Traditional and Alternative Risk: 
An Application to Hedge Fund Returns

Claudio Boido*, Antonio Fasano†‡

December 15, 2013

Abstract

The impact of investment policies, styles and strategies, on hedge fund performances spanning the period from January 2006 to December 2012 are analysed. The performance analysis is not carried out at the level of a single fund but on equal-weighted reference portfolios, built from “peer funds”, intended here as those following the same style or strategy. To this end we establish rules to segregate hedge funds into a range of investment styles. After classifying and aggregating HF returns, we investigate how and if the grouping is consistent in term of return dynamic and if it can explain the risk premium of this asset class. To enforce the analysis we take into account the risk factor by developing and testing performance measure the alternative investment excess return, weighted with downside risks. All the metrics proposed are tested thoroughly with comparative statistics and inference tests. Testing of grouping is carried out under different stress conditions, as those showed during the recent financial crisis.

JEL Classifications: G11, G12, G23

Keywords: Hedge Funds, Fund-of-Funds, Performances, Alpha.

Part I

1 Introduction

A hedge fund is an investment that offers risk and return opportunities not easily obtained with any other asset class. A hedge fund is an alternative asset class, where asset managers disclose only the main outlines of their strategy without going into technical details so as to keep their potential profits. If the financial markets are not efficient, active managers are

*Claudio Boido (boido@unisi.it) is Associate Professor (Professore Associato) of Financial Market and Institutions at the University of Siena. He is in charge of management of the Financial Market Area for the University Master of Economics and Banking in Siena.
†Corresponding author.
‡Antonio Fasano (afasano@unisa.it) earned his Ph.D. at the University of Rome “La Sapienza”. Currently Associate Professor (Professore Associato) of Financial Market and Institutions at the University of Salerno and Rome (LUISS).
1This working paper is the final result of a research study carried by both authors. We can attribute the first part to Claudio Boido and the second one to Antonio Fasano.
able to beat markets to make a profit from their strategy. Each hedge fund follows a different strategy in terms of risk/return. These strategies reflect on the individual manager’s skill and the performances are not linked to the benchmarks as mutual funds. The heavy use of leverage can have a potential effect on expected returns in terms of volatility even if in some strategies, as market neutral, the volatility tends to be lower. Therefore hedge fund returns combine a manager’s skill and the features of the selected strategy which have shown a strong correlation with equity markets than other over the year. Most investors classify hedge funds as alternative assets even if, during the fall of financial markets of 2008, the correlation between hedge fund indexes and equity indexes was approximately 0.82, that is the hedge fund strategies had a bad impact on the financial crisis as well as on traditional asset classes. The explanation lies in the fact that the performance of each strategy is based on the exposure to the asset classes traded on the different markets, that is, when there is a market crisis, the same crisis is reflected on hedge fund returns. Billio et al. [2010] noted that there was an apparent interconnectedness among hedge funds, banks, brokers, and insurance companies during the financial crisis of 2007-2009, which amplified shocks into systemic events. Their results suggested that while hedge funds can provide early indications of market dislocations, they may not give significant contributions to systemic risk as those of banks, insurance companies and brokers. According to Eureka hedge publications, the assets managed by the hedge fund industry declined by $470 billion between June 2008 and April 2009. Fortunately, the flight of investors from the hedge fund industry stopped in May 2009, when inflows into the hedge fund industry exceeded outflows for the first time since June 2008. Given the shock experienced by the industry in 2008, it is certainly interesting to understand if and how the financial crisis altered the determinants of the flows experienced by hedge funds. In the first quarter of 2010, hedge fund–related trading accounted for almost one–third of the average U.S. daily share volume Harris [2010]. In 1990, according to estimates from International Financial Services (2010) some 530 hedge funds managed about $50 billion in assets; by the end of 2009, more than 8,000 hedge funds were managing $1.6 trillion. The strategy mix of the hedge fund industry has also changed. In 1990, the industry was dominated by funds that followed a global macro strategy; in 2008, a sizeable majority of funds followed equity based strategies, such as long–short equity and event-driven approaches. As the events of 2008 have demonstrated, a large part of the average hedge fund beta boils down to straight equity and credit risk. In addition, particularly during the 1990s, the alternative part in hedge fund beta was passed off as absolute return (or alpha), which unreasonably cast a favourable light on the entire hedge fund sector. The growth of hedge funds in the last decade is justified by the following reasons. The first is the unprecedented wealth creation during periods of strong performance in the equity markets, which significantly expanded the base of wealthy private investors. Secondly, institutional investors, together with high-net-worth individuals, started showing greater interest in the hedge fund market. Thirdly, there was a need for effective instruments for diversification during periods of falling equity and bond markets. A “stable absolute return” during periods of different market conditions became more and more a target for both private and institutional investors. Indeed hedge funds have less restriction on the use of constraints as leverage, short selling, and derivatives than more regulated vehicles such as mutual funds.

2 Hedge fund strategies classification

There is not yet a large consensus on a unique classification of the hedge funds. Many researchers follow the strategy class and the specific strategy used by main large index
providers (Hedge Fund Research, CSFB/Tremont, MSCI and Standard & Poor). There are at least eight distinct styles or philosophies of asset management currently employed by hedge funds, and risk exposure depends very much on style affiliation. It can be classified according to a systematic (strategies based on computer program) or a discretionary (strategies based on the opinion and sentiment of the fund managers) approach. Others classify hedge funds according to their geographical location, using this term in a different meaning of Fung and Hsieh, Fung and Hsieh [1997a] that is in a more traditional meaning (Euro area, Emerging markets, U.S and U.K. markets). The large opportunity to classify hedge fund strategies also allows taking into considerations multi-strategies, when it is difficult to understand the risk/return relationship with a unique strategy.

The following classification was used by Eichengreen and Mathieson [1998] and there are 8 categories of hedge funds with 7 differentiated styles and a fund-of-funds category.

- **Event driven funds.** These are funds that take positions on corporate events, taking an arbitrated position when companies are undergoing re-structuring or mergers. For example, hedge funds would purchase bank debt or high yield corporate bonds of companies undergoing re-organization (often referred to as ‘distressed securities’). Another event-driven strategy is merger arbitrage. These funds seize the opportunity to invest just after a takeover has been announced. They purchase the shares of the target companies and short-sell the shares of the acquiring companies. Occasionally, they do the opposite if the deal is likely to fail.

- **Global funds** is a catch-all category of funds that invest in non-US stocks and bonds with no specific strategy reference. It has the largest number of hedge funds. It includes funds that specialise in the emerging markets

- **Global/Macro funds** refer to funds that rely on macroeconomic analysis to bet on major risk factors, such as currencies, interest rates, stock indices and commodities.

- **Market neutral funds** refer to funds that bet on relative price movements using strategies such as long-short equity, stock index arbitrage, convertible bond arbitrage and fixed income arbitrage. Long-short equity funds use the strategy of Jones by taking long positions in selective stocks and going short on other stocks to limit their exposure to the stock market. Stock index arbitrage funds trade on the spread between index futures contracts and the underlying basket of equities. Convertible bond arbitrage funds typically capitalize on the embedded option in these bonds by purchasing them and shorting the equities. Fixed income arbitrage bet on the convergence of prices of bonds from the same issuer but with different maturities over time. This is the second largest grouping of hedge funds after the Global category.

- **Sector funds** concentrate on selective sectors of the economy. For example, they may focus on technology stocks if these are over-priced and rotate across to other sectors.

- **Short-sellers** focus on engineering short positions in stocks with or without matching long positions. They play on markets that have raised too fast and on mean reversion strategies.

- **Long-only funds** take long equity positions typically with leverage. Emerging market funds that do not have short-selling opportunities also fall under this category.

- **Fund of funds** refer to funds that invest in a pool of hedge funds. They specialize in identifying fund managers with good performance and rely on their good industry relationships to gain entry into hedge funds with good track records.

On a risk-adjusted basis (dividing the mean return by the standard deviation), the category of fund that ranks highest is the market neutral funds followed by event-driven funds.
This classification is based only on self identification and it does not consider the risk and return profile of each hedge fund.

3 Related Literature

Many researchers have proposed different kinds of classification of hedge fund strategies, seeking to understand the guidelines of their investments.

Fung and Hsieh [1997b] have classified hedge fund strategies following “style” (that is the type of position the fund manager is taking) and “location” (that is the asset class where hedge funds invest, for example equity, fixed income, commodities, currencies). Amenc and Martellini [2003] have proposed a distinction between “return enhancer” and “risk reducer strategies”. In the first group we find hedge funds which seek higher expected returns but, according to the portfolio selection, they increase the volatility (that is distressed securities, event driven, macro funds). In the second group active fund managers want to obtain positive excess returns reducing portfolio volatility (that is convertible arbitrage, fixed income arbitrage, long/short and short selling funds).

In literature we can find some researchers who attempt to explain hedge fund returns on the type of asset classes and the styles of hedge fund managers.

Schneeweis, Spurgin, and McCarthy [1996] use a large set of indicators (returns, volatility by equity index, interest rates, exchange rates, commodity prices and future contracts) to show that hedge funds depend on different sets of factors compared to traditional asset classes (bond and equity). The same method is applied by Martin [2002], according hedge funds to a specific strategy through a cluster analysis. He uses a regression analysis to ascertain a link between the performance of the different strategies and the selected economic factors. His results show that a significant correlation exists when each strategy is pooled as an index.

Brown and Goetzmann [2001] examine the risk/return relationship of each hedge fund strategy cluster and so they use a cluster analysis algorithm to identify hedge fund strategy groups through a likelihood based ratio test to identify the optimal number of groups.

Asset allocation management style developed by Sharpe [1992] shows some problems in its application to hedge funds returns:

- It’s not clear which the right market indices as explanatory variables are;
- Hedge fund managers use dynamic strategies, so the standard regression model is not valid because asset exposures change quickly;
- Sharpe ‘s coefficient restrictions are not adequate for hedge funds, that is leverage and derivative allow the exposures to go over the limit.

Agarwal and Naik [2000] insert, as explanatory variables, options to improve the non-linearity in returns. Authors associate location, trading strategies and leverage, identifying the sources of returns (as Fung and Hsieh [1997a,b]). A similar approach is followed by Mitchell and Pulvino [2001] who run a regression of returns on an equity market index. Brown and Goetzmann [2001] find that a difference in investment style explains about 20% of the cross sectional variability in hedge fund performance and they confirm that the Sharpe model is used to measure exposure to investment management style. In a more recent study, Bali, Brown, and Caglayan [2011] Gregoriou and Pascalau [2012] argued that funds of hedge funds are, in general, over diversified.

---

2 This is a special application of regression analysis of fund returns, where the explanatory variables are returns on a set of market indices. It states that the regression coefficient is non negative and the sum of the coefficients must add up to one. In this way the coefficient on all markets indices can be understood as the percentage of the portfolio selected to that asset class.
Shawky, Dai, and Cumming [2012] investigate a significant relationship between diversification and performance and they find a considerable positive relationship across sectors and asset classes, while diversification across styles and location show a significant negative association with hedge fund returns.

Even before the financial crisis, as written by Brown, Goetzmann, and Ibbotson [1999], Liang [2000], Agarwal and Naik [2000], and Kosowski, Naik, and Teo [2007], the failure risk in the hedge fund industry is high and rather persistent. Moreover, Kim [2008] examined failure risk in the hedge fund industry and found that hedge funds and funds of hedge funds fail as poor performance over an extended period of time (typically six to nine months) leads to significant withdrawals by investors. Cao et al. [2011] show that hedge funds adjust their market exposure in the light of changing market liquidity, which potentially also affects their returns. Moreover, Sadka [2010] documents that unexpected changes in stock market liquidity are a priced risk factor in the cross-section of hedge fund return.

Finally, Boyson, Stahel, and Stulz [2010] find that large declines in stock market liquidity can be contagious to hedge fund returns. Mancini, Ranaldo, and Wranglemeyer [2011] show that foreign exchange liquidity decreased significantly during the 2007–2009 period and that FX liquidity is correlated with measures of equity liquidity. A number of papers have argued that hedge funds that restrict investor liquidity perform better than hedge funds that do not restrict investor liquidity, including Schneeweis, Kazemi, and Martin [2002], Aragon [2007], Aragon, Liang, and Park [2008], Ding et al. [2009], and Agarwal, Daniel, and Naik [2009].

Although there is some evidence that hedge funds that restrict investor liquidity did better during an earlier time period (1994–2001), this result completely disappears during a more recent time period (2002–2009) and does not hold for the 1994–2009 sample period as a whole.

The literature on hedge funds characteristics is too large to be covered in detail here. The consensus so far is that the following hedge fund characteristics are related to performance.


Many hedge fund indices have returns that are far from normally distributed, with negative skewness and a high kurtosis. In addition, many hedge funds claim that their return distributions are not normal due to the use of derivatives, short sales, and rebalancing rules. However, it is now well understood that hedge fund returns can be highly non-normal, which makes traditional performance measures, based on the assumption of normality, unsuitable.

Several researchers (Fung and Hsieh [1997a,b], Fung and Hsieh [2004], Mitchell and Pulvino [2001], Brooks and Kat [2002], Lamm Jr [2003], Agarwal and Naik [2004], Malkiel and Saha [2005], Eling and Schuhmacher [2007], Perelle [2007], Kat and Mifre [2008] have shown the inadequacy of mean variance framework for measuring the performances of hedge fund portfolio, because it is valid only if the returns are normally distributed and that the distributions of returns
are asymmetrical. Other authors (Bacmann and Pache 2003, Liang and Park 2007, Perelló 2007) have proposed the use of downside risk measure (Sortino ratio, Omega measure, Upside potential ratio) as a solution to the asymmetry of returns to value the performance of hedge funds. As regard the newer approaches we can single out:

- Measures based on lower partial moments (Omega, Sortino, Kappa)
- Measures based on drawdown (Calmar ratio, Sterling ratio, Burke ratio)
- Measures based on value at risk (Excess return, Conditional Sharpe ratio Modified Sharpe ratio)

Some authors (Eling and Schuhmacher 2007) conclude that the choice of performance measurement has no impact on the ranking of hedge funds and that the use of Sharpe Ratio, based on the first two moments (i.e. mean and variance) provide an adequate description of the return distribution.

**Part II**

4 Alternative Risk and Traditional Risk

The comparison among style returns is meaningful in as far as we take into account the different risks the investors take for the different fund classes and the excess risk involved — if any — when moving from a traditional to an alternative portfolio.

Indeed, a common inquiry for the HF investor would be how those funds, as an alternative investment, compare with a traditional equity investment. Therefore we will first analyse inter-style performance to identify the possible different style-dynamics, and subsequently we will risk-adjust this performance and designing an *ad hoc* alternative risk adjusted measure.

By doing so, we will try to answer the question concerning whether the risks taken and subsequent results are randomly drawn or they can be considered the effect of a systematic and disciplined management policy.

The performance variability among fund groups, expressing the different managers’ investment policies and also proxing their skills, can be measured by means of F-tests, where one assumes as the null-hypothesis that on average groups are indistinguishable.

4.1 Alternative Risk: the stylised facts

Alternative investments have a reputation for showing aggressive, or however different, return patterns when compared with traditional asset classes. Anyway it bears noting that data upon which time series are built are self-reported by managers, which entails some well-known biases.

While addressing the reader to Ackermann, McEnally, and Ravenscraft [1999], Liang [2000], Fung and Hsieh [2000], Fung and Hsieh [2002] for a detailed analysis, we here point out three critical biases whom an investor (or a scholar) should be aware of:

- **Self-selection bias**: Data base inclusion is on a voluntary basis and manager tend to include only those funds with a good track record.

---

3 We suggest to consult the original paper (Sortino and Price 1994, Sortino, Meer, and Plantinga 1999, Keating and Shadwick 2002) for more details.

4 Eling and Schuhmacher 2007 for more details.
• **Backfill bias**: It is also named “instant history” bias and is a consequence of the previous: when a funds starts showing a good track record, managers include them backfilling the data base with the previous records.

• **Survivorship bias**: Here we have the other side of the self-selection: fund indices are normally based only on operating funds so the bias arises as unsuccessful funds are more likely to close. The phenomenon is amplified by the practice of “driven deaths”, this happens when a fund is performing badly and managers close it by moving its assets to another fund.

With these issues in mind, we will now review data from a sample of funds over the period January 2006–December 2012\(^5\).

The geographic scope of the sample dataset is shown in Figure 1. Despite for a somewhat high number of funds the country is not reported, data confirm that many funds prefer to incorporate overseas for reasons possibly ascribable to legal shields and tax exemption.

Figure 2 builds upon Figure 1, breaking down the sampled universe by denominations of currency. Primary denominations are Euro and US Dollar. For better interpretation of these data, it might worth noting that the British Virgin Islands and Cayman Islands are British overseas territories, but the GB Pound is not their legal tender: the former prints its own currency, the latter use the US Dollar. Therefore, despite the huge number of companies domiciled in these countries, the GB Pound still has a relatively low weigh, if compared with Euro or US Dollar.

---

5The data base by which the sample was extracted, kindly provided by MondoAlterntive, consists of 990 funds. It is a global HF database, not targeted toward a specific geographic area or purpose.
It bears mentioning that, while hedge fund are often seen as a single investment class, they are in fact a very broad one, so they are normally distinguished by the investment policies they apply, regarded as strategies and styles. Hence, it makes sense to investigate how the different policies compare to one another on the performance point of view, and particularly during phases of market crisis.

In the context of investment policies the two terms ‘Strategy’ and ‘Style’ can be used in a rather loosey-goosey way. In this work we will denote as strategies those formally declared by the management and reported in the fund database records; while styles will denote clusters of strategies recognised as similar by the practitioners. As it can be easily guessed, criteria for segregating hedge funds into standardised investment styles are quite challenging, since there is no formal consensus on a single taxonomy and research institutes tend to provide their own categories.

In this regard Table (1) shows, the reported strategies found in our data set, after specific outlier treatment (see section (4.2)). These strategies were further grouped by style. Despite we believe that the asset allocation discipline creates natural links among some strategies, the classification remains subjective.

In particular, as for the funds of hedge funds, their assets are themselves hedge funds (which can have the same investment style or even different ones), therefore they are not an actual investment style, but we took the liberty to group these funds as a style on its own to simplify the reporting of research output.

Whatever the style (or strategy), a defining hallmark of the hedge funds, as an alternative investment class, are the common active, if not ‘aggressive’, management policies. Both the supplementary costs and the expected benefits of these policies are supposedly charged to the investors in terms of fees. To check out informally these claims, Figure (3), based on the style clusters set in Table (1), shows fee structures distributed by style and currency. Fees are intended as performance fees, calculated as a percentage of the increase in the gross
4.2 Methodology for Outlier Detection and Management

The biases already described in the previous section might easily yield to a fast and possibly not faithful switching return dynamics, which will be observed in the form of outliers in the historical records.

Apart from this, there is a more technical reason beyond the non-smooth dynamic of hedge funds, documented by Lo [2005] as “phase-locking” behaviour. This behaviour, well known in the physical and natural sciences, is a situation when two normally uncorrelated pattern suddenly become perfectly synchronized. In financial market the dramatic switch in correlations can be the consequence of market crises. Hedge funds exhibit non linearity and are exposed to phase-locking behaviour, which in turn will involve sudden jump in returns, observable as outliers.

Finally, the lack of regulations and reporting standards often determines a change of scales in fund NAVs, which is observable in terms of dramatic outliers. This effect can be easily checked out by a human reading of the observation records, but is not manageable with ordinary statistical procedures.
Following this, we now describe the procedure used to produce the workable time series. First, we selected a continuous time span, that is, we selected only those funds having a continuous track record from January 2005 through December 2012 (thus the data set was deeply downsized to obtain comparable time series). Given this criterion, to deal with backfill bias, the actual time span used started from January 2006 (so previous observation where discarded, while selected funds were still required to report them). As for survivorship bias, by the scope of this analysis, we are interested in those funds actually available to the investors during the whole time covered by our analysis, that is, how they perform provided that they do (see Dor, Berkovitz, and Xu 2012 for similar arguments).

Given this preliminary data set, a popular option for outlier treatment is the extreme studentised deviate, more familiarly spelled as the “three-sigma edit rule”, that is, any point distant more than $t$ standard deviations from the mean of its neighbours is an outlier. Formally, an outlier does not belong to the interval:

$$[\hat{\mu} - t\hat{\sigma}, \hat{\mu} + t\hat{\sigma}]$$

where $\hat{\mu}, \hat{\sigma}$ are the sample mean and deviation of the reference neighbourhood.

For normal data, it is customary to take 3 as the threshold value $t$. In fact, for a normal distribution, observing values more than three standard deviations away from the mean is only 0.3% likely. Detected outliers are usually replaced with their reference mean ($\hat{\mu}$).

There is a significant bias with this approach: the mean, and even more the standard deviation, are themselves distorted by outliers!

As an alternatives to the “local three-sigma” rule one can use the Median Absolute Deviation:

$$M_X = \text{median} \left( |X - \text{median}(X)| \right), \quad (1)$$
where \( m_X \) is the median of \( X \).

In this case, (1) uses the median of neighbouring observations as a reference value, and uses the deviation from the median (MAD) as an alternative measure of distance. Given (1), an outlier is such when lying more than \( t \) times the MAD from the median of its neighbours. Such a filter is commonly addressed as Hampel filter (see Pearson [2011]).

In this research we used a version of the standard Hampel filter slightly modified as follows:

- After outlier detection, the median was replaced with the average observations before and past the outlier;
- The filter was adjusted at the borders of the series, to avoid cutting data;
- The MAD was adjusted by a factor of about 1.5, as it can be shown that for normal data:
  \[
  \sigma_X \approx 1.4826 m_X
  \]

Moreover the cleaning stage involved removing funds with missing values or incoherent values (e.g. negative NAVs). We also did not include “Asset based lending” and “Mortgage-Backed Securities” funds, since beyond the scope of this research.

### 4.3 Aggregating Funds by Investment Policy

In this section we detail the methodology used to aggregate returns.

Since every fund has a reported strategy (and only one), let the hedge fund \((i, j)\) denote \( i \)-th fund among those reporting the \( j \)-th strategy. Therefore \( r_{ij}(t) \) denotes the monthly return observed on month \( t \)-th (of the time series) for the \((i, j)\) fund. With reference to the time window \( T = t_b - (t_a - 1) \), the average return for the \((i, j)\) fund will be:

\[
\tilde{f}_{ij}(t_a, t_b) = F_{ij}(T) = \frac{\sum_{h=a}^{b} r_{ij}(t_h)}{T}
\]

In the same period the average for \( j \)-the strategy will be:

\[
s_j(t_a, t_b) = S_j(T) = \frac{\sum_{i=1}^{n_j} F_{ij}(T)}{n_j} \frac{N_S}{N_F}
\]

where \( n_j \) is the number of funds comprising \( j \)-th strategy.

The average for all funds and strategies during \( T \) are resp.:

\[
F(T) = \sum_{i,j=1}^{n_j,n_s} \frac{F_{ij}(T)}{N_F}; \quad S(T) = \sum_{j=1}^{N_s} \frac{S_j(T)}{N_s} \frac{N_S}{N_F}
\]

where \( N_F, N_S \) are the total number of funds, strategies. Obviously \( N_F = \sum_{j=1}^{N_s} n_j \).

The average for all funds during \( T \) does not necessarily equal to the average for all strategies, unless the latter are equinumerous, in fact assuming every \( n_j = n \):

\[
S(T) = \sum_{j=1}^{N_s} \frac{\sum_{i=1}^{n} \frac{F_{ij}(T)}{nN_s}}{N_s} = \sum_{j=1}^{N_s} \sum_{i=1}^{n} \frac{F_{ij}(T)}{nN_s} = \sum_{j=1}^{N_s} \frac{F_{ij}(T)}{N_F} = \sum_{j=1}^{N_s} S_j(T)
\]
As for styles, given the previous strategy-style mapping (see Table 1), the styles can be thought as a partition of the set $A$ of all strategies. If $M_k \subset A$ is the set of the strategies in $A$ mapped to the $k$-th style, the average return for this style, with reference to the time window $T = t_b - (t_a - 1)$, is:

$$q_k(t_a, t_b) = Q_k(T) = \frac{\sum_{i=1}^{n_j} F_{ij}(T)}{m_k}$$

where $m_k$ is the number of strategies mapped to the $k$-th style.

As seen for strategies, if the style sets, $M_k$, have the same size, the average returns for all funds during $T$ equals the average among styles.

As the hypothesis of equinumerosity is not consistent with our data set, for both strategy and style partitions, it makes sense to compare the average returns obtained by aggregating the funds. Therefore Figure 4 compares the annualised returns for grouped and non-grouped funds. Similarly Table 2 shows the correlation between the series obtained with the different grouping levels. Both the plots and the figures show the differences in the average returns are negligible.
Table 2: Measuring the Grouping Effect: Correlation

<table>
<thead>
<tr>
<th></th>
<th>Ungrouped</th>
<th>By Strategy</th>
<th>By Style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ungrouped</td>
<td>1.00</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>By Strategy</td>
<td>0.95</td>
<td>1.00</td>
<td>0.96</td>
</tr>
<tr>
<td>By Style</td>
<td>0.98</td>
<td>0.96</td>
<td>1.00</td>
</tr>
</tbody>
</table>

4.4 HF (Reported) Returns: notes and descriptive statistics

To give the reader a preliminary knowledge of returns’ dynamic, we will show some descriptive statistics related to the return time series obtained by means of the monthly reported NAV, after applying the outlier treatment from section (4.2).

As observed in section (4.3), despite average performance for all funds will not necessarily be equal to the average among styles or strategies, for our data set return time series proved almost overlapping visually and highly correlated analytically. So we can measure overall return starting from grouped funds with negligible bias.

We start by analysing spatial and serial correlations resp. with Figures (5). Some styles appear highly uncorrelated each other leaving large room to diversification opportunities. On a monthly basis all styles exhibit some serial correlations too, except for ‘Futures’ and ‘Global Macro’.

Figure (7) shows distributions of style returns for the time range scrutinised by means of notched boxplots. On average, over the whole period January 2006–December 2012, styles exhibit good performances, with a limited shortfall area. ‘Futures’ and ‘Multistrategy’ have a null median, but positive returns tend to outweigh negative ones. Indeed, by looking at the boxplot medians, we can say that returns are not normal (as the theory suggests) and skewed toward the positive side.

4.5 Crisis Window

When investigating the risk-return dynamics, the importance of market crashes on the outcome of the different investment policies is easily agreed upon, but settling the proper crisis window correspondingly is not straightforward. If we want to mold our analysis in a pre and post crisis fashion, it is crucial to have an objective identification of the crisis window, in order to avoid a bias due to an ad hoc definition of the crisis.

With reference to the period scrutinised, a number of events can be considered as the outbreak of the crisis: most notably the collapse of the investment bank Bear Stearns in March 2008, which triggered a contagion overwhelming several large financial institutions (including Lehman Brothers, Merrill Lynch, Fannie Mae, Freddie Mac, Wachovia, Citigroup).

Anyway for a formal, non-subjective, definition we resolved to refer to the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER). The Committee maintains a business cycle chronology with reference to the U.S. It defines a “recession” as any period occurring between a ‘peak’ and a ‘trough’. According to September 2010 meeting:

“A trough in business activity occurred in the U.S. economy in June 2009. The trough marks the end of the recession that began in December 2007 and the beginning of an expansion. The recession lasted 18 months, which makes it the longest of any recession since World War II.” (Committee 2010)

Despite our analysis is globally scoped, the business cycle reference stems from US economy. We believe this choice is sounder than the identification of a global, but not as much
Figure 5: Cross correlation among style returns (monthly). The upper triangle shows numerical values. To increase readability of the results, the more the correlation among styles the more the font dimension. The inferior triangle shows the same data in the form of dispersion plots. The fitted line (red) for each plot helps in assessing visually the degree of correlation between styles.
Figure 6: Correlograms of monthly return time series for each style. The auto-correlations allow us to assess memory effects, that is the degree of correlation of present values of the series with past ones. The correlations are reported for different time lags and they can be considered negligible when contained between the horizontal bars.
Figure 7: The overall shape of returns is shown for each style as a notched boxplot. Each notched box extends between the quartiles so that it contains the middle half of the data, as such the more its height the more the variability of the style performances. The line in the middle of the box represents the median, and from its position one can assess the distribution symmetry. “Whiskers” extend outside the box for 1.5 times the inter-quartile range, and they let us identify easily extremal values, which are located outside of the “fences”.

objective, time window. At any rate, since any analysis involving time series is affected by the selected time window, some specific tests will be undertaken to assess the impact of differing time spans, in order to confirm that the crisis period is actually noteworthy.

By applying NBER business cycle findings, we are able to plot the charts in Figures (8) and (9). The visual clues we get by those plots show a heterogeneous reaction to cycle switching, therefore demanding the supplementary investigations following ahead.

5 Performance Tests

How different are the simple performances (i.e. without taking into account the risk) between the different management policies? And particularly how different are under stressed market conditions? As we have observed above (cf. Figures (8) and (9)) performance patterns are rather heterogeneous among styles and possibly amplified during the crisis window. In as far as the good performance of management styles is a proxy of management skills, talented managers could make a higher difference just when the market conditions are tougher and even the crisis can pose profit opportunities for those competent investors capable to time the market.

To test in a formal way if the alleged differences are true we turn to statistical inference. Using an ANOVA analysis, we check if the differences among groups are significant. We will replicate this analysis with reference to the crisis cycle, to test if results held or possibly

---

6All statistical tests implemented in this research were carried out via the R language. Main source code is available for interested readers in the Appendix of this paper.
Figure 8: Average Style Performances with respect to NBER crisis period — December 2007 throughout May 2009.
Figure 9: Performance dynamics with respect to NBER crisis period — between the red lines. The dashed line is the overall mean.
More formally we employ an F-test where we assume, as the null-hypothesis, that on average strategies are indistinguishable, this is to say that mean returns are the same. In Table (3) we show the result of this test. Here MSB can be considered as a measure of the performance variability among strategies and this variability can be seen as an effect of the diverse management strategies. As usual meaningfulness is assessed by means of p-value, involving rejection of the null hypothesis.

In term of p-value there is a clear effect of the crisis on the variability of the returns: market stress implies a better capacity to distinguish one management policy (strategy) by another.

As the NBER-identified crisis window lasted for 18 months, we further analyse the moving average over 18 months for the whole data set period. In Figure (10) the black line is the average annualised return for each month, while the red line — for each month — is the moving average starting in that month. As the red plot shows, the worst 18-month period — in term of hedge fund performance — is compatible with NBER definition.

By Table (3), we can assess that the variability during the NBER-window is higher than that related to the whole period, but can an equally significant result be obtained also for a different time window? To check this hypothesis we take all possible 18-months time windows for the data set period and run the same test for all them. The battery of tests provide us with a p-value for each period; therefore in Figure (11) we present, for each 18-months window, both the average observed return and the p-value assessing the test significance (i.e. the level of return variability among strategies for the given period). To plot visually the p-values, we plot $-\log_{10}(p)$. In this way, a value like ‘2’ can be considered as a significant level.

Figure (11) partly confirms previous results, but also adds new insights. The p-values follow the trend of the related period and they appear connected to both upward and downward peaks. The policy difference is significant in cycles which show non-average returns, in both cases the management policies can make difference in results.

<table>
<thead>
<tr>
<th></th>
<th>OMR</th>
<th>MSB</th>
<th>F</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 2006 - Dec 2012 (84) Whole range</td>
<td>0.0046</td>
<td>0.0007</td>
<td>0.808</td>
<td>0.66066</td>
</tr>
<tr>
<td>Dec 2007 - May 2009 (18) Crisis only</td>
<td>-0.0053</td>
<td>0.0042</td>
<td>3.202</td>
<td>0.00011</td>
</tr>
</tbody>
</table>

Table 3: The performance variability among style is measured by means of an F-test. As usual the OMR is the overall mean and MSB the means squares between groups, with reference to monthly return.

6 Alternative Approach to Risk

As it was pointed out, the comparison among policy returns is financially meaningful only if backed by the risk born by each policy. This suggested to build a measure embedding both the excess premium and risk possibly involved in an alternative investment.

To this end, we introduce a measure, denoted Alternative Excess Return (AER), trying to assess the premium of the alternative investment over a traditional one, while weighting the extra risk carried. In schematic terms:

$$\frac{A - T}{DD}$$

Here the numerator is the excess return of the alternative return ($A$) over the traditional return ($T$); the denominator is the “dynamic downside risk” of the premium ($DD$). It is the
Figure 10: Average and Moving Average Strategy Returns.
Figure 11: Comparing p-values over Different Time Lags.
Figure 12: Flat AER histogram, that is without the risk factor. Each overlapping histogram is the AER distribution of a given colour-coded style.

risk of the alternative return being below the traditional return. As the latter is non-static, the downside is not absolute, but it is relative to the market condition: in a bear market it is sufficient not do as bad as the (traditional) market; but in a bull market a downside is measured when the alternative asset is unable to get returns as high as a traditional portfolio.

We tested historically this measure, proxing the traditional returns with the S&P500 index.

Figure 12, 13 show the distribution of the flat AER premia, that is without downside risk, by style. As it appears, there is a huge variability among styles.

We checked the significance of the alternative premium with two statistical tests: a binomial test, where the “success” is defined as obtaining a positive AER instead of a negative one; a Student t-test, measuring the intensity of this success. Therefore the former is influenced only by the sign of the premium and the latter by its dimension too.

Table 4 shows the results of these tests. As we see, for some styles during the period under investigation, the alternative premium is positive, even considering the risk factor, besides the average mean is significant. In terms of t-test, i.e. of “intensity” of the premium, two styles, Event Driven and Equity Hedge, show a significant better risk adjusted performance with respect to traditional investment.
Table 4: The significance of the AER measure by style. The first two columns report a binomial test of AER positivity: they are resp. the realised percent of “success” and its significance. The following columns measure the intensity of AER: they are resp. the mean AER and its significance.
7 Conclusions

The aim of this study is to analyse the risk-return dynamic of hedge fund alternative investments also in the presence of distressed market conditions and with reference to the crisis window set by NBER Business Cycle Dating Committee. We also distinguished the analysis by the management policies applied, despite there is no formal taxonomy in this regard. The latter can be intended also as a proxy of the management portfolio selection skills.

After analysing the management policies of hedge funds, we introduced a new metric for assessing the alternative investment premium over the traditional assets. First we found that the crisis had the effect of emphasizing the differences among the management policies. Under market pressure management styles are not all equal, the different approaches turn into different possibility of ruin, safety, or even profit.

To better assess the quality of these policies we developed a measure of the excess return gained going alternative (AER), scaled with the downside risk dynamically targeted to a traditional investment portfolio. The alternative investment is rewarding, for some styles, that is, we found that the AER is historically positive in mean and statistically significant.
References


A Main R Code

# nav is a zoo object with fund NAV's strategy, style char vector with
# strategy, styles for each fund

# 1) Set STARTDATE and ENDDATE opportunistically as to get less NAs 2)
# Select funds already reporting 1 year before STARTDATE (cleandate) 3)
# Clean data from NA's 4) Select obs one month before STARTDATE (to get the
# first return) Step 2 is to avoid fund raising effect (backfill):

# 1) Set STARTDATE and ENDDATE
STARTDATE = "Jan 2006"
ENDDATE = "Dec 2012"
dim(nav)

# 2) Select funds reporting at least since cleandate and up to ENDDATE
cleandate = as.yearmon(STARTDATE) - 1  #Backfill effect

nav = nav[time(nav) >= cleandate]
nav = nav[time(nav) <= ENDDATE]
dim(nav)

# 3) Clean data from NA's count NAs write.table(apply (is.na(nav), 2,
# function(x) length(which(x)))) Select funds (cols) if not NA
x = apply(!is.na(nav), 2, function(x) all(x))
nav = nav[, x]
dim(nav)

# Select funds with NAV>0 (some report NAV==0)
x = apply(nav > 0, 2, function(x) all(x))
nav = nav[, x]
dim(nav)
rm(x)

# 4) Select obs one month before STARTDATE
nav = nav[time(nav) >= as.yearmon(STARTDATE) - 1/12]
dim(nav)

# After clean the fund returns we need to align/trim other series
# accordingly

# Align market indices to trimmed fund dates
sp500 = sp500[time(sp500) >= as.yearmon(STARTDATE) - 1/12]
sp500 = sp500[time(sp500) <= ENDDATE]
dim(sp500)

ret = diff(nav)/coredata(nav[time(nav) < end(nav)])
dim(ret)

ret.m = diff(sp500)/coredata(sp500[time(sp500) < end(sp500)])
dim(ret.m)

# number of tolerated outliers
out.tol = 6
# A revised Hampel filter

```r
hampel2 = function(x, k, t0 = 3) {
  n <- length(x)
  y <- x
  ind <- c()
  L <- 1.4826
  for (i in 1:n) {
    a = ifelse(i < k + 1, 1, i - k)
    b = ifelse(i > n - k, n, i + k)
    x0 <- median(x[a:b])
    S0 <- L * median(abs(x[a:b] - x0))
    if (abs(x[i] - x0) > t0 * S0) {
      a = ifelse(i == 1, 2, i - 1)
      b = ifelse(i == n, n - 1, i + 1)
      y[i] <- mean(c(x[a], x[b]))
      ind <- c(ind, i)
    }
  }
  list(y = y, ind = ind)
}
```

# Count outliers

```r
out.count = apply(ret, 2, function(x) length(hampel2(x, 2)$ind))
```

# Accept only `out.tol' outliers or less

```r
ret = ret[, out.count <= out.tol]
```

# Fix remaining outliers via Hampel

```r
x = apply(ret, 2, function(x) hampel2(x, 2)$y)
coreda(ret) = x
dim(ret)
```

# After clean the fund returns we need to align/trim other series accordingly

# Align descriptions to trimmed fund number

```r
fund.names = names(ret)
x = des$name %in% fund.names
des = droplevels(des[x, ])
dim(des)
```

# Aggregate returns by style and strategy and compute mean for each

# Style and Strategy attributes

```r
fstrat = factor(des$strat) #strategies as a factor
strat.names = levels(fstrat)
strat.num = length(strat.names)
fstyle = factor(des$style) #styles as a factor
style.names = levels(fstyle)
style.num = length(style.names)
obs = NROW(ret)
```
# put funds on rows to match des obj
xx = data.frame(t(coredata(ret)))

# Get average for each strategy
x = aggregate(xx, list(Strategy = fstrat), mean)
strat.ret = zoo(t(x[, -1]), time(ret))
names(strat.ret) = x[, 1]

# Get average for each style
x = aggregate(xx, list(Style = fstyle), mean)
style.ret = zoo(t(x[, -1]), time(ret))
names(style.ret) = x[, 1]

# Standard F-test for assessing differences among groups
ftest = function(f) {
  # Group num
  (ngroup = ncol(f))

  # Observations per group
  (nobs = nrow(f))

  # Total observations
  (Nobs = ngroup * nobs)

  # Within group mean
  (ym = apply(f, 2, mean))

  # Overall mean
  (ymm = mean(ym))

  # Sum of squares between
  (SB = sum(nobs * (ym - ymm)^2))

  # Between-group degrees of freedom (number of groups -1)
  (freebet = ngroup - 1)

  # Between-group mean square
  (MSB = SB/freebet)

  # Centering the data in each group, i.e. get deviations
cg = as.matrix(f) - matrix(1, nobs) %*% ym
  # matrix(1,nobs) means £1_{\text{nobs}}£ vector

  # Sum of squares between
  (SW = sum(cg^2))

  # Within-group degrees of freedom
  (freewit = Nobs - ngroup)

  # Within-group mean square
  (MSW = SW/freewit)

  # F-Ratio
  (F = MSB/MSW)
}
# p-value
pval = 1 - pf(F, freebet, freewit)
data.frame(OMR = ymm, MSB, F, pval)

# F test output in a data frame with labels
ftest.lab = function(dat, com) {
  # create a start/end range label
  lab = paste(start(dat), "-", end(dat))

  # add range length
  x = length(seq.Date(as.Date(start(dat)), as.Date(end(dat)), by = "months"))
  lab = paste0(lab, " (", x, ")")

  # add comment if any
  lab = paste(lab, com)

  # output as labeled row
  data.frame(ftest(dat), row.names = lab, check.names = F)
}

# Test performance differences among styles or strategies
# ==============================================================

# Set crisis window value ex nber
crisis.start = "Dec 2007"
crisis.end = "May 2009"

# Test return diffs on several subperiods
# -------------------------------------------------------------
retTestf = function(retm) {
  # Stack results of f tests here
  parstack = data.frame()

  # test on all available observation range
  x = ftest.lab(retm, "Whole range")
  parstack = rbind(parstack, x)

  # test on obs with crisis only
  retm.cr = retm[time(retm) >= crisis.start & time(retm) <= crisis.end]
  x = ftest.lab(retm.cr, "Crisis only")
  parstack = rbind(parstack, x)

  # test on obs before crisis


```r
retm.bc = retm[time(retm) < crisis.start]
x = ftest.lab(retm.bc, "Before crisis")
##### parstack= rbind (parstack, x)

# test on obs after crisis
retm.ac = retm[time(retm) > crisis.end]
x = ftest.lab(retm.ac, "After crisis")
##### parstack= rbind (parstack, x)

# Test on a random subsample of returns with length = obs in crisis

# Get lengths
n = NROW(retm.cr)
nn = NROW(retm)

# Vector of start dates
test.rows.start = 1:(nn - n + 1)
# exclude crisis period
test.rows.start = test.rows.start[-which(time(retm) == crisis.start)]
for (i in test.rows.start)
{
y = seq(i, length = n)
  # test on sample rows and stack result
  x = ftest.lab(retm[y], "")
  ##### parstack= rbind (parstack, x)
}

names(parstack)[4] = "Pr(\textgreater F)"
parstack
```

```
retTestf.s = function(retm, n) {

  # Stack results of f tests here
  parstack = data.frame()

  # Test on a random subsample of returns with length = obs in crisis

  # Get lengths retm.cr= retm[time(retm) >= crisis.start & time(retm) <= crisis.end ]
  n =NROW(retm.cr)

  nn = NROW(retm)

  # Vector of start dates
  test.rows.start = 1:(nn - n + 1)
  # exclude crisis period
  test.rows.start = test.rows.start[-which(time(retm) == crisis.start)]
  for (i in test.rows.start)
  {
    y = seq(i, length = n)
    # test on sample rows and stack result
    x = ftest.lab(retm[y], "")
    parstack = rbind(parstack, x)
  }

  names(parstack)[4] = "Pr(\textgreater F)"
  parstack
}
```
# Test significance of alternative extra returns
# ==============================================
# vec=function(x) as.matrix (c (as.matrix(x) ))

# Success tests
stest = function(ret.f, nam) {
    er = ret.f - ret.m$Close
    # er = vec(er)
    erplus = er[er > 0]
    size = length(er)
    x = length(erplus)
    btest = binom.test(x, size, 0.5, alternative = "greater")
    ttest = t.test(ret.f, ret.m$Close, alternative = "greater")
    data.frame(`%success` = length(erplus)/size, `Bin. p-value` = btest$p.value,
               `Mean AER` = mean(er), `t p-value` = ttest$p.value, row.names = nam,
               check.names = F)
}

# ====== Style AER tests ======
# Stack results of tests here
parstack = data.frame()

# Loop trough styles
for (i in 1:NCOL(style.ret)) {
    x = stest(style.ret[, i], names(style.ret)[i])
    parstack = rbind(parstack, x)
}

x = order(parstack$`t p-value`)
parstack = parstack[x, ]

# Test significance of Risk Adjusted AER
# =======================================
# Semi-deviation
sv = function(ret.f) {
    m = mean(ret.f)
    r = ret.f[ret.f < m]
    sqrt(sum((r - m)^2)/length(r))
}

# Success tests with Risk Adjusted AER
stestr = function(ret.f, nam) {
er = ret.f - ret.m$Close
er = er/sv(ret.f) - ret.m$Close/sv(ret.m$Close)
erplus = er[er > 0]
size = length(er)
x = length(erplus)
btest = binom.test(x, size, 0.5, alternative = "greater")
x = ret.m$Close/sv(ret.m$Close)
ttest = t.test(ret.f/sv(ret.f), x, alternative = "greater")
data.frame(success = length(erplus)/size, "Bin p-value" = btest$p.value,
`M. RA-AER` = ttest$estimate[1] - ttest$estimate[2], `t p-value` = ttest$p.value,
row.names = nam, check.names = F)

# ==== Style AER tests ====

# Stack results of tests here
parstack = data.frame()

# Loop through styles
for (i in 1:NCOL(style.ret)) {
    x = stestr(style.ret[, i], names(style.ret)[i])
    parstack = rbind(parstack, x)
}

x = order(parstack$`t p-value``)
parstack = parstack[x, ]