge fund index against the portfolio that is constructed by 34% to TOPIX and 25% to SMB factors. The explanatory power as measured by R^2 is 0.81. We judge this to be sufficient and therefore make use the two factors for the replication of the AsiaHedge Japan Long/Short Index.

When we attempt to replicate the hedge fund index in practice, we cannot use future data. Therefore, we estimate the hedge fund returns exposure on the identified risk factors by using past data. The easiest estimation method is known as rolling window regression, where we make use of 24 monthly returns data for example. For the purposes of this replication, we set the intercept in the regression at zero. As mentioned

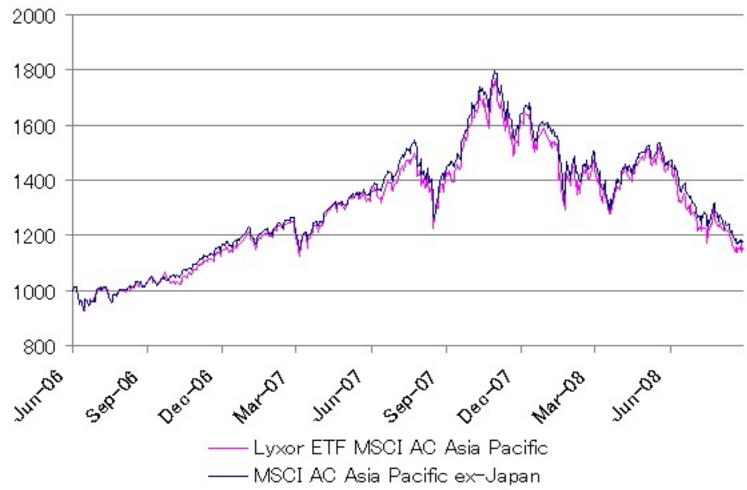

Figure 8. Tracking of Lyxor ETF to MSCI AC Asia Pacific Index (Data source: Bloomberg)

previously, the use of 24 monthly returns data is only reflective of the average exposure during past two years and does not capture recent exposure changes. To remedy this problem, we also estimated the exposure changes by using Kalman filter, which is able to capture dynamic exposure changes. The model and estimation procedures used in this replication are as follows.

We model the exposure change by state space model. Let R_t denote the hedge fund returns, F_t be the K dimensional vector that represents factor returns, β_t be unobservable K dimensional factor exposure vector. Let us assume that the exposures change as follows.

$$\beta_t = \beta_{t-1} + \epsilon_t, \tag{2}$$

$$R_t = \beta_t' F_t + \eta_t, \tag{3}$$

where ϵ_t is representative of white noise that follows K dimensional Gaussian distribution with mean $\mathbf{0}$, and covariance matrix Σ, and η_t is a random variable that follows one dimensional normal distribution with mean 0 and variance σ_η^2. Our aim is to estimate unobservable factor exposures by using observable return data of R_t and F_t. For a replicating purpose, we assumed that expected hedge fund alpha is zero.

We use Kalman filter for the estimation. Let $Y_t = \{R_1, \cdots, R_t, F_1, \cdots, F_t\}$ be observed values. We represent conditional expectation and covariance matrix as

$$\beta_{t|s} = E[\beta_t | Y_s],$$

$$V_{t|s} = E[(\beta_t - \beta_{t|s})(\beta_t - \beta_{t|s})' | Y_s].$$

$\beta_{t|s}$ is called prediction if $s < t$, and filtering if $s = t$. We can obtain these values sequentially by repeating prediction for one period ahead and filtering one after another.

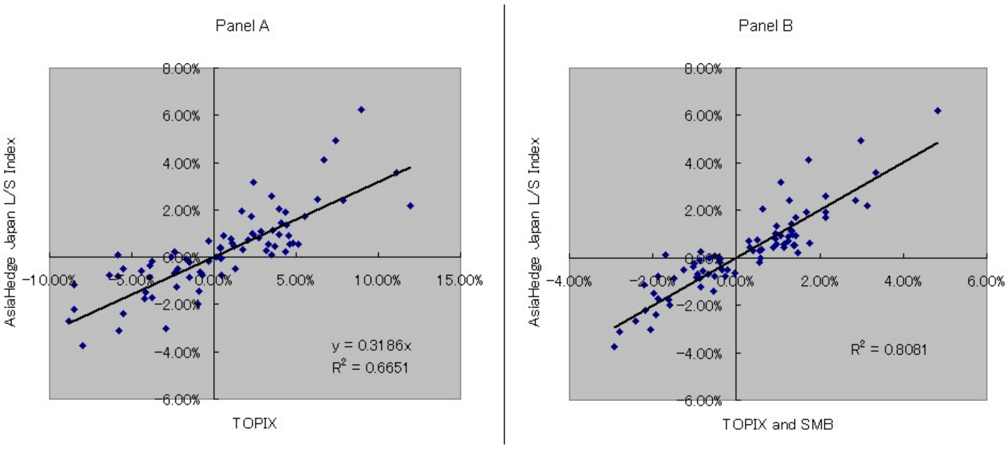

Figure 9. Monthly excess return of AsiaHedge Japan Long/Short Index against TOPIX and the portfolio of 34% TOPIX and 25% SMB (Jan/2002-May/2008)

First, the prediction for the following period is obtained by

$$\beta_{t|t-1} = \beta_{t-1|t-1},$$

$$V_{t|t-1} = V_{t-1|t-1} + \Sigma.$$

In the prediction algorithm, $\beta_{t|t-1}$ is same with filter $\beta_{t-1|t-1}$, while $V_{t|t-1}$ consists of two terms. The first is filter $V_{t-1|t-1}$ and the second is the effect of white noise η_t. Then, filtering is calculated as

$$\beta_{t|t} = K_t R_t + (I - K_t F_t')\beta_{t|t-1},$$

$$V_{t|t} = V_{t|t-1} - K_t F_t' V_{t|t-1},$$

where $K_t = (F_t' V_{t|t-1} F_t + \sigma_\eta^2)^{-1} V_{t|t-1} F_t$, which is referred to as Kalman gain. Therefore, filtering $\beta_{t|t}$ is a weighted average of new observation R_t and prediction $\beta_{t|t-1}$, where the weight on the new observation is Kalman gain. The second term for filtering $V_{t|t}$ represents the improvement of state estimation accuracy due to new observation. We estimate the parameters Σ and σ_η^2 by maximum likelihood, and then exposures to risk factors from January 2002 to May 2008. Figure 11 graphs the exposure changes of AsiaHedge Asia ex-Japan Index to MSCI Asia Pacific excluding Japan index estimated using rolling window regression and Kalman filter. Figure 12 does the same comparison between AsiaHedge Japan Long/Short Index and the two factors (TOPIX and SMB). These graphs show that on average, the Kalman filter captures exposure changes earlier than the rolling window regression.

Finally, we proceed to back test the performance of our replication models through making investments into the identified ETFs at levels estimated by the statistical methods. Figure 13 and 14 exhibit statistics of the replication results for AsiaHede Asia ex-Japan and Japan Long/Short indices respectively. In-sample figures refer to static

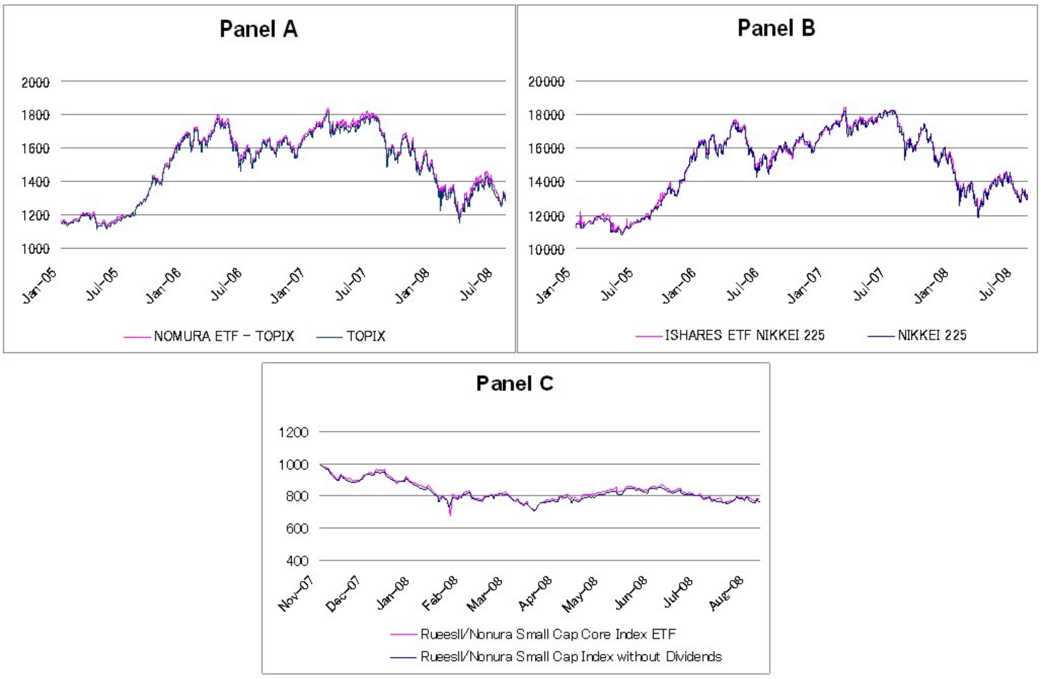

Figure 10. Tracking of ETFs to Japanese benchmark indices (Data source: Bloomberg)

exposures estimated through regression analysis using data from January 2002 to May 2008 period. Data of future periods is not available in practice, but we have included the estimated value using such data for comparative purposes. Rolling-window and Kalman filter capture dynamic exposures using past data. These are available methods in practice. One of the most important statistics is the correlation with the replication target, which is shown at the bottom of the table. The replicated returns by Kalman filter have higher correlation than rolling-window regression for both cases. Other statistics also show Kalman filter to be the better replicator. Especially, using Kalman filter succeeded in replicating skew and kurtosis better than rolling window. We conclude that the current exposure of both AsiaHedge hedge fund indexes can be captured more accurately by Kalman filter than by rolling-window regression. Moreover, statistics of Kalman filter is better than that of in-sample regression. In-sample regression uses future data, which does not allow for it to capture changing exposures within the period, leading to worse results. Rolling-window regression underperforms in-sample regression, which effectively means that the loss incurred by the delay of capturing exposure change is larger than by inability to capture dynamic changes of exposure.

If one could find appropriate risk factors and capture the change of exposures to them quickly, he can replicate the target hedge fund returns on a month-to-month basis. Although exposure estimation by Kalman filter was the best in this example, it can be insufficient for some cases. The state space model that Kalman filter uses is for

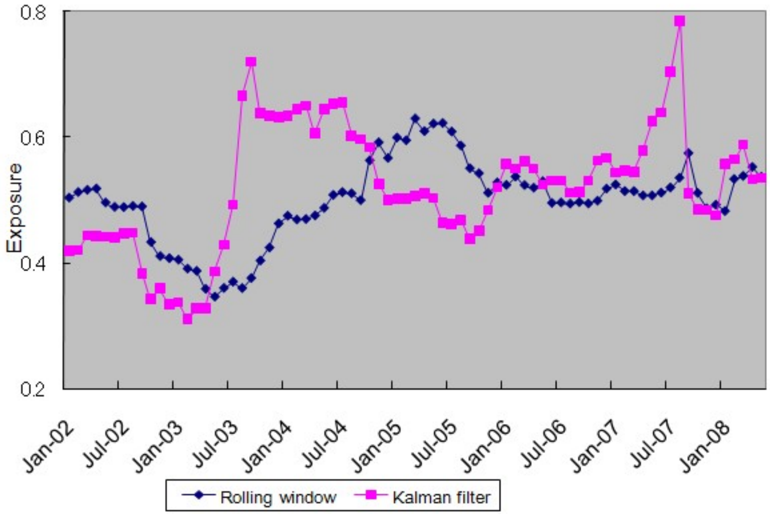

Figure 11. Exposure change of AsiaHedge Asia ex-Japan Index to MSCI Asia Pacific excluding Japan index estimated by rolling window regression and Kalman filter

the case of normal white noise. Therefore, this model cannot capture drastic exposure changes. [21] estimated the exposure changes of mutual funds for non-Gaussian white noise cases using the Monte Carlo filter. Similar research should be done for hedge funds using Monte Carlo and other non-linear filtering methods to catch drastic exposure changes appropriately. Moreover, as long as we use monthly hedge funds return data, we cannot take into account dynamic exposure changes that occur within a month. Hedge fund managers monitor risk on a weekly, daily, or a more frequent basis, and adjust their investment positions according to these risk levels especially when prominent financial events occur. Some investable hedge fund indices report daily returns, but it is on an estimation basis. In many cases, these daily estimated returns are revised at a future date. The development of methods that incorporate dynamic exposure change within a month is also challenging task.

6 Distribution replicating approach

This section discusses the distribution replicating methodology that aims to replicate the joint distribution of an investor's portfolio and hedge fund returns. Unlike the factor-based approach, the distribution replication approach does not aim to replicate the target hedge fund returns on a month-to-month basis. Instead, this method aims to generate returns that have the same distribution pattern as the hedge fund returns. When it is difficult to find to risk factors that explain the target hedge fund returns, factor-based approach cannot work. This approach would be a powerful replication method for that case. Since investors expect hedge funds to have low correlation with

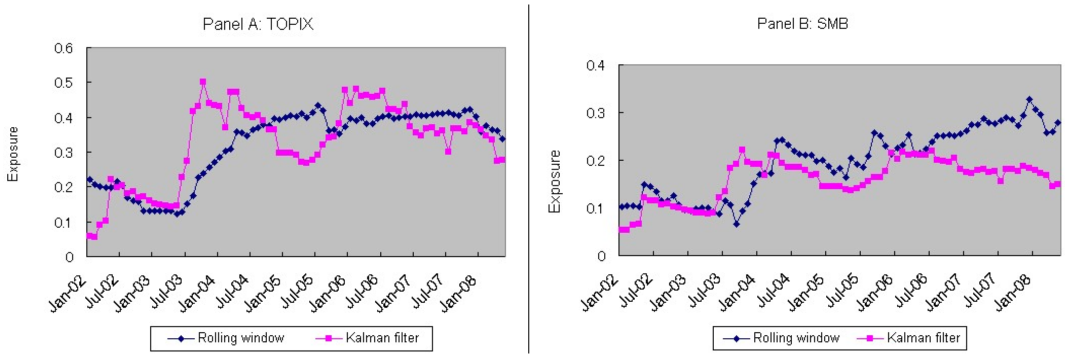

Figure 12. Exposure change of AsiaHedge Japan Long/Short Index to SMB estimated by rolling window regression and Kalman filter

	AsiaHedge Asia ex Japan Index	In-sample	Rolling window	Kalman filter
Annual return	14.16 %	11.07 %	10.50 %	11.71 %
Annual std. dev.	9.53 %	8.66 %	8.28 %	8.57 %
Sharpe ratio	1.16	0.92	0.90	1.01
Skew	-0.77	-0.55	-0.39	-0.58
Kurtosis	3.70	3.33	3.48	3.80
Max drawdown	12.33 %	11.58 %	10.61 %	12.12 %
Correlation		0.89	0.86	0.93

Figure 13. Statistics for replicated returns and the AsiaHedge Asia ex-J Index

their existing portfolio, it is also desirable to replicate the dependence between the investor's portfolio and hedge fund returns. One way of achieving this is to replicate the joint distribution of investor's portfolio and hedge fund returns.

The above mentioned methodology is proposed by [20]. In short, the basic concept is to price the distribution of hedge fund's payoff and replicate it through the dynamic trading of tradable securities. The theoretical framework is based on [5]'s payoff distribution pricing model. A given payoff distribution can be replicated by many different payoff functions. It showed that the cheapest way to get a payoff distribution is to allocate terminal wealth as a decreasing function of the state price density. [20] extended the payoff distribution pricing model to a bivariate setting, and considered the joint distribution with investor's portfolio. Details of the theoretical background are described in the papers. [27] proposed an alternative way to perform [20]'s replication methodology.

The following explains their attempt at using the distribution replicating approach. First, we describe the settings. Let us assume asset S^1 is the investor portfolio, asset S^2 is a reserve asset and asset S^3 is a hedge fund. Our goal is to replicate the joint

	AsiaHedge Japan Long/Short Index	In-sample	Rolling window	Kalman filter
Annual return	6.41 %	5.42 %	4.71 %	5.39 %
Volatility	6.00 %	5.44 %	5.15 %	5.51 %
Sharpe ratio	0.55	0.43	0.32	0.42
Skew	0.61	0.20	0.05	0.36
Kurtosis	4.68	2.90	3.14	4.37
Max drawdown	10.78 %	9.26 %	10.92 %	9.70 %
Correlation		0.90	0.88	0.95

Figure 14. Statistics for replicated returns and the AsiaHedge Japan Long/Short Index

distribution of monthly returns of S^1 and S^3 through the dynamic trading of S^1 and S^2. We set 0 as the start of the month and T as the end of the month. The monthly logarithmic returns of the three assets are represented by R_i for $i = 1, 2, 3$. We denote the distribution function for R_i by F_i and the copula function with respect to the joint distribution of $\left(R_1, R_3\right)$ by C. To dynamically trade S^1 and S^2, we need to identify the asset price process of S^1 and S^2. They are denoted by $\{S_t^1\}_{t=0}^T$ and $\{S_t^2\}_{t=0}^T$. We set $S_0^1 = S_0^2 = 1$.

To implement the replicating strategy, we need to proceed with the following steps. We first obtain the joint distribution of investor's portfolio and hedge fund returns $\left(R_1\right.$ and $\left.R_3\right)$. This is the target distribution to replicate. Next, we choose a reserve asset S^2 and infer the stochastic processes $\{S_t^1\}_{t=0}^T$ and $\{S_t^2\}_{t=0}^T$. Following that, we also obtain the joint distribution of R_1 and R_2. We then determine the payoff function, which transform the joint distribution of $\left(R_1, R_2\right)$ to that of $\left(R_1, R_3\right)$. Finally, we price the payoff distribution and replicate the payoff through the dynamic trading of S^1 and S^2. In the following, we explain the each step more in detail.

There are several ways to determine the joint distribution of R_1 and R_3. The easiest is to model them using the bivariate probability distribution. However, hedge funds exhibit different return characteristics to traditional asset classes, such as skews and fat-tails, and are nonlinearly related in some cases as shown in section 2. Therefore, [20] proposed modeling investor's portfolio and hedge fund returns separately, and connecting them by copula. Some copulas can capture asymmetric dependence structure flexibly. For example, Clayton copula has more dependence in the lower tail than in the upper tail. This allows for the copula to capture the strong dependence in bear markets and weak dependence in bull markets. Risk arbitrage funds tend to exhibit a similar dependence structure.

For hedge fund returns, it is desirable to use the distribution class that can capture its skewness and fat-tails. For example, the previous research by [20] and [27] used Gaussian distribution, Student-t, Gaussian mixture and Johnson distribution. We denote the distribution function of standard normal distribution by $N(\cdot)$, and the density of Gaussian distribution with mean μ and variance σ^2 by $\phi(x; \mu, \sigma^2)$. Gaussian distribution is determined by only mean and variance, and therefore does not capture skewness and fat-tails. Student-t captures can capture some kurtosis but cannot model skewness. Gaussian mixture with m regimes and parameters $\{\pi_k\}_{k=1}^m$, $\{\mu_k\}_{k=1}^m$, $\{\sigma_k^2\}_{k=1}^m$ has the density

$$f(x) = \sum_{k=1}^{m} \pi_k \phi(x; \mu_k, \sigma_k^2),$$

where $\phi(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{ - \frac{(x-\mu)^2}{2\sigma^2} \Big\}$ is the density of Gaussian distribution with mean μ and variance σ^2. The intuitive interpretation of this distribution is as follows. It is assumed that there are m regimes (states), which happen with probability $\{\pi_k\}_{k=1}^m$. If regime k occurred, the conditional distribution is Gaussian whose density is $\phi(x;\mu_k,\sigma_k^2)$. Allowing multi-regimes, this distribution class can capture skewness and kurtosis. The distribution functions of Johnson distributions are given by

$$F(x) = N\left(\gamma + \delta h\left(\frac{x-\xi}{\lambda}\right)\right),$$

where γ and δ are shape parameters, ξ and λ are location and scale parameter respectively, and $h(\cdot)$ is one of the following functions:

$$h(z) = \begin{cases} \log z & \text{(log-normal)} \\ \log\left(z + \sqrt{z^2+1}\right) & \text{(unbounded)} \\ \log\left(\frac{z}{1-z}\right) & \text{(bounded)} \end{cases}$$

Johnson distributions are often used for analysis of non-normal behavior of hedge fund returns as seen in [19], [28], and [29].

For investor's portfolio, we have to use a well-fitted probability distribution of monthly returns, and model its stochastic process thereafter. To carry this out, we should model the monthly returns to be consistent with the stochastic process. [27] pointed out that [20] had modeled the monthly return and stochastic process inconsistently. To correct the inconsistency, [27] modeled the daily and monthly returns using a Gaussian mixture. The most consistent way is to model the stochastic process of the investor's portfolio and then derive the probability distribution of the month end logarithmic returns from the stochastic process.

Copula integrates the marginal distribution of investor's portfolio and hedge fund returns, and defines the joint distribution. The copula function C of the random vector (R_1, R_3) is the joint distribution function of the random vector (U, V), where $U = F_1(R_1)$ and $V = F_3(R_3)$. Then, the joint distribution function of (R_1, R_3) satisfies

$$F_{1,3}(r_1, r_3) = C\Big(F_1(r_1), F_3(r_3)\Big).$$

The previous works used Gaussian, Student-t, Gumbel, Clayton, Frank, and symmetrised Joe-Clayton copulas. The characteristics of these copula is explained in [20]. Further details are argued in [18] and [26].

There are some ways to estimate the parameters and select model. [20] and [27] used different parameter estimation and model selection procedures. [20] estimated the parameter for copula and marginal distribution of R_1 and R_3 all together by maximum likelihood. They then selected the final model by AIC. On the other hand, [27] estimated the parameters separately. First, they estimated the parameter for the marginal distributions and selected the model using a goodness-of-fit test. They then ranked the time-series data of the returns to determine the copula. Through this method, they hoped to work around the copula selection without the effect of mis-specification and estimation error of the marginal distributions. Further detail regarding these procedures is described in the papers.

After we determine the copula, we need to infer the stochastic process of investor's portfolio S^1 and reserve assets S^2. Remember that the time-series structure of these assets is required because we aim to replicate the payoff of hedge funds through the dynamic trading of these assets. The strategy is derived through using methods of pricing and hedging of derivatives such as tree models, Monte Carlo simulation and others. The easiest way is to model the asset price process by geometric Brownian motion and estimating the parameters. In using the tree model, we approximate the stochastic process using a bivariate binomial or trinomial tree. For both these cases, it is only necessary to estimate some of the statistics of instantaneous (or daily) return such as mean, standard deviation, and correlation. On the other hand, if one uses Monte Carlo simulation, one has to make assumptions and infer the stochastic process so as to fit the sample data.

After estimating the parameters and selecting models of the asset returns and price processes, we have to next determine a payoff function. Since we are trying to obtain the joint distribution of investor's portfolio and hedge fund returns through the dynamic trading of the investor's portfolio and reserve assets, we require a function $\tilde{g}$ such that

$$P\Big(R_1 \leq x,\ \tilde{g}\Big(R_1, R_2\Big) \leq y\Big) = P(R_1 \leq x, R_3 \leq y) \ \text{ for any } x, y,$$

or equivalently,

$$P\Big(\tilde{g}\Big(R_1, R_2\Big) \leq y \Big| R_1 = x\Big) = P(R_3 \leq y | R_1 = x) \ \text{ for any } x, y.$$

Then, $\tilde{g}(\cdot, \cdot)$ is given by

$$\tilde{g}(x, y) = Q\Big\{x, P\Big(R_2 \leq y | R_1 = x\Big)\Big\},$$

where $Q(x, \alpha)$ is the order α quantile of the conditional law of R_3 given $R_1 = x$. In terms of asset prices, the payoff function is represented as

$$g(S^1_T, S^2_T) = \exp\Big\{\tilde{g}\Big(\log S^1_T, \log S^2_T\Big)\Big\}. \tag{4}$$

Given the payoff function, the problem comes down to pricing and hedging of contingent claim. Therefore, we can utilize methods developed for option pricing. The present value of the objective payoff can be obtained by the discounted expectation of the payoff under risk-neutral measure. If one modeled the market to be complete, the payoff can be perfectly replicated through a delta hedging strategy. [20] recommended to implement these processes using Monte Carlo simulation or a trinomial tree, while

[27] proposed the use of a quadratic hedging strategy with a combined method of tree model and Monte Carlo simulation. If the present value of the payoff is equal to one, the replication strategy delivers a return that has the same distribution with that of the target hedge fund. If it exceeds one, we need additional cost to replicate the payoff. Otherwise, we are able to get hedge fund returns at lower cost.

Through the steps described above, one can theoretically replicate the distribution of hedge fund returns. However, there can exist replication errors due to the following reasons. First, the estimation error can exist in the inference of both the hedge fund return distributions and the stochastic process of investor's portfolio and reserve assets. Furthermore, even when the returns are generated from known population distribution data and parameters are estimated using simulated data, these estimated parameters can include some errors especially in the case of small samples. The inference process is performed using some assumptions such as stationarity. If there is a structural break in the sample data period used for parameter estimations, the inference process would fail. Second, approximation errors can occur when the tree model is used to calculate both the value of the payoff and size of position, and numerical errors can occur if Monte Carlo simulation is used instead. The accuracy of replication can be evaluated through comparing some statistics of the target hedge fund returns and the replicated returns.

[20] studied the effects estimation errors, derived from inference using small sample data. They assumed the true population distribution of asset returns to be stationary. The investor's portfolio and reserve assets log returns were assumed to follow a bivariate Gaussian distribution with a parameter set. As for hedge fund returns and the copula between those of hedge fund and investor's portfolio, they considered two cases. The first is that the hedge fund log returns follow a Gaussian distribution and the copula is a symmetrised Joe-Clayton copula with some parameter sets. This copula's dependence in the upper tail is stronger than in the lower tail. The second is that hedge fund log returns follows a Johnson unbounded distribution with parameters that set skewness levels at 2.0 and excess kurtosis levels at 10. The copula in this case is a Gaussian copula, which has a dependence structure similar to that of a bivariate Gaussian distribution. In their study, return samples were drawn from the population using different sample sizes of N (= 24, 48, 72, 96, 120, 240), to examine the effect of small samples. Using the data samples, they inferred return distributions, and derived payoff functions by (4). They took 2000 observations from the investor's portfolio and reserve assets, substituted them into the payoff function, and obtained the payoff outcomes.

They then calculated the mean, standard deviation, skewness, kurtosis, and correlation with the investor's portfolio. These procedures were repeated 100 times for each N, and the different sets of statistics were compared and evaluated. They showed that, for the case of Gaussian distribution and symmetrised Joe-Clayton copula, the larger the sample, the more accurate the payoff function is. Even with a relatively small sample the procedure still works quite well. For the case of Johnson unbounded distribution and Gaussian copula, they could not conclude that a larger sample improves the payoff function accuracy as like the previous example. Moreover, the replication accuracy of skewness and kurtosis was low, because small samples do not contain enough information on tail events. These results imply that if we succeed in calculating the present value of the payoff and position size, we can replicate the joint distribution of investor's portfolio and hedge fund return with good accuracy excluding tail events.

Following their simulation analysis, [20] attempted to replicate an actual fund of hedge funds and a further two hedge funds on an out-of-sample basis. To achieve this, they estimated parameters and selected models using only preceding sample data of the month they tried to replicate. Therefore, their results are reflective of actual replication that can be done. They assumed that the investor's portfolio consisted 50% of a S&P 500 tracking portfolio and 50% of a long-dated US Treasury bonds, and used the nearest Eurodollar futures as the reserve asset. As for replicating the fund of hedge funds, they succeeded in cloning the joint distribution of investor's portfolio and the fund returns in most cases except for three large losses from 1987 to 2004. As for replicating the two hedge funds, [20] generated similar statistical properties as the fund returns, although the returns appeared in a different order. Further detailed results are described in the paper.

[27] tested their replication accuracy using an in-sample test. For example, they performed parameter estimation and model selection using sample data from 2002 to 2006, and attempted to replicate hedge fund returns during that period. Note that we cannot get the result when we implement their replication strategy in practice. Their methodology was applied to HFRI and EDHEC hedge fund indices by strategy. The volatility and the dependence coefficients (Kendall's tau) were replicated with great precision, and the skewness and kurtosis were with slightly less accuracy. The choice of reserve asset impacted the initial cost of investing in the replicating portfolio, leading to the difference of mean returns.

Separately, we attempted to clone Eurekahedge global CTA/managed futures and macro indices on an out-of-sample basis. [17] showed that it was difficult to find risk factors that explained returns of hedge funds for these strategies. Hence, we used a distribution replicating approach instead of attempting to replicate the returns of these indices on month-to-month basis using a factor-based approach. The inception month of these indices is January 2000. We used the data from the first two years since inception to estimate the return distributions while the monthly log-returns from January 2002 to May 2008 were replicated. We assumed the investor's portfolio S^1 to be composed of 50% Japanese stocks and 50% Japanese government bonds (JGB). Since we are required to dynamically trade these assets, TOPIX futures and long-term JGB futures were used as proxies. Both these securities are listed on the Tokyo Stock Exchange. The reserve asset S^2 is made up of 25% S&P 500 futures, 25% NYMEX WTI crude oil futures, 25% COMEX gold futures, and 25% JPY against USD spot currency. All data is obtained from Bloomberg. Log-returns on futures are calculated by rolling the front contract. The front contract is rolled on the last trading day of the maturity month. Our base currency is USD. And since TOPIX and JGB futures are denominated in JPY, we applied a currency hedge. Accordingly, the log returns of these assets are adjusted by the difference between the interest rates of USD and JPY. Libor rates were used for the interest rates. We ran a linear regression of the monthly log-returns of the two indices against the above mentioned six assets using data from January 2002 to May 2008. The adjusted R^2s for the global CTA/Managed futures and macro indices were 0.22 and 0.48 respectively. The statistics for monthly log-returns of investor's portfolio and reserve asset January/2002-May 2008 are shown in Figure 15, where Kendall's tau is a dependence coefficient defined by

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0],$$

where (X_1, Y_1) and (X_2, Y_2) are independent and identically distributed random vectors.

	Investor's portfolio	Reserve asset
Mean	0.75 %	1.19 %
Std. Dev.	2.04 %	2.68 %
$\frac{\text{Mean}}{\text{Std. Dev.}}$	0.37	0.44
Skew	-0.15	0.05
Kurtosis	2.68	2.33
Correlation with investor's portfolio	1.00	0.35
Kendall's tau with investor's portfolio	1.00	0.19
Min	-3.87 %	-4.17 %
Max	5.37 %	7.39 %
Number of positive return months (Total: 77)	53	50

Figure 15. Statistics for the log-returns of investor's portfolio and reserve asset (January/2002-May/2008)

We implemented the following replication procedure for each month using past data. First, parameter estimation and model selection for the marginal distribution of monthly log-returns of hedge fund were performed using the same method mentioned in [27]. The best fitting model is chosen from a Gaussian mixture with m regimes ($m = 1, 2, 3, 4, 5$) and Johnson unbounded distribution. Next, the copula model between monthly log-returns of hedge fund and investor's portfolio is estimated and selected in the same manner as in [27], which is through the ranking of time-series data. The copula is selected from Gaussian, Student, Clayton, Frank and Gumbel. Following that, we modeled the stochastic process of the investor's portfolio and reserve asset through a bivariate log-normal process;

$$\begin{aligned} \frac{dS_t^1}{S_t^1} &= \mu_1 dt + \sigma_{11} dW_t^1, \\ \frac{dS_t^2}{S_t^2} &= \mu_2 dt + \sigma_{21} dW_t^1 + \sigma_{21} dW_t^2 \end{aligned}$$

with $S_0^1 = S_0^2 = 1$. The parameters were estimated using daily data. Then, the monthly log-return of investor's portfolio is obtained by $N(\mu_1 - \sigma_{11}^2/2, \sigma_{11}^2)$ as a byproduct. Finally, we replicated the payoff function (4) by using a delta-hedging strategy on a daily basis. We calculated the present value and delta of the payoff using Monte Carlo simulation. Figure 16 describes the obtained investment weights on the investor's portfolio and reserve asset. Panel A and B are the weights for global CTA/Managed futures and global macro respectively.

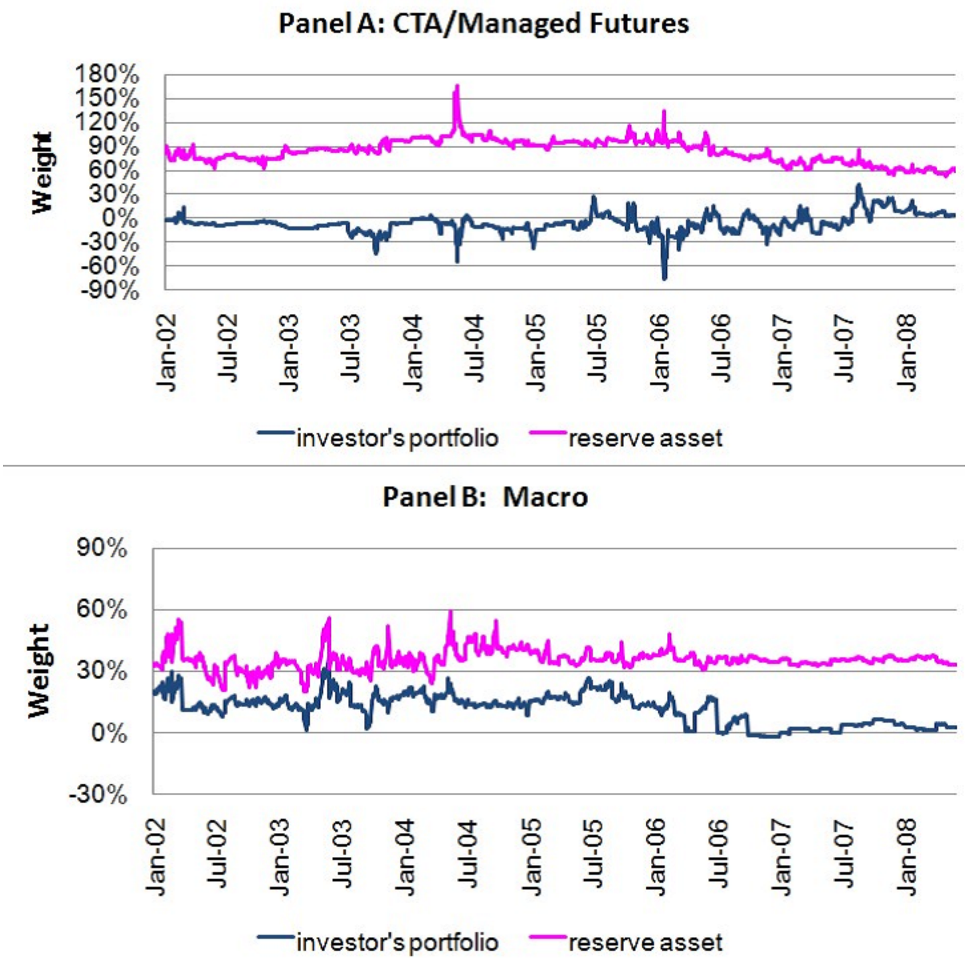

Figure 16. The investment weights on the investor's portfolio and reserve asset

The results of our replication process are as follows. Figure 17 exhibits the monthly log-returns and the growth of the asset value of the target hedge fund indices and the clones. Panel A is the result for Eurekahedge global CTA/Managed futures, and Panel B is that for Eurekahedge global macro. Panel 1 describes the time-series of monthly log-returns. The correlation between the log-returns of target hedge fund index and its clone are included in the panel. The correlations for CTA/Managed futures and global macro are 0.38 and 0.52 respectively, which show that the clones do not replicate the returns of target on a month-to-month basis. Panel 2 exhibits the growth of asset values of the target hedge fund indices and the clones. The CTA/Managed futures clone achieved the same amount of profits with the target in the long run. Though, the shapes of their equity curves are similar, the timing of their returns differs. On the other hand, the global macro clone earns less return than the target. In this case, though their equity curves are also similar, there are differences in return levels. Figure

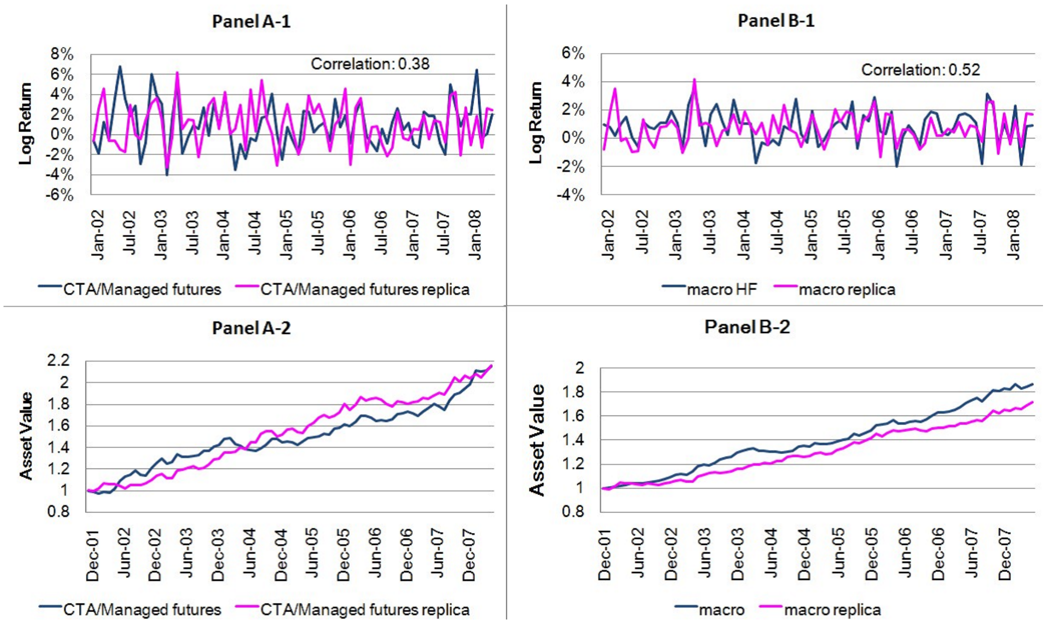

Figure 17. Monthly log-returns and growth of asset values of the target hedge fund indices and the replicas

18 exhibits the statistics for the returns of the target hedge fund and the replicas. For the case of CTA/Managed futures, mean and standard deviation are replicated with great accuracy. Minimum and maximum returns and the number of positive returns are also close. The skewness, kurtosis and the dependence coefficients (correlation and Kendall's tau) were slightly less accurate. For the case of global macro, the result is similar to that of CTA/Managed futures except for the mean and skewness. The mean return is slightly less than the target, and the sign of skewness is opposite from the target. However, many of the other statistics are replicated with better accuracy than in the case of CTA/Managed futures. These include standard deviation, kurtosis, dependence coefficients, and minimum value.

In summary, the above mentioned methodology enables us to obtain the joint distribution of the investor's portfolio and the target hedge fund's returns. The merit of this approach is that we can gain access to the statistical properties of hedge fund returns even if we are not able to find risk factors that explain hedge fund returns. This method,

	CTA/ Managed futures	CTA/Managed futures replica	Macro	Macro replica
Mean	0.99 %	1.00 %	0.81 %	0.70 %
Std. Dev.	2.27 %	2.18 %	1.16 %	1.11 %
$\frac{\text{Mean}}{\text{Std. Dev.}}$	0.44	0.46	0.70	0.63
Skew	0.29	0.14	-0.20	0.46
Kurtosis	2.99	2.36	3.08	3.32
Correlation with investor's portfolio	0.04	0.28	0.41	0.52
Kendall's tau with investor's portfolio	0.07	0.16	0.24	0.32
Min	-4.05 %	-3.28 %	-1.99 %	-1.35 %
Max	6.77 %	6.22 %	3.49 %	4.19 %
Number of positive return months (Total: 77)	49	50	59	57

Figure 18. Statistics for log-returns of the target hedge fund and the replicas (January/2002-May/2008)

since it clones the dependence structure between the investor's portfolio and hedge fund returns, is able to satisfy investor's expectation that hedge fund returns have low correlation with their existing portfolio. Moreover, this methodology allows positions to be adjusted dynamically. However, there are still areas for improvement within this methodology. For example, the reserve asset is constructed through an equal weighted portfolio of assets that it is assumed that the target hedge funds invest in. The selection of these assets is an important determination in the construction of the reserve asset. To determine the appropriate assets is a challenging task. Going forward, developments in modeling and estimating the asset returns or price processes, and dynamic replication techniques will further improve the performance of replication.

7 Conclusion

In this paper, we discussed hedge fund replication. As long as the return characteristics of hedge funds are appealing, hedge fund clone products would prove beneficial and useful. Since hedge fund returns cannot be replicated perfectly, a number of different methods have surfaced. These methods are classified into three approaches: Rule-based, Factor-based, and Distribution replication based approaches. These approaches aim to replicate different aspects of hedge fund returns. In the above sections, we explained each method in detail. The methodologies are work-in-progress. Future research may lead to improvisation of these techniques that in turn will generate replicated returns with properties much closer to target hedge fund returns than current clones. We also eagerly await the development of hybrid methods that are able to combine the advantages of each of the various mentioned methods. Furthermore, new developments in this arena should not only be limited to existing hedge fund replications. It is going to be important to develop new investment strategies, which meet a wide range of investors' needs, through the use of transparent and liquid assets.

References

[1] Agarwal, V. and Naik, N.Y. (2004). Risks and Portfolio Decisions Involving Hedge Funds. *The Review of Financial Studies*, Vol. 17(1), pp. 63-98.

[2] Bowler, B. (2007), The Emergence of Synthetic Hedge Funds. http://www.pacificprospect.com/jsp_2007/downloads/a/2.pdf

[3] Brown, S.J. and Goetzmann, W.N. (2003). Hedge Funds With Style. *The Journal of Portfolio Management*, Vol. 29(2), pp. 101-112.

[4] Duarte, J., Longstaff, F. and Yu, F. (2007). Risk and Return in Fixed Income Arbitrage: Nickels in front of a Steamroller? *The Review of Financial Studies*, Vol. 20(3), pp. 769-811.

[5] Dybvig, P. (1988a). Distributional Analysis of Portfolio Choice. *Journal of Business*, Vol. 61, pp. 369-393.

[6] Dybvig, P. (1988b). Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market. *Review of Financial Studies*, Vol. 1, pp. 67-88.

[7] Fung, W. and Hsieh, D.A. (1997). Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds. *The Review of Financial Studies*, Vol. 10(2), pp. 275-302.

[8] Fung, W. and Hsieh, D.A. (2000). Performance Characteristics of Hedge Funds and Commodity funds: Natural vs. Spurious Biases. *The Journal of Financial and Quantitative Analysis*, Vol. 35(3), pp. 291-307.

[9] Fung, W. and Hsieh, D.A. (2001). The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *The Review of Financial Studies*, Vol. 14(2), pp. 313-341.

[10] Fung, W. and Hsieh, D.A. (2002). The Risk in Fixed-Income Hedge Fund Styles. *The Journal of Fixed Income*, Vol. 12, pp. 6-27.

[11] Fung, W. and Hsieh, D.A. (2003). Benchmark for Alternative Investments: A Risk Factor-Based Approach. Working Paper.

[12] Fung, W. and Hsieh, D.A. (2004). Extracting Portable Alphas from Equity Long-Short Hedge Funds. *Journal of Investment Management*, Vol. 2, pp. 57-75.

[13] Fung, W. and Hsieh, D.A. (2007). Will Hedge Funds Regress towards Index-like Products? *Journal of Investment Management*, Vol. 5(2), pp. 46-55.

[14] Fung, W. Hsieh, D.A., Naik, N.Y. and Ramadorai, T. (2008). Hedge Funds: Performance, Risk and Capital Formation. *Journal of Finance*, Vol. 63(4), pp. 1777-1803.

[15] Géhin, W. (2007). Hedge Fund Clones: Key Issues in Distinguishing between Them. http://www.edhec-risk.com/latest_news/RISKArticle.2007-10-11.3936?newsletter=yes

[16] Getmansky, M., Lo, A. and Makarov, I. (2004). An Econometric Analysis of Serial Correlation and Illiquidity in Hedge-Fund Returns. *The Journal of Financial Economics*, Vol. 74, pp. 529-609.

[17] Hakamada, T., Takahashi, A. and Yamamoto, K. (2007), Selection and Performance Analysis of Asia-Pacific Hedge Funds. Journal of Alternative Investments, Vol. 10(3), pp. 7-29.

[18] Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman and Hall.

[19] Kaplan, P.D. and Knowles, J.A. (2004). Kappa: A Generalized Downside Risk-Adjusted Performance Measure. Working Paper.

[20] Kat, H.M. and Palaro, H.P. (2005), Who Needs Hedge Funds? A Copula-Based Approach to Hedge Fund Return Replication. Working Paper.

[21] Kobayashi, T., Sato, S., and Takahashi, A. (2005). Style Analysis Based on a General State Space Model and Monte Carlo Filter. Working Paper.

[22] Lo, A. and Hasanhodzic, J. (2006). Can Hedge-Fund Returns Be Replicated?: The Linear Case. Working Paper.

[23] Malkiel, B.G. and Saha, A. (2005). Hedge Funds: Risk and Return. *The Financial Analysts Journal*, Vol. 61(6), pp. 80-88.

[24] Mitchell, M. and Pulvino, T. (2001). Characteristics of Risk in Risk Arbitrage. *Journal of Finance*, Vol. 56, pp. 2067-2109.

[25] Naik, N.Y., Ramadorai, T., and Stromqvist, M. (2007). Capacity constraints and hedge fund strategy returns. *European Financial Management*, 13(2), pp. 239-25.

[26] Nelsen, R.B. (1999). *An introduction to copulas: Vol. 139 of Lecture Notes in Statistics*. Springer-Verlag, New York.

[27] Papageorgiou, N., Remillard, B. and Hocquard, A. (2008). Replicating the properties of hedge fund returns. *Journal of Alternative Investments*, forthcoming.

[28] Passow, A. (2005). Omega Portfolio Construction with Johnson distributions. *Risk*, Vol. 18(4), pp. 85-90.

[29] Pérez, P. G. (2004). An approach to the non-normal behavior of hedge fund indices using Johnson distributions. Working Paper.

[30] Sharpe, W. (1992). Asset Allocation: Management Style and Performance Measurement. *The Journal of Portfolio Management*, (Winter,1992), pp. 7-19.

[31] Tremont/TASS (2004), Fourth Quarter Asset Flows Report. Tremont Capital Management, New York, NY.