Original Article

Hedge fund replication

Received (in revised form): 31st May 2011

John Grummitt
is a quantitative analyst at Infinity Capital Markets.

Steve Satchell
is a fellow of Trinity College Cambridge University, UK and visiting professor at Sydney University, Australia. His specialities are finance and econometrics, on which he has written at least 80 papers. An academic advisor and consultant to a wide range of financial institutions covering such areas as actuarial valuation, asset management, risk management and strategy design. Dr Satchell has doctorates from the Universities of Cambridge and London.

Correspondence: John Grummitt, Infinity Capital Markets Ltd, Salters Hall 4 Fore Street, London EC2Y 5DB, UK
E-mail: J.Grummitt@infinitycm.co.uk

ABSTRACT **In this article, we look at hedge fund replicators. These are products sold by financial institutions, which aim to match the returns of hedge funds. Hedge funds are able to make a positive return even in falling markets; however, they charge high fees and often require investors to maintain their investment in the fund for a certain minimum period. Hence, having a liquid replicator fund with low fees and the same returns as the hedge fund may be a preferable investment to holding the hedge funds themselves. We look at the some of the hedge fund replicators which have been proposed in the literature, focusing on the replicator proposed by Hasanhodzic and Lo in their 2006 paper, we repeat their analysis for the more volatile period of the previous few years. We find a reasonable match between the replicator and the Dow CSFB index of self-reporting hedge funds. The replication is not perfect which could be due to problems in the replicator or that the index rulebook does not fully capture the hedge fund market. Using the published index construction methodology and a C++ simulation of a hedge fund market, we find some evidence that the difference could be explained by the index rules failing to capture all of the aspects at play in the hedge fund sector. Continuing our analysis, we investigate whether true hedge fund returns are closer to the returns of the replicator than the index of reported returns. We seek to explain why a difference exists between the two by proposing a model in which hedge fund managers 'benchmark' themselves against some expected return and look at the probability of them not reporting returns if they fail to meet these expectations. We then try to discern possible candidates for this benchmark. Using this model, we find evidence consistent with the hypothesis that hedge fund managers are selectively reporting in order to present their returns in a more favourable light. The model suggests that if hedge fund managers are selectively reporting returns, one in 20 hedge fund managers would not report a monthly underperformance of 1 per cent.**

Journal of Derivatives & Hedge Funds (2011) **17,** 219–236. doi:10.1057/jdhf.2011.17;
published online 7 July 2011

Keywords: hedge funds; hedge fund replication; index construction

INTRODUCTION

Hedge funds are a class of investment vehicle aiming for *absolute return* – making a profit regardless of the market[1–3] conditions. They[4,5] often have a very loose mandate which means they are unconstrained in terms of the strategies they may employ and the markets they may enter. They may use over-the-counter derivatives tailored to a particular exposure profile, highly leveraged positions and short selling. This makes them a desirable investment for many high net worth individuals and institutions seeking to diversify their portfolio with an investment uncorrelated the major asset classes.

They are often highly protective of their exact methods and investments, claiming to protect their strategies and expertise from imitators. This secrecy and absolute return goal means that they can charge high fees to investors – '2 and 20' is typical – 2 per cent a year management and 20 per cent of any outperformance.[6] They are also usually highly illiquid with lock-in periods of several years meaning investors may be stuck in a failing hedge fund. For bodies with regulatory oversight (for example, pension funds) who invest in hedge funds, the extra work required in due diligence for an investment with little information in the public domain may be overly onerous. These issues have led many to question whether we can get the returns of hedge funds without these downsides – if we can reproduce these returns without the secrecy of hedge funds, then who needs hedge funds?

The attempt to replicate hedge funds has its roots in academia, however, many institutions are now offering hedge fund replication products which trade in such a way as to mimic the returns from hedge funds (see[7] for a list of some of the available products). In this article, we review some of the history of hedge fund replicators and construct our own replicators based on a 2006 paper by Hasanhodzic and Lo.[8] We continue the investigation by questioning whether what hedge funds actually report to the market is accurate or whether the replicators themselves are more indicative of true hedge fund returns than the indices compiled from data provided by self-reporting hedge funds.

HEDGE FUND REPLICATORS

Those attempting to construct hedge fund replicators have employed a number of different approaches; these can be loosely classified into three categories. The first is to try and find out the exact positions the hedge funds hold and buy these into replicating portfolios. This involves looking at information filed by companies the hedge funds may invest in and at any information the hedge fund itself provides.

The second replication method aims to produce returns, which are drawn from the same distribution as the target hedge fund or else have similar statistical properties. This area of work began with a 2005 paper by Kat and Palaro.[9] Kat and Palaro proposed considering the payoff distribution of an investor's portfolio containing hedge funds and other investments. In a complete market, this payoff distribution can be replicated by tradable assets, investing in these assets then provides a hedge fund replication strategy. Although the returns of the replicating portfolio will not match the returns of the hedge fund on a month by month basis, the returns will be drawn from the same distribution and as such will have all of the properties that investors like about hedge fund returns (in particular their low correlation with the capital markets).

The third approach to hedge fund replication is to create factor models. This involves expressing the hedge fund return for a particular period as a weighted sum of other time series over the period.

$$H_t = \sum_{i=1}^{k} \beta_{i,t} \Upsilon_{i,t} \qquad (1)$$

where H_t is the hedge fund return for period t, k is the number of factors, $\beta_{i,t}$ is the hedge fund sensitivity to factor i in period t and $r_{i,t}$ is the return of ith factor for period t.

The logic of this approach is that if hedge fund returns can be explained by a series of sensitivities to factors and these factors can be linked to liquid investments then a replicator can be constructed by investing in these market-traded investments in the proportions dictated by the sensitivities.

Factor models for hedge funds were inspired by Sharpe's 1992 paper[10] in which he showed that mutual fund return could be decomposed as a series of factors representing returns on bills, intermediate-term government bonds, long-term government bonds, corporate bonds, mortgage related securities, large capitalisation value stocks, large capitalisation growth stocks, medium capitalisation stocks, small capitalisation stocks, non-US bonds, European stocks and Japanese stocks. In 1997, Fung and Hsieh[11] applied Sharpe's model to hedge funds. They found that the analysis and fit of the regression varied greatly from fund to fund. This seemed to suggest that the set of factors chosen by Sharpe was not appropriate to hedge funds. In an example from their paper

> George Soros's Quantum fund was long US stocks and short Japanese stocks in the October 1987 stock market crash, short the British pound in September 1992, long precious metals in April 1993 … and long the US Dollar/short Japanese Yen in February 1994.[11]

However, even though Soros is investing in the same assets as Sharpe's mutual funds, Fung and Hsieh found that he has a low correlation to the returns of the asset classes. They went on to suggest this was because of Soros's '… *dynamic use of leverage and choice of asset exposure*'. That is, it is more important *how* the hedge fund trades than *where* (in which markets) they operate. Investigating this further, they found that hedge fund managers employing similar strategies have correlated returns. This seemed to indicate that, at least within a strategy,[12] there are one or more common return drivers which can be used as a basis for a replicator.

Examining trend following strategies in particular, Fung and Hsieh found that these returns were correlated with the return on a lookback option. Trend followers make a return through following the market trends – investing when the price is rising and selling when it is falling. A good trend follower will be able to identify an up- or down-trend early and invest appropriately – they will make a return from entering and exiting the market at the best times to capture the most of any price movements. As such, a lookback option, which pays the best return over an historical period, represents a perfectly executed market timing strategy. Any fund aiming for such a return will capture some percentage of the lookback option return. As such Fung and Hsieh postulated that trend following funds could be 'replicated' by investing in lookback options. A similar approach was taken for other strategies, for example, selling uncovered index put options for merger arbitrage[13] strategy funds,[14] long stock market and small caps, short large caps for equity long/short strategies,[15] yield spreads for fixed income arbitrage funds[16] and so on.

Hedge funds often self-classify themselves into one of these strategies but there is usually little to prevent a hedge fund manager diversifying by using other strategies or using hybrids or multi-strategies. There are also funds of hedge funds that may invest in more than one strategy and index providers who construct hedge fund indices from hedge funds following different strategies. If we have a series of instruments each replicating one strategy, the next logical step would be to create a weighted portfolio of these instruments to replicate multi-strategy hedge funds, hedge fund indices and funds of hedge funds.

Replicating the more aggregated information in hedge fund indices and funds of hedge funds is useful because the data hedge funds provide about themselves is seen as unreliable, suffering from self-selection and survivorship bias, as well as style drift – a hedge fund that reports that it is using a particular strategy may actually be using this along with other strategies and this may change with time. As such, it is useful to have a replicator that can accommodate more than one strategy.

Hedge funds are also under no obligation to report information about themselves and it could be suggested usually only do so to present themselves in a favourable light (for example,[17] found discontinuities in the distribution of reported returns and[18] found evidence that hedge funds inflate their December returns to boost annual return figures and ensure higher fees). In contrast, funds of hedge funds will probably still report returns even if a few of their hedge funds are underperforming. Likewise, hedge fund indices try to incorporate mechanisms to mitigate this bias (the Dow CSFB index will remove funds which stop reporting returns and, in an attempt to capture failing as well as successful funds, once a fund is included in the index, it will be tracked until liquidation).

Therefore, constructing a factor model that relates to these more aggregated instruments allows us to better calibrate the replicator to what we believe is truly happening in the market.

The aim of the hedge fund replicators is to capture as much of the hedge fund return as possible, using easily traded instruments. Sharpe suggested that any factor selection should be

> 1) Mutually exclusive, 2) exhaustive and 3) have returns that 'differ'. Pragmatically, each should represent market capitalization weighted portfolio of securities …'[10]

Approaches to constructing a factor model usually start with a shortlist of investments, which the replicator constructor believes capture the return of a particular hedge fund strategy or set of strategies. They then perform some kind of filtering to determine those factors with the highest explanatory power. For example, the following hedge fund replicators/factor models have been proposed

Amenc, El Bied, Martellini (2003)[19]
With factors:

- Change in yield on a 3M T-Bill (related to expectations of future stock returns),

- Volatility,
- Market volume (NYSE),
- Oil price,
- Moving average US stock market,
- Moving average world stock market.

Fung and Hsieh (2004)[20]

- Credit spread,
- Bond lookback options,
- Currency lookback options,
- Commodity lookback options,
- US stock market,
- Small caps – Large cap spread,
- Change in yield of 10Y T-Bond.

Hasanhodzic and Lo (2006)[8]

- US dollar,
- The bond market,
- Credit spreads,
- The US stock market,
- Commodities.

Goodworth and Jones (2007)[21]

- A hundred factors covering themes of equity, debt, credit, volatility, commodities, currencies and corporate events.
- They applied various screening techniques to get down to 5–8 factors.

Takahashi and Yamamoto (2008)[22]

- Small caps – Large cap spread,
- High book-to-market (value) – Low book-to-market (growth) spread,
- Momentum,
- At-the-money and out-of-the-money euro calls and puts on the S&P 500.

THE HASANHODZIC AND LO REPLICATOR

Of the models listed above, the Hasanhodzic and Lo replicator suggests itself as an obvious choice to turn into a commercial product such as the ones listed in.[7] Although some of the other models aim to generate static weights that can be held for many years, use information that would not have been available at the time the weights would be invested or require time-consuming data collection and filtering. The Hasanhodzic and Lo replicator uses factors from a range of markets that can be easily traded through futures or index trackers, allows weights to change every month to account for changes in the hedge fund market, does not require filtering and as such is easily adoptable as a valid trading strategy by market practitioners concerned with transparency and ease of use.

The Hasanhodzic and Lo replicator is a simple OLS regression performed every month of the factor returns on the hedge fund return using monthly data for the preceding 24 months. The βs of the regression provide the weights for the replicating portfolio.

Hasanhodzic and Lo's original analysis covered the period from 1986 to 2005. This means it preceded the volatile markets of the previous few years. To investigate how this replicator would perform in the markets of the recent past, we repeated Hasanhodzic and Lo's methodology to analyse the period from July 2006 to June 2010. We found that the replicator suggested weights as in Figure 1.

Therefore, we can see, for example, that at the end of 2008, we would reduce our exposure to the stock market in favour of bonds.

Investing in these weights should lead to returns, which mirror the returns of the hedge fund index. Performing this analysis, we find

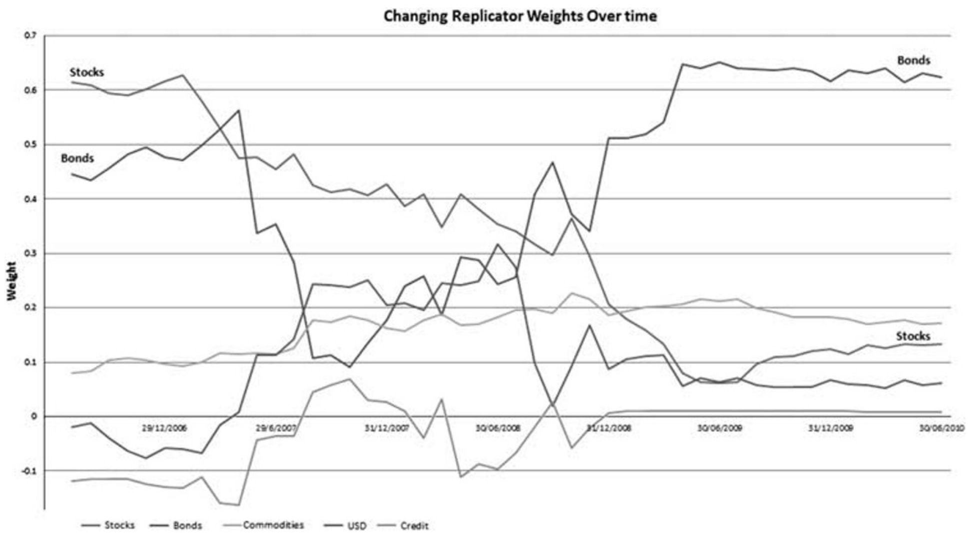

Figure 1: Weight invested in each asset by the replicator over time.

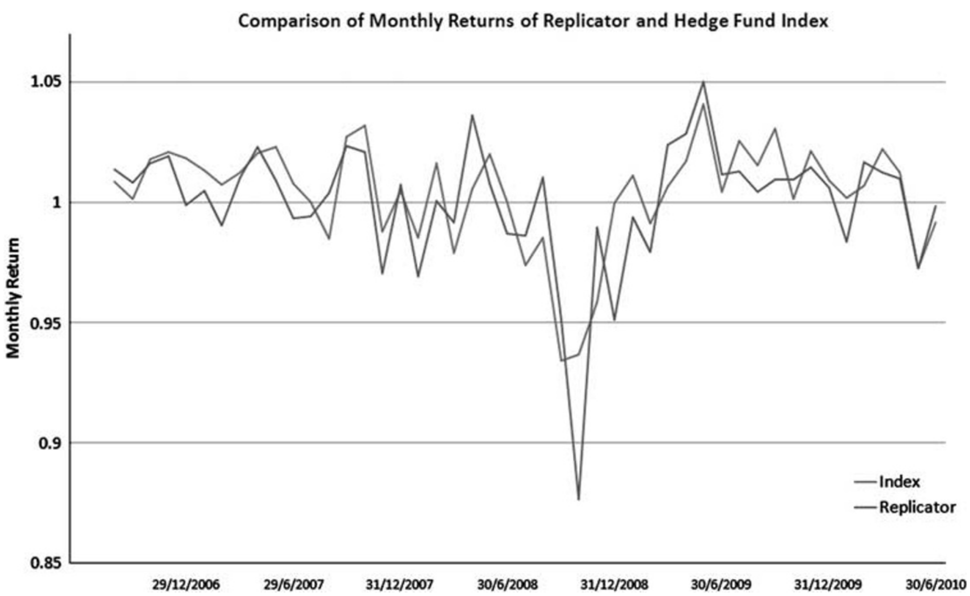

Figure 2: Monthly returns of the replicator and hedge fund index plotted over time.

that the time series for the replicator and index appear as in Figure 2.

Comparing the two, we see the drawdown of the replicator is considerably larger than the index. The maximum drawdown of the index over the period is 24 per cent but for the replicator it is 30 per cent.

This maximum occurs towards the end of 2008 during the Lehmans collapse. Plotting the cumulative return over the period in Figure 3,

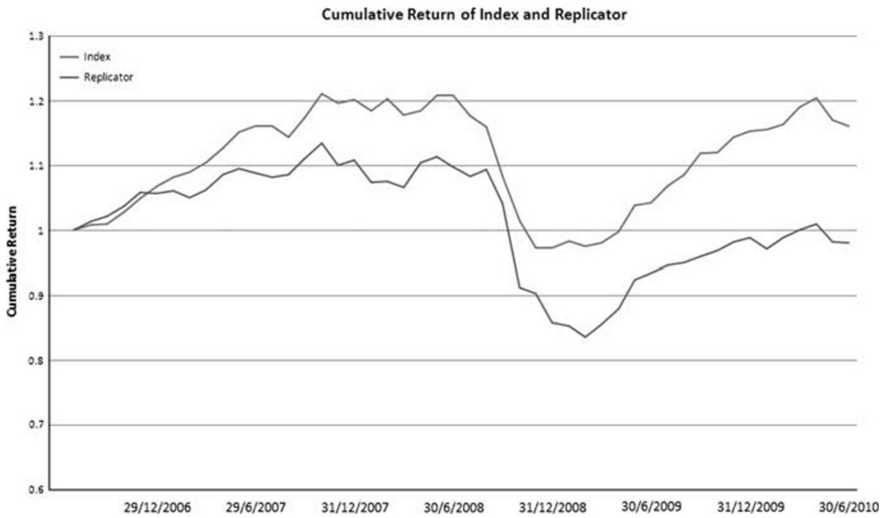

Figure 3: How the cumulative return of the index and replicator change over time.

we can see the difference between the two is even more pronounced.

Reading the final returns from the graph, it appears that investing in hedge funds over the period would have resulted in a 16 per cent profit whereas investing in the replicator would have led to a 2 per cent loss.

We can see from the graphs that the replicator is close (although with a lower return) to the index but how 'close' is this? We chose to quantify how well the replicator follows the index using the Kolmogorov–Smirnov statistic. This is a test of the null hypothesis that the returns are drawn from the same distribution against the alternative hypothesis that they are from different distributions. We estimate the cumulative distribution functions of the two time series using

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{Xi \leqslant x} \tag{2}$$

(as described in, for example, Bickel and Doksum.[23])

where $I_{Xi \leqslant x}$ is the indicator function, equal to 1 if $X_i \leqslant x$ and 0 otherwise. The distributions of the index and replicator using this estimator appear as in Figure 4.

Given, these distributions, it can be shown (for example, Pestman[24]) that under the null hypothesis, the Kolmogorov–Smirnov test statistic will tend to the same distribution – the Kolmogorov distribution – for all continuous $F(x)$. The Kolmogorov–Smirnov test statistic, using estimates of the two distributions, is defined as

$$D_n = sup|\widehat{F}_{1,n}(x) - \widehat{F}_{2,n}(x)| \tag{3}$$

The statistic for these two time series is 0.12766. Values closer to 0 indicate that
they are drawn from the same distribution – for example, using the tables in[24] (and that
we have 47 months in the sample) we have the critical value $P(D_{47} \leqslant 0.1946) = 0.95$. Therefore, we are quite a way from rejecting the null hypothesis (that they are drawn from the same distribution).

Another measure – the coefficient of determination, R^2 – is 0.379222.

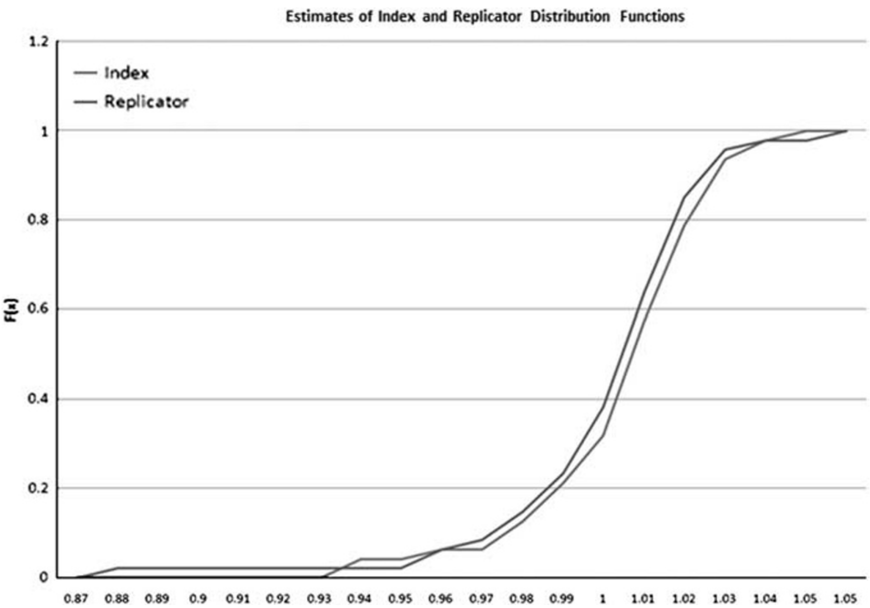

Figure 4: Index and replicator distribution functions estimated using Equation (2).

Hasanhodzic and Lo's next step was to see if they could get a closer match with the index through matching volatilities. They did this by investing partially in the replicating portfolio and partially in the risk free asset. We applied this to our study by estimating a monthly standard deviation of replicator and index using a rolling window sample of the preceding 12 months. We then invested a percentage of our wealth in the replicator with a weight given by

$$W_t = \frac{\sigma_{I,t}}{\sigma_{R,t}} \qquad (4)$$

where W_t is the proportion of portfolio invested in the replicator in period t, $\sigma_{I,t}$ is the standard deviation of monthly index returns using a sample of the previous 12 months, and $\sigma_{R,t}$ is the standard deviation of monthly replicator returns using a sample of the previous 12 months.

The remainder is assumed to be held as cash and earn no return. When we did this, we found replicator weights ranging from 57 per cent to 120 per cent with an average of 85 per cent. In 30 out of 36 months, we had to hold some cash to mitigate the higher volatility of the replicator.

This suggests that the hedge funds in the index as well as being better performing also have lower volatilities than the replicator. Plotting the returns of this volatility matched replicator in Figure 5, we can see, for example, that the hedge funds navigated the events from the end of 2008 to mid-2010 with much lower volatility than the replicator.

The Hasanhodzic and Lo replicator is designed as a realistic trading strategy, as such all weights are traded out of sample. To test how valid the Hasanhodzic and Lo replicators are, we can try to apply the weights to in sample returns. This should give us an idea of the extent to which hedge fund returns can be explained by the factors. We do this by investing in the factors

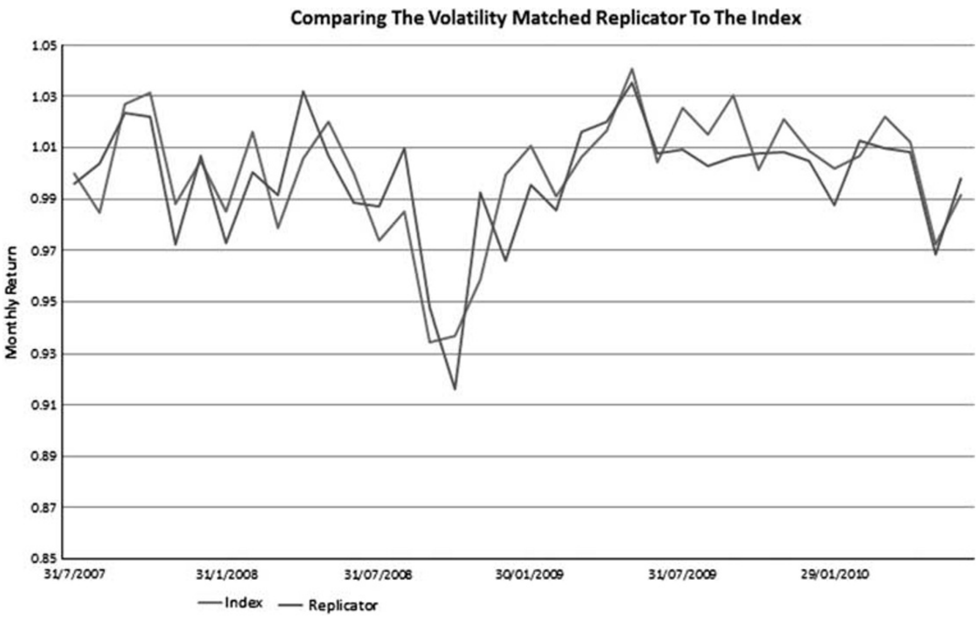

Figure 5: Comparison of the monthly returns of the volatility matched replicator and the index plotted over time.

1 year before the date suggested by the Hasanhodzic and Lo replicator. This leads to the results shown in Figure 6, plotted with the out of sample replicator from above and the index.

The Kolmogorov–Smirnov statistic for the in sample replicator is 0.170213, which is larger than the out of sample replicator and is only slightly less than the critical value for rejecting the null hypothesis at the 90 per cent confidence level – 0.1748. However, the R^2 is considerably higher at 0.798062.

Looking at the graph, it seems that although the in sample replicator provides a better match with the index than the out of sample one, the peaks are still not quite as high as the index of hedge funds and when there are losses, the replicators both fall further. This leads to a cumulative return for the period, which at 10 per cent still cannot get close to the index's return of 16 per cent.

Performing a similar volatility matching as for the out of sample regressions, we find that investing in the factors over the period would have given a better return for the same amount of risk. This can be seen in Figure 7.

INVESTIGATING THE HEDGE FUND INDEX

We can observe from our analysis that while the Hasanhodzic and Lo replicator provides a reasonable match with the index in normal times, it seems to diverge from the index at the peaks and troughs. This movement is always in the hedge funds' favour – they make more in good times and lose less in bad times. Bollen and Pool[17] have found some evidence that hedge funds may misreport returns to the market. More specifically, Agarwal, Daniel and Naik[18] have shown that hedge fund performance could be being smoothed around the year end with

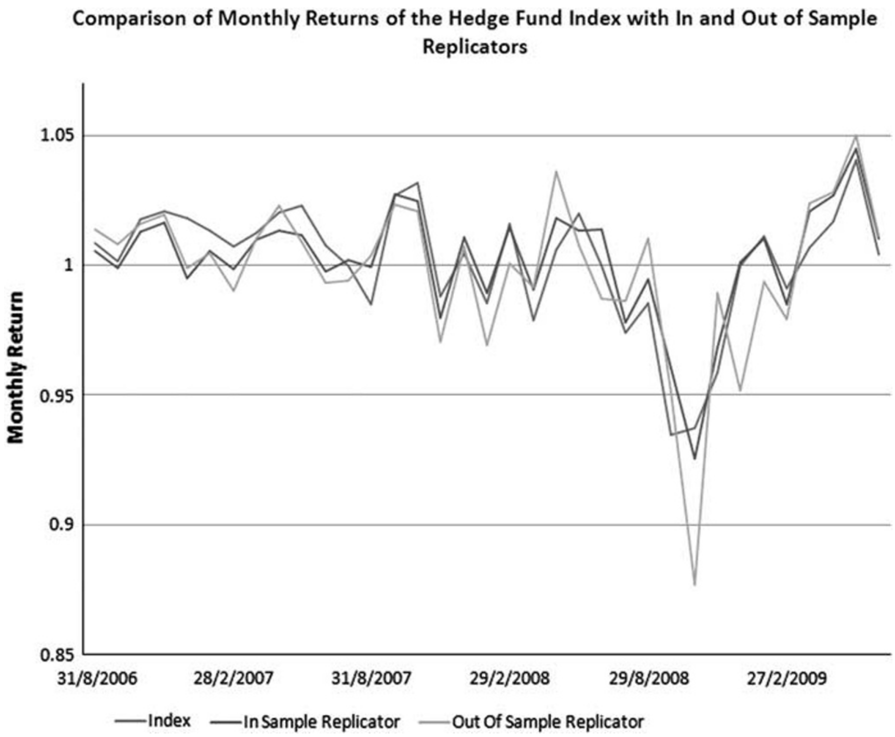

Figure 6: Comparison of the monthly returns of the two replicators with the index over time.

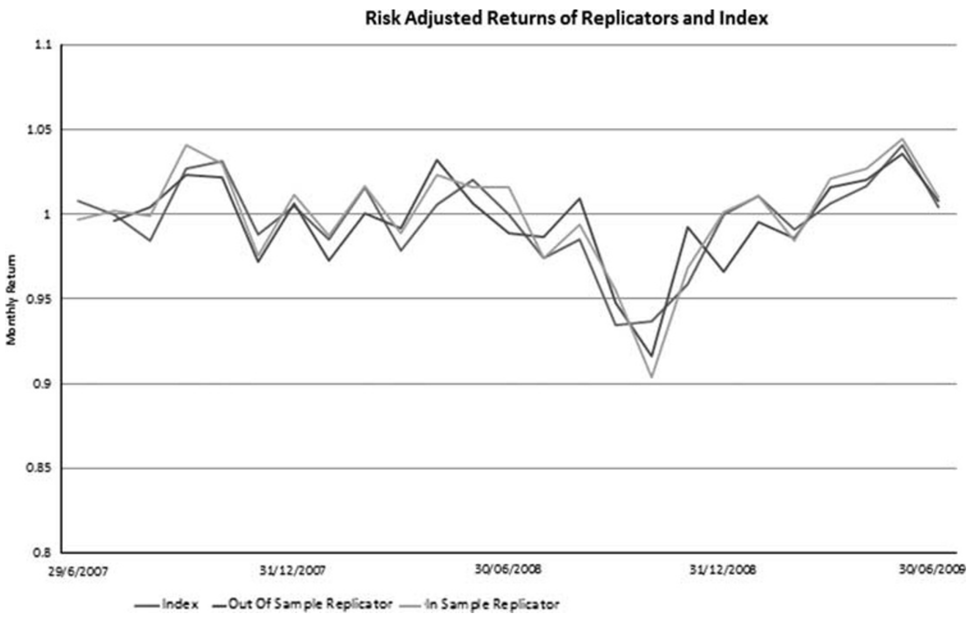

Figure 7: The changing risk adjusted returns of the two replicators and the hedge fund index.

some performance 'borrowed' from the following January. This is to ensure that the fund can report large annual returns and attract high fees.

We explore the possibility that differences between the index and replicator we have seen are because of survivorship and self-selection bias in the index. It is more likely that hedge funds which are doing well will want to advertise their returns and report than those doing less well who might choose not to report to the index. We expect this effect to be increased during difficult times in the markets and if the work in[18] is correct at the year ends. The prime example of this would be at the end of 2008 where we do indeed see a difference between the two time series.

The index provider incorporates rules to mitigate survivorship and self-selection bias. From the Dow CSFB methodology rule book:

- *Any hedge fund that fails to provide return data for 2 consecutive months, or fails to provide assets under management (AUM) data for 6 consecutive months, is deemed to have failed to meet the reporting requirements of the index and is dropped from the index.*
- *Member funds remain in the index unless they have liquidated or failed to meet the reporting requirements.*[2]

This, however, still allows funds to skip bad months as long as they do not do it twice in a row. For missing months, their return is '… estimated as the average reported returns of the member funds' [2]– effectively removing them from the index for a month. This could be when they have their largest negative effect on the index. This kind of smoothing could well explain the differences between replicator and index.

To test this idea, we built a simulation of the index. We believe that the hedge funds are exposed to the markets represented by the factors with a sensitivity proportional to the replicator weights. To simulate this, we created hedge funds which held the factors in the weights specified by the in sample replicator. To capture slight differences in market timing and mismatches between the return of the factor and the returns of the investments the funds actually hold, we added a Gaussian noise[25,26] to the return of each factor return in each hedge fund.

We created 465 such hedge funds (the number currently in the index) and started each hedge fund with a random amount of AUM.[27–30] This is required because the hedge fund index is AUM weighted – larger hedge funds have a greater impact on the index return.

After this initial distribution is generated, the hedge funds' AUMs would grow or shrink based on the returns of the factors they held.[31,32]

For a typical hedge fund (#233) in the simulation, we have plotted the effect of altering the Gaussian noise parameter on the subsequent hedge fund returns (Figure 8).

Finally, an index was created using the construction methodology provided by the index provider.[2]

We were then in a position to include survivorship and self-selection bias. Whether a hedge fund reports or not could depend on a variety of factors. We approached it by introducing a benchmark, which represents what the hedge fund manager believes the fund's returns should be if he were to report, we then compare the simulated returns to this benchmark.

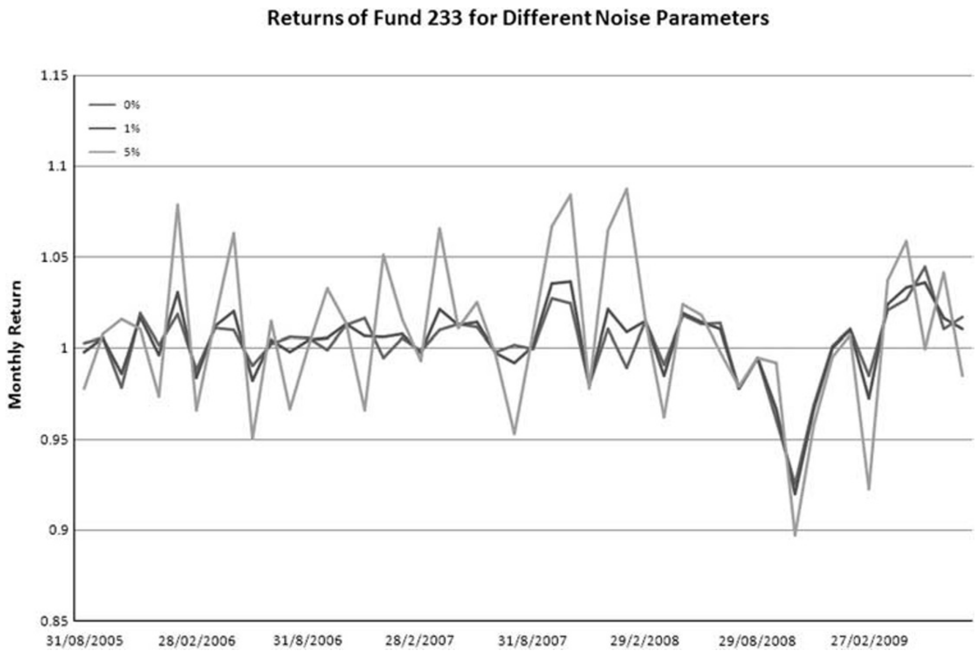

Figure 8: The monthly returns of a representative fund in the simulated index for different values of the noise parameter.

The benchmark used represents what the hedge funds would realistically benchmark themselves against. At what point would they feel that performance was not good enough to report? In our initial investigations, we take as our benchmark the hedge fund index itself. Although it might be argued that this makes things somewhat circular, it is reasonable to suppose that this is what is actually happening in the market. After all, hedge fund managers are competing with other hedge funds for investors, therefore it is sensible to assume that they would want to report higher returns than their peers.

Another factor we considered was that the funds may benchmark themselves by past performance – if they suffer their worst ever loss, they may not report even if the loss is not that great or other hedge funds in the market are reporting larger losses. There is also the question of what should happen if the hedge fund underperforms the benchmark. Should they leave the index immediately or wait a few months to see if things improve or does this vary from hedge fund to hedge fund?

To decide how to model hedge funds leaving, we took as our benchmark the average monthly return of the hedge fund index for the preceding 12 months. We used the following non-reporting rules:

- *Simulation A*: Funds leave the index (and do not return) if they underperform the benchmark for 3 consecutive months.
- *Simulation B*: Funds do not report a return if it is their worst loss in the period. If they do not report for 2 consecutive months, they are

removed according to the rules of the index provider.

- *Simulation C*: Funds will leave the index if they underperform the benchmark for 3 consecutive months and will not report if it is their worst loss up to that point.
- *Simulation D*: If a fund underperforms the benchmark, it may or may not report with a probability of it not reporting increasing with the size of the underperformance. If they do not report for 2 consecutive months, they are removed according to the rules of the index provider.

Running these simulations, we found the results in Table 1.

Therefore, starting with the in sample replicator which had a Kolmogorov–Smirnov statistic of 0.170213, we built simulations involving hedge funds selectively reporting returns. This led to a better fit in terms of Kolmogorov–Smirnov in all simulations. In the case of simulation D, the Kolmogorov–Smirnov came down to 0.104167 and the R^2 increased from 0.798062 to 0.81095. This increase in goodness of fit when we incorporate non-reporting is consistent with the hypothesis that hedge funds may be selectively reporting returns.

Table 1: Comparison of simulations and replicator returns against the reported returns of the hedge fund index

	Kolmogorov–Smirnov	*Coefficient of determination*
In sample replicator	0.170213	0.798062
Simulation A	0.125000	0.792458
Simulation B	0.125000	0.761076
Simulation C	0.166667	0.696809
Simulation D	0.104167	0.810950

If hedge funds are misreporting in this way it may be possible to develop a sophisticated model of manager behaviour perhaps aided by the large analogous literature of tax non-compliance and misreporting (for example, see[33,34] and[35]). As an attempt to further investigate the decision to not report, we can look at the probability of a hedge fund not reporting as used in simulation D. We believe this probability distribution should be some function of the size of the underperformance with hedge fund managers more likely to not report if the loss is larger. For the purposes of the simulation, we assume a simple relation where

$$P(\text{Hedge fund does not report a monthly loss of size } L) = kL \quad (5)$$

where k is some constant multiplier related to how important the size of the loss is to the decision to not report. We can see the effect of changing this k on the simulation in Figure 9.

As we would expect, if the parameter is high (for example, $k = 50$ in the above), hedge fund managers are unlikely to report even small losses and therefore the simulated index return is higher in bad months than if k were smaller.

We can investigate from this, if a hedge fund has a bad month, what the probability of them not reporting that return will be. Repeating simulation D with different values for this input parameter, we find the results shown in Table 2.

We can see a peak in the Kolmogorov–Smirnov statistic, which seems to indicate that the correct value for k is around 5. Substituting a typical loss into Equation (5), and setting k equal to 5, this means that we expect one in 20 hedge fund managers who have made 1 per cent less

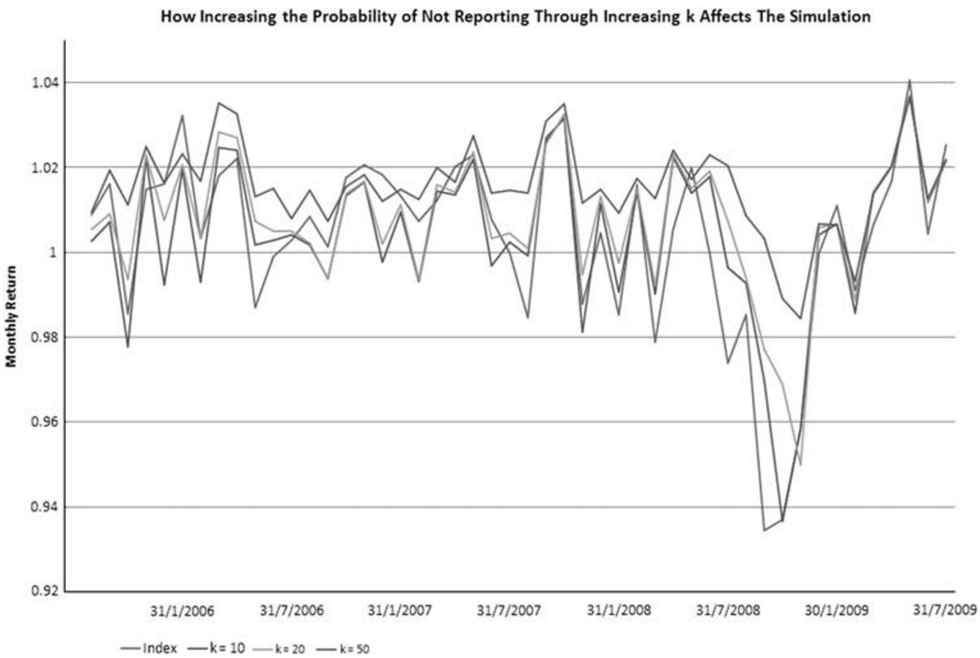

Figure 9: Monthly returns of the simulated index for different values of the not reporting factor, *k*.

Table 2: Comparison of simulated returns with the reported returns of the hedge fund index for different parameter values *k*

Probability of not reporting factor (k)	*Kolmogorov–Smirnov*	*Coefficient of determination*
2	0.125000	0.805287
4	0.125000	0.806687
5	0.104167	0.810950
6	0.125000	0.811151
7	0.125000	0.809903
10	0.166667	0.786610
20	0.333333	0.635488

than they thought they should in a month to not report.

So far we have assumed that hedge fund managers would benchmark themselves against other hedge fund managers through the 12 month rolling average of the hedge fund index. The previous 12 months performance may be too long – financial consultancy firms such as Allenbridge adjust their rankings tables[36] every month and where information is released by the hedge fund itself, it will usually be on a monthly basis. It is therefore advantageous for the fund if its reported monthly returns outperform other funds' monthly returns. We may also want to consider a benchmark like the index value minus 0.2 per cent, meaning the hedge fund manager would not consider leaving the index if he only underperforms by a small amount.

Alternatively, we may argue that hedge fund managers would see their competition as wider than just other hedge fund managers and may be concerned that reporting a poor performance will lead to investors removing their investment from the hedge fund sector completely. As such,

Table 3: Comparison of simulated returns with the reported returns of the hedge fund index for simulations with different benchmarks

Benchmark	*Kolmogorov–Smirnov*	*Coefficient of determination*
Rolling average of last 12 months of index returns	0.104167	0.810950
Rolling average of last 3 months of index returns	0.125000	0.776168
Hedge fund index	0.145833	0.778550
Hedge fund index − 0.2%	0.125000	0.821924
Absolute Return	0.125000	0.821924
S&P 500	0.083333	0.700151

an alternative hedge fund benchmark may be something like the S&P500. Or the benchmark could be even simpler, as hedge funds emphasise absolute return regardless of the market conditions perhaps their benchmark is just to get some positive return, no matter the size. In the above (Table 3), we start with our original simulation D and change the benchmark to see what effect it will have on our simulation and therefore see if we can say anything more about the fund managers' choice to not report a return.

Therefore, it looks like the original choice of the previous 12 months of hedge fund index performance gives quite a good fit in terms of both measures; however, in terms of the Kolmogorov–Smirnov statistic, it seems that using a benchmark of the S&P500 gives a much better fit with the index. This may suggest that hedge fund managers are motivated to not report returns because of performing less well than the stock market.

CONCLUSIONS

In this article, we looked at hedge fund replicators with a special focus on Hasanhodzic and Lo replicators of the Dow CSFB hedge fund index. We found that to a certain extent the returns of the index could be replicated by a portfolio of liquid, market-traded instruments. Examining the differences in drawdown structure and cumulative returns, we used a simulation to test if there was evidence of hedge funds misreporting returns to the hedge fund index provider. We found that if the replicator returns represent the true hedge fund returns, we can construct a hedge fund index, which matches the Dow CSFB index by having these funds selectively report returns. This is consistent with the hypothesis that there is some misreporting of hedge fund returns.

We then went on to examine the difference between the returns the model suggests hedge funds achieve and those they report. We found that this difference could be explained to a degree by them not reporting if they underperformed some benchmark. We assumed a simple probability distribution, used the simulation to calibrate this and found that hedge fund managers underperforming the benchmark by 1 per cent would not report with probability of approximately 5 per cent. Continuing with this analysis, we tried to discover which benchmark hedge fund managers would consider important when deciding whether to report losses and found that this decision may be effected by the performance of the stock market.

From this limited study, many avenues for further investigation are suggested. It would be interesting to see if any more information could be gained about why hedge fund managers do not report returns from the difference between the reported returns and the model. We assumed a simple distribution for the probability of a hedge fund manager not reporting returns – can we say anything further about this distribution?

We have looked at the Hasanhodzic and Lo replicator in particular, however, there are quite a few other replicators and factor models to which we could apply this analysis. In Goodworth and Jones' 2007 paper,[21] they perform detailed analysis of hedge fund returns starting with a broad range of 100 factors and editing this down with a series of quantitative and qualitative filters. Their aim was to create a factor model for calculating portfolio value at risk but there is no reason not to use their factors and factor selection methodology as part of a replicator. We might also use other replication targets, other than the Dow CSFB index. Most of the hedge fund index providers produce hedge fund indices broken down by strategy, if the work of Fung and Hsieh is correct that replication strategies should be based around replicating funds by strategy then this may provide further insight. It is worth mentioning that this is the approach taken by Hasanhodzic and Lo in their 2006 paper applying their replicators to individual strategies.

Géhin and Vaissié in 2006[37] proposed improvements to the factor models based on simple linear regression – primarily using Kalman filters to get a more accurate picture of hedge fund factor exposure. This could also be incorporated into our analysis.

One final aspect could be to look at how hedge funds are affected by volatility in the market and how they then report this volatility. There is an observed positive autocorrelation in hedge fund returns,[3] and it has been suggested that in illiquid markets, hedge fund administrators may simply use the last available price which leads to a lower volatility and therefore better performance in terms of things like Sharpe ratios than there would be otherwise. It may be a useful exercise to see what effect this has on reported returns and on our model. It would also be interesting to see how hedge funds respond to changes in volatility in the markets. Our simulations could incorporate this information through making the standard deviation of the Gaussian distribution of factor returns a variable noise factor; perhaps this factor would be shocked in months where the change in a volatility index such as VIX was large.[38]

REFERENCES AND NOTES

1 The Dow CSFB index is one of many hedge fund indices available. Each will have different calculation methodologies, for example.[2] See[3] for a list and description of the most commonly used indices.

2 Dow Jones Credit Suisse Hedge Fund Indices as well as information on methodology of the index construction and characterisation of hedge funds, http://www.hedgeindex.com.

3 Brooks, C. and Kat, H.M. (2002) The statistical properties of hedge fund index returns and their implications for investors. *Journal of Alternative Investments* 5(2): 26–44.

4 Roncalli and Weisang's 2008 paper[5] gives a review of some problems found in these replicators; in particular slow reactivity, problems modelling non-linear returns, capturing tactical allocations, capturing alpha and matching higher moments of returns.

5 Roncalli, T. and Weisang, G. (2008) Tracking problems, hedge fund replication and alternative beta, http://ssrn.com/abstract=1325190.
6 Jaeger, R.A. (2003) *All About Hedge Funds*. USA: McGraw Hill, ISBN 0-07-139393-5.
7 Amenc, N., Géhin, W., Martellini, L. and Meyfredi, J. (2008) Passive hedge fund replication: A critical assessment of existing techniques. *The Journal of Alternative Investments*: 69–83, (Fall).
8 Hasanhodzic, J. and Lo, A.W. (2006) Can hedge-fund returns be replicated? The linear case, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=924565.
9 Kat, H.M. and Palaro, H.P. (2005) Who needs hedge funds? A copula-based approach to hedge fund return replication. Alternative Investment Research Centre Working Paper no. 27, Cass Business School Research Paper, http://ssrn.com/abstract=855424.
10 Sharpe, W. (1992) Asset allocation: Management style and performance measurement. *The Journal of Portfolio Management*, 18(2): 7–19.
11 Fung, W. and Hsieh, D.A. (1997) Empirical characteristics of dynamic trading strategies: The case of hedge funds. *The Review of Financial Studies* 10(2): 275–302.
12 See[2] for a full description of the various hedge fund strategies.
13 Merger arbitrage funds concentrate on securing absolute return through betting on whether an announced merger will take place.
14 Mitchell, M. and Pulvino, T. (2001) Characteristics of risk and return in risk arbitrage. *Journal of Finance* 56(6): 2135–2175.
15 Fung, W. and Hsieh, D.A. (2004) Extracting portable alphas from equity long-short hedge funds. *Journal of Investment Management* 2: 57–75.
16 Fung, W. and Hsieh, D.A. (2002) The risk in fixed-income hedge fund styles. *Journal of Fixed Income* 12(2): 16–27.
17 Bollen, N.P.B. and Pool, V.K. (2009) Do hedge fund managers misreport returns? Evidence from the pooled distribution. *The Journal of Finance* 64(5): 2257–2288.
18 Agarwal, V., Daniel, N.D. and Naik, N.Y. (2009) Do hedge funds manage their reported returns? http://ssrn.com/abstract=891169.
19 Amenc, N., El Bied, S. and Martellini, L. (2003) Predictability in hedge fund returns. *Financial Analysts Journal* 59(5): 32–46.
20 Fung, W. and Hsieh, D.A. (2004) Hedge fund benchmarks: A risk based approach. *Financial Analysts Journal* 60(5): 65–80.
21 Goodworth, T.R.J. and Jones, C.M. (2007) Factor-based, non-parametric risk measurement framework for hedge funds and fund-of funds. *The European Journal of Finance* 13(7): 645–655.
22 Takahashi, A. and Yamamoto, K. (2008) Hedge fund replication, http://ssrn.com/abstract=1420158.
23 Bickel, P.J. and Doksum, K.A. (2001) Mathematical statistics, basic ideas and selected topics. Vol.1, 8–9, Prentice Hall, ISBN 0-13-850363-X.
24 Pestman, W.R. (1998) Mathematical statistics: An introduction. 284-6, Walter de Gruyter, ISBN 3-11-015356-4.
25 Gaussian noise constructed using a Box Muller algorithm as described in[26].
26 Miller S. Introduction to C++, Course Notes, Academic Year 2008–2009, Economics, Mathematics and Statistics, Birkbeck College.
27 The initial AUMs were generated using the Pareto algorithm from the Boost library,[28] with parameters set using input from.[29] The Pareto distribution was proposed for wealth distributions[30] and therefore is appropriate to model AUM. See[31] for a discussion of this distribution.
28 Boost C++ Library, http://www.boost.org/, with a description of the implementation of the Pareto distribution at, http://live.boost.org/doc/libs/1_39_0/libs/math/doc/sf_and_dist/html/math_toolkit/dist/dist_ref/dists/pareto.html, and source code located within the Boost library at boost/math/distributions/pareto.hpp.
29 *The Hedge Fund Elite*, New York Magazine, 9 April 2007, http://nymag.com/news/features/2007/hedgefunds/30342/.
30 Wold, H.O.A. and Whittle, P. (1957) A model explaining the Pareto distribution of wealth. *Econometrica* 25(4): 591–595.
31 AUM can also vary because of subscriptions and redemptions; however, it is assumed that investors will invest in proportion to the hedge fund's size. As the index is AUM weighted, this means that there will be no net effect on the index return. This is not altogether realistic – it is likely that subscribers will concentrate in the historically more successful hedge funds and depart from those failing hedge funds. See Satchell[33] for further discussion of the links between investor preference, subscriptions, returns and AUM.
32 Satchell, S. (2009) IIR Fund Manager Selection Conference May 2009, Zurich, Switzerland.
33 Chander, P. (2007) Income tax evasion and the fear of ruin. *Economica* 74(294): 315–328.
34 Cai, H. and Liu, Q. (2009) Competition and corporate tax avoidance: Evidence from Chinese industrial firms. *The Economic Journal* 119(537): 764–795.
35 Cruz, C.A., Shafer, W.E. and Strawser, J.R. (2000) A multidimensional analysis of tax practitioners' ethical judgements. *Journal of Business Ethics* 24(3): 223–244.
36 Allenbridge Investment Consultants, http://www.allenbridge.co.uk/downloads_ranking_tables.php.

37 Géhin, W. and Vaissié, M. (2006) The right place for alternative betas in hedge fund performance: An answer to the capacity effect fantasy. *Journal of Alternative Investments* 9(1): 9–18.

38 We assume it is the change in VIX that would affect the hedge fund returns as any constant level of volatility could be hedged away by the hedge funds. However, when an unexpected change in volatility occurs they may not be able to change their exposures immediately leading to higher volatility in the hedge fund returns.