Cass Business School
City of London

WHO NEEDS HEDGE FUNDS?

A Copula-Based Approach to Hedge Fund Return Replication

Harry M. Kat*

Helder P. Palaro**

This version: November 28, 2005

Please address all correspondence to:

Harry M. Kat
Professor of Risk Management and
Director Alternative Investment Research Centre
Cass Business School, City University
106 Bunhill Row, London, EC2Y 8TZ
United Kingdom
Tel. +44.(0)20.70408677
E-mail: harry@harrykat.com

---

*Professor of Risk Management and Director Alternative Investment Research Centre, Cass Business School, City University, London. **PhD Student, Cass Business School, City University, London.

The authors like to thank CSI and Rudi Cabral for allowing us to access the CSI futures database. Comments from Andrew Clare, John Hatgioannides, Roel Oomen and participants of the 2005 FMA Annual Meeting in Chicago are gratefully acknowledged.

WHO NEEDS HEDGE FUNDS?

A Copula-Based Approach to Hedge Fund Return Replication

Abstract

In this paper we develop and demonstrate the workings of a copula-based technique that allows the derivation of dynamic trading strategies, which generate returns with statistical properties similar to hedge funds. We show that this technique is not only capable of replicating fund of funds returns, but is equally well suited for the replication of individual hedge fund returns. Since replication is accomplished by trading futures on traditional assets only, it avoids the usual drawbacks surrounding hedge fund investments, including the need for extensive due diligence, liquidity, capacity, transparency and style drift problems, as well as excessive management fees. As such, our synthetic hedge fund returns are clearly to be preferred over real hedge fund returns.

1. Introduction

Rising from relative obscurity, over the last 15 years hedge funds have become increasingly popular with high net worth individual as well as institutional investors. As a result, the number of hedge funds has risen from around 500 in 1990 to an estimated 8000 in 2005. Over the same period, assets under management are estimated to have increased from $50 billion to $1 trillion. Apart from the success of the prime brokerage concept, one of the crucial factors behind the spectacular growth of the hedge fund industry has been the rise of the fund of funds structure as the preferred way of investing in hedge funds. Currently, most money invested in hedge funds flows through funds of funds, with the total number of such funds being estimated at around 4000.

Initially, hedge funds were sold on the promise of superior performance, the story being that hedge fund managers' long experience and proven investment skills were a virtual guarantee for superior returns. Especially high net worth investors proved sensitive to these arguments and fuelled much of the early growth of the industry. Towards the end of the 1990s the story began to change, however. No longer were hedge funds sold on the promise of superior performance, but more and more on the basis of a diversification argument, pointing at hedge fund's relatively low correlation with stocks and bonds and the beneficial effects on risk and return from including hedge funds in the traditional investment portfolio. The reason for this rather remarkable change in sales tactics was twofold. Firstly, starting in the late 1990s, hedge fund performance took a turn for the worst, with every next year being worse than the year before. According to the HFRI Fund of Funds Composite Index, the average fund of funds only returned a meagre 3.85% over the first ten months of 2005. Secondly, driven by historically low interest rates, substantial losses in the equity markets, and keen to be seen taking action, institutional investors started to look more seriously at hedge funds as well. Given institutions' emphasis on risk management, the hedge fund story changed to accommodate this new clientele[1].

[1] The current trend is to introduce retail investors to hedge funds as well. Since the typical retail investor is unlikely to appreciate the special nature of hedge fund investment, this will intensify the call for more profound regulation, which in turn will force the industry to reshape itself once again.

Hedge fund managers typically put a lot of effort into generating their returns. However, with the industry implicitly admitting, and more sophisticated performance studies confirming[2], that hedge fund performance is not truly superior (anymore), the question arises whether it is possible to generate similar returns in a much more mechanical way and with less effort. More precisely, is it possible to design dynamic trading strategies, mechanically trading cash, stocks, bonds, etc., that generate hedge fund-like returns? If indeed we could design such strategies, this would solve a respectable number of problems typically surrounding hedge funds (as well as many other "alternative" investments), including:

The need for extensive due diligence

Without any publicly available information and research, investors are forced to invest substantial amounts of time and energy in visiting hedge fund managers, asking questions, interpreting the answers and doing background checks. A number of third parties do provide these services on a stand-alone basis, but at significant costs.

Lack of liquidity

Most hedge funds use lock-up structures to tie in new investors for periods ranging from 6 months up to 5 years[3]. After the lock-up has expired, investors typically need to give one or three months notice if they want to disinvest. In addition, some funds charge departing investors an additional fee of up to 5% "to compensate remaining investors for the costs of having to rebalance the fund portfolio". It is hard to see why a fund would require an exit fee if there is already a notice period in place though. Given proper notice, freeing up money should not cost an arm and a leg. Imposing an exit fee therefore seems nothing more than a subtle way of extending the lock-up period.

Lack of transparency

All hedge funds claim to do something highly exclusive and proprietary and anxiously guard their trading secrets. Although transparency has improved with the arrival of institutional investors, hedge fund investors are seldom told what exactly goes on

[2] See for example Amin and Kat (2003), Bailey et al. (2004) or Fung et al. (2005).

[3] Investors are becoming increasingly resistant to lock-up periods. According to Dyment et al. (2005), in 2004 68% of investors would only invest with managers with lock-ups of one year of less. In 2005, this rose to 77%.

inside the black box. As a result, it can sometimes be very hard to properly assess the risk-return characteristics of a fund[4].

Lack of capacity

In an attempt to preserve the level of returns, successful hedge funds may close for new investors or close for new money altogether[5]. This, however, does not prevent money from flowing to other managers in the same category. As a result, when opportunities are in limited supply, performance may come under pressure. This is especially true for arbitrage-type strategies, where the arrival of more money and/or managers will significantly increase market efficiency. Recently, convertible bond arbitrage has suffered quite badly from this form of over-investment, reporting an average return of –2.65% for the first ten months of 2005 (HFRI Convertible Arbitrage Index).

Excessive management fees

The average hedge fund charges its investors "2 plus 20", i.e. a flat management fee of 2% plus an additional incentive fee equal to 20% of any profits over a hurdle rate. Funds of funds tend to charge an additional "1 plus 10" on top of this. With interest rates and hedge fund performance at historically low levels, this means that nowadays pre-fee hedge fund returns are split more or less equally between investors and fund (of funds) managers.

Style drift

Hedge fund managers may sometimes change their style or strategy, which in turn may cause a significant change in a fund's risk-return profile. When not explicitly notified of this change and without sufficient transparency, investors can only find out about this from the returns that the fund generates. With returns reported on a monthly basis, however, it could take a long time before it becomes clear that something has changed.

---

[4] Dyment at al. (2005) report that only 14% of investors requires full transparency. Even more surprisingly, 19% of investors do not require any transparency at all.

[5] Anticipating closure, according to Dyment et al. (2005), in 2005 45% of investors required future capacity rights when investing in hedge funds.

Of course, we are not the first to attempt to replicate hedge fund returns. Following the work of Sharpe (1992) on equity mutual funds, previous authors have primarily relied on the use of factor models to replicate month-to-month returns[6]. In theory, the factor model approach should work well. Once the relevant risk factors have been identified and the fund's sensitivity to these factors has been determined, one can construct a portfolio of stocks, bonds, and other securities with the same factor sensitivities as the fund in question. Since it has the same factor sensitivities, the resulting portfolio will generate returns that are similar to those of the fund.

The problem when applying the above approach in a hedge fund context is that in practice we often have little idea how hedge fund returns are actually generated, i.e. which risk factors to use. As a result, factor models typically explain only 25-30% of the variation in hedge fund returns, which compares quite unfavourably with the 90-95% that is typical for mutual funds. Although the procedure works better for portfolios of hedge funds, funds of funds and hedge fund indices, where much of the idiosyncratic risk is diversified away, factor models do not appear to offer a particularly fruitful alternative when looking to replicate hedge fund returns accurately[7].

Given the failure of the factor model approach, we took a step back and reconsidered the problem at hand. When an investor likes a hedge fund, it is (or should be at least) because of the statistical properties of the fund's returns, i.e. their mean, standard deviation, etc. and their relationship with the returns his existing portfolio. This implies that we do not necessarily have to replicate a fund's month-to-month returns. For most applications it will be enough if we can *generate returns with the same statistical properties as the returns generated by the fund.*

So far, there has only been one study, which followed the above route. Based on the early theoretical work of Glosten and Jagannathan (1994) and Dybvig (1988a, 1988b), and primarily aimed at evaluating hedge fund performance, Amin and Kat (2003) developed mechanical trading strategies, trading the S&P 500 and cash, which aim to

[6] See for example Schneeweis et al. (2003) or Agarwal and Naik (2004).

[7] Despite the lack of explanatory value, several parties have recently announced the intention to launch factor model based products that aim to provide investors with hedge fund like returns at lower costs and in a more convenient format.

generate returns with the same marginal distribution as the returns of a given hedge fund. Although interesting from a theoretical perspective, from a practical perspective only replicating the marginal distribution is not enough, though. Most of today's investors are attracted to hedge funds because of their relatively weak relationship with traditional asset classes and their own portfolio in particular. To properly replicate hedge fund returns we therefore not only have to replicate the marginal distribution, but also the relationship between a fund and the investor's existing portfolio. In this paper we develop a procedure, which does exactly that.

The basic idea behind the proposed procedure is straightforward. From the theory of dynamic trading it is well known that in the standard theoretical model with complete markets one can perfectly hedge any payoff function. Therefore, if we can find a payoff function which, given the probability distribution of the underlying index or indices, implies the desired distribution, we will also have found the dynamic trading strategy which generates (returns that are drawings from) that distribution.

Of course, there are a number of serious hurdles to take. First, we are not interested in just any strategy. To maximize expected return, we want the cheapest strategy possible. Second, since we are aiming to replicate not only a fund's marginal return distribution but also its relationship with the investor's existing portfolio, we are confronted with bivariate distributions, which can take on a large variety of shapes and forms. Third, real markets are a lot less well behaved than assumed in the standard theoretical model. As a result, an inconsistency may arise between the determination of the desired payoff function, which is a purely empirical matter, and the subsequent derivation of the dynamic trading strategy generating that payoff. A second consequence of relying on an abstract model is that in practice our dynamic trading strategies may not be able to exactly generate the desired payoff. We therefore perform extensive out-of-sample tests of our strategies, using daily data over the period 1985-2005.

The replication procedure concentrates on replicating a fund's risk profile without explicitly considering the fund's expected return. The underlying assumption is that, in an efficient market, in the longer run investors will receive a return in line with the risk that they have taken. This is why the empirical finding that hedge fund returns are not truly superior is fairly crucial. If they were superior, we would still be able to replicate their risk profile, but we could not expect to replicate their average as well. If it is superior, it can't be replicated and vice versa. The latter observation points at another application of the replication technique developed in this paper: the evaluation of hedge fund returns. Explicitly constructed to offer the same risk profile, when the average replicated return is significantly higher than the average fund return, the fund is the inefficient alternative. We will investigate this line of thought further in a forthcoming companion paper[8].

The present paper proceeds as follows. In the next section we briefly discuss the theoretical setting in the form of Dybvig's (1988a) Payoff Distribution Pricing Model (PDPM), which we extend to a bivariate setting. In section 3, we discuss the determination of the desired payoff function, i.e. the payoff function, which, given the distribution of the assets to be traded, implies the desired return distribution. In section 4 we carry out a number of simulation-based analyses, investigating how the size of the available data sample influences the accuracy of the procedure. In section 5 we discuss the practical implementation of the procedure and the results of some out-of-sample tests, replicating the returns on three well-known hedge funds (of funds). Section 6 concludes. Proofs relating to the univariate and bivariate PDPM can be found in Appendix I.

2. Theoretical Setting

In principle, a given payoff distribution can be generated by many different payoff functions. Different payoff functions come with different price tags, however. We therefore need to know more about the general characteristics of the cheapest alternative. This is where Dybvig's (1988a) Payoff Distribution Pricing Model (PDPM) comes in. The PDPM can be derived from a simple set of primitive assumptions:

1) Investors' preferences depend only on the probability distribution of terminal, i.e. end-of-horizon, wealth.

---

[8] See Kat and Palaro (2006a).

2) Perfect capital markets.
3) Investors prefer more to less.

This set of assumptions allows investors' preferences to depend on all moments of the distribution of terminal wealth.

Suppose there are *n* possible states of the world. The *state price* of state *i* is the price of an elementary security which pays \$1 if state *i* occurs and 0 otherwise. The *state-price density* is defined as the price per unit of probability of terminal wealth in a particular state, and is given by the ratio of the state price and the probability of occurrence of that state. In Dybvig (1988a), the author shows that the cheapest way to obtain a given payoff distribution is to allocate terminal wealth as a decreasing function of the state-price density. In Dybvig (1988b), the author applies this result, assuming a binomial tree model for the underlying index, and shows that for a payoff function to be efficient it should *allocate terminal wealth as a non-decreasing function of the final value of the underlying index*.[9] Intuitively, this is a plausible result as it implies that payoff and index will be positively correlated, which, when it comes to actually generating the payoff, will serve to keep the required rebalancing trades down.

We now propose a more general set of assumptions. Suppose that apart from being concerned about the terminal wealth obtained from some new investment opportunity, investors are also concerned about the dependence between this investment and their existing portfolio. This means replacing assumption 1 by:

1) Investors' preferences depend only on the joint probability distribution of terminal wealth derived from the investment and their existing portfolio.

Equivalently, since the distribution of the investor's existing portfolio will be given, we can say that:

---

[9] Proof can also be found in Appendix I A.

1) Investors' preferences depend only on the probability distribution of terminal wealth derived from the investment conditional on the distribution of terminal wealth derived from their existing portfolio.

The non-satiation and perfect capital markets assumptions remain unchanged. Given this new set of assumptions, it is possible to derive an allocation rule for the cheapest payoff function similar to the univariate case[10]. This time, however, the rule depends on the value of the investor's existing portfolio, which makes it a little more awkward to incorporate in the replication procedure. We will return to this issue in the next section.

Another important paper in this context is Cox and Leland (2000). The latter show that in a Black-Scholes (1973) world all path-dependent payoff functions are inefficient because they generate payoff distributions that can also be obtained with a path-independent payoff function, but at lower costs. Our replicating payoffs will therefore not only have to allocate terminal wealth in a specific manner, but always be path-independent as well.

Strictly speaking, the above results are only valid in the relatively simple theoretical setting from which they are derived. For the purpose of our replication procedure, however, we will assume that they are also valid in a more complex world where asset returns may be non-normally distributed, with highly unusual patterns of dependence. Of course, this need not be true. Unfortunately, without the availability of a more sophisticated theoretical framework, this is the best one can do.

3. Determination of the Replication Strategy

The replication procedure consists of a number of distinct steps. First, we collect return data on the fund to be replicated, the investor's portfolio, and the reserve asset (see Appendix I B). Second, we analyse the data to infer the joint distribution of the fund return and the investor's portfolio return. We refer to this as the 'desired distribution'. We do the same for the joint distribution of the investor's portfolio

[10] Proof is provided in Appendix I B.

return and the return on the reserve asset, which we refer to as the 'building block distribution'. Third, we determine the cheapest payoff function, which turns the building block distribution into the desired distribution. Fourth, we price the latter payoff function. Fifth, we derive the required allocations to the investor's portfolio and the reserve asset from the resulting value function.

In this section we discuss the above steps in more detail. Before we do so, however, we provide a brief introduction to copulas and their use in multivariate dependence modelling. As will become clear, copulas are a crucial ingredient in the replication procedure as they allow us to easily capture a large variety of non-normal dependence structures.

Copulas

Recent research in finance has uncovered various deviations from not only univariate, but also multivariate normality[11]. One powerful and at the same time convenient way to model this is by the use of copulas, as it allows the decomposition of any *n*-dimensional joint distribution into *n* marginal distributions and a single copula function[12]. Assume a random vector of two random variables. A bivariate copula can then be defined as follows.

Definition 1: *The copula C of the random vector (X,Y) is the joint distribution of the random vector (U, V), where $U = F_X(X)$ and $V = F_Y(Y)$, and where F_X, F_Y, are the distribution functions of X and Y respectively.*

The above definition implies that:

$$F_{X,Y}(x,y) = C(F_X(x), F_Y(y)),\ \forall x \in \Re, y \in \Re\,, \tag{1}$$

[11] Longin and Solnik (2001) for example find clear evidence of asymmetric dependence in international equity markets. A similar conclusion can be found in Ang and Chen (2002) with respect to US stocks.

[12] Copulas have been widely used in the statistical literature. Joe (1997) and Nelsen (1999) provide a good introduction. Cherubini et al. (2004) discuss copulas in a finance context.

where F_{XY} is the joint distribution of the random vector *(X,Y)*. Intuitively, the copula function divides the characteristics of the joint distribution between the marginal distributions, which contain the univariate characteristics of each random variable, and the copula, which contains all information concerning the dependence between these random variables.

Next, we present a key result in copula theory[13]. Let $\overline{\Re} = \Re \cup \{-\infty, \infty\}$ denote the extended real line.

Sklar's Theorem: *Let $F_{X,Y}$ be a 2-dimensional joint distribution function with marginal distributions F_X, and F_Y. Then there exists a copula C such that for all (x,y) in $\overline{\Re}^2$ $F_{X,Y}(x,y) = C(F_X(x), F_Y(y))$. If F_X and F_Y are continuous then C is unique; otherwise, C is uniquely determined on Ran(F_X) $\times$ Ran(F_Y). Conversely, if C is a copula and F_X, F_Y are distribution functions, then the function F_{XY} defined by (1) is a joint distribution with margins F_X, F_Y.*

From a multivariate financial modelling perspective, it is the converse of Sklar's Theorem that is most interesting, as it implies that any combination of two univariate distributions and a copula defines a valid bivariate distribution. This solves the problem that in statistics, although we do have a large set of flexible parametric univariate distributions available, the set of parametric multivariate distributions is quite limited.

Estimation of the Desired and Building Block Distributions

In the replication procedure we allow three different marginal distributions (Normal, Student-t and Johnson SU)[14] and six different bivariate copulas. The first two copulas are part of the class of *elliptical copulas*, since they are derived from elliptical distributions. The *normal copula* is extracted from the bivariate normal distribution. If we combine the bivariate normal copula with two normal marginal distributions, we end up with the bivariate normal distribution. However, if either one or both marginal

[13] Proof of this theorem can be found in Nelsen (1999, p. 18).

[14] See Johnson (1949, 1965) for details on the Su distribution.

distributions are non-normal, then the joint distribution produced will be a completely different distribution. The *Student-t copula*, which is extracted from the bivariate Student-t distribution, is also an elliptical copula, but it differs from the normal copula in that it allows for some extreme dependence in the lower and upper tails. Since the Student-t copula is symmetric, however, this dependence must be the same for both tails.

The next three families of copulas, Gumbel, Cook-Johnson and Frank, are part of the *Archimedean* copulas class, a rich class of copulas that allows for very different types of dependence. The *Gumbel copula* is asymmetric. It has more dependence in the upper tail than in the lower tail. The *Cook-Johnson copula*, also known as the Clayton copula, is also asymmetric, but with more dependence in the lower tail than in the upper tail. As shown by Longin and Solnik (2001) and Ang and Chen (2002), this is quite common behaviour in equity market returns. The *Frank copula* implies the same dependence between positive returns as between negative returns. Like the Normal and Student-t copulas, it allows for positive and negative dependence. The sixth and final copula is the s*ymmetrised Joe-Clayton (SJC) copula*, proposed by Patton (2005a). It is the most flexible of the copulas discussed here. It has two parameters, which separately control the dependence in the lower and upper tail. As a result, this copula can fit data with very different patterns of dependence in the tails.

<< Insert Figure 1 Here >>

Figure 1 shows 500 simulated drawings from six bivariate joint distributions. In all cases, the marginal distributions are standard normal and the linear correlation is 0.7. Despite this, the plots show six different patterns of dependence, underlining the impact and different characteristics of each of the six copula families. Only in the bivariate normal case is the linear correlation coefficient sufficient to fully describe the observed dependence structure[15].

[15] Kat (2003) discussed this point in a hedge fund context.

The estimation method that we use is known as the Inference Functions for Margins (IFM) method[16]. It is a two-step maximum likelihood method. Let *(X,Y)* be a vector of two random variables with joint distribution function F_{XY} and marginal distribution functions F_X and F_Y respectively. The marginal distribution F_X depends only on the set of parameters Θ_X and the same for F_Y and Θ_Y. Let Θ_C be the vector of parameters of the bi-dimensional copula C. So the unknown vector of parameters is given by $\Theta = (\Theta_X, \Theta_Y, \Theta_C)$. We know from Definition 1 that $F_{XY}(x,y;\Theta) = C(F_X(x;\Theta_1), F_Y(y;\Theta_2);\Theta_C)$. So the joint distribution F_{XY} is completely specified by the vector of parameters *Θ*. Differentiating with respect to both variables, we have $f_{XY}(x,y) = c(F_X(x), F_Y(x)) f_X(x) f_Y(y)$, where $c(u,v) = \frac{\partial C(u,v)}{\partial u \partial v}$ is the copula density.

For a bivariate random sample of size T $\{(x_i, y_i)\}_{i=1}^T$, the log-likelihood function is therefore given by:

$$l(\Theta) = \sum_{t=1}^{T} \ln c(F_X(x_t;\Theta_X); F_Y(y_t;\Theta_Y);\Theta_C) + \sum_{t=1}^{T} \ln f_X(x_t;\Theta_X) + \sum_{t=1}^{T} \ln f_Y(y_t;\Theta_Y).$$

Estimating all parameters at the same time would be very cumbersome and time-consuming. We therefore do so in two consecutive steps. First, we estimate the marginal set of parameters Θ_X and Θ_Y (separately) by maximum likelihood. Subsequently, we create the series $\hat{u}_t = F_X(x_t;\hat{\Theta}_X)$ and $\hat{v}_t = F_Y(y_t;\hat{\Theta}_Y)$ and estimate Θ_C by maximum likelihood using the likelihood function $l(\Theta_C) = \sum_{t=1}^{T} \ln c(\hat{u}_t;\hat{v}_t;\Theta_C)$.

With three possible candidates for the marginal distribution and six for the copula, we have 54 possible joint distributions to choose from. To select the final model, we use the Akaike information criterion (AIC)[17]. We considered some other selection criteria as well, including the quadratic distance between the estimated copula and the

---

[16] See Xu (1996) and Patton (2005b) for details on the statistical properties of this method.

[17] See Akaike (1973) for details.

empirical copula for example. The advantage of the AIC, however, is that it penalises models with a large number of parameters.

Determination of the Desired Payoff Function

Having selected the desired and building block distributions, the next step is to determine the cheapest payoff function, which turns one into the other, i.e. the cheapest function g^* such that:

$$P(S_P \leq x, g^*(S_P, S_R) \leq y) = P(S_P \leq x, S_I \leq y), \forall x, y, \quad (2)$$

with S_I denoting the end-of-month payoff of the fund, S_P the end-of-month payoff of the investor's portfolio, and S_R the end-of-month payoff of the reserve asset.

We start by assuming the current value of all assets is equal to 100. Rescaling to log-returns, this means looking for the cheapest function $g(x,y) = \log\left(\frac{g^*(100\exp(x), 100\exp(y))}{100}\right)$ such that:

$$P(X_P \leq x, g(X_P, X_R) \leq y) = P(X_P \leq x, X_I \leq y) = F_{P,I}(x,y), \forall x, y, \quad (3)$$

with $X_I = \log\left(\frac{S_I}{100}\right)$, $X_p = \log\left(\frac{S_P}{100}\right)$, and $X_p = \log\left(\frac{S_P}{100}\right)$. Or equivalently, the cheapest function g such that:

$$P(g(X_P, X_R) \leq y \mid X_P = x) = P(X_I \leq y \mid X_P = x) = F_{I|P}(y \mid x), \forall x, y \quad (4)$$

From Appendix I B., we know that the cheapest payoff function depends on the conditioning value x. As a result, the bivariate function g may not be a 'smooth' function, i.e. the derivatives of this function will 'jump' around the line $x=x_{min}$, making the execution of the replication strategy derived from the payoff function quite awkward. From Appendix I B. we find that the desired payoff function should only be a non-decreasing function of the reserve asset if

$$\frac{\mu_R - r}{\sigma_R} > \rho\left(\frac{\mu_P - x}{\sigma_P}\right), \tag{5}$$

for $\forall\rho \in [-1,1]$. The expression on the left is nothing more than the Sharpe ratio of the reserve asset. From (5) it therefore follows that as long as the Sharpe ratio of the reserve asset is high enough and the correlation with the investor's portfolio low enough, the desired payoff function should be a non-decreasing function of the reserve asset.

Assuming the reserve asset satisfies the above condition[18], the function g in expression (4) is given by:

$$g(x, y) = F_{I|P}^{-1}(F_{R|P}(y \mid x) \mid x), \forall y \in \Re \tag{6}$$

where $F_{I|P}^{-1}(y \mid x)$ denotes the pseudo-inverse of $F_{I|P}(y \mid x)$. This is a composed function, with two non-decreasing components. The composition is therefore also non-decreasing, as required.

Next, we have to prove that (4) holds:

$$\begin{aligned} &P(g(X_P, X_R) \leq y \mid X_P = x) = P(g(x, X_R) \leq y \mid X_P = x) = \\ &P(F_{I|P}^{-1}(F_{R|P}(X_R \mid x) \mid x) \leq y \mid X_P = x) = P(F_{I|P}^{-1}(U \mid x) \leq y \mid X_P = x), \end{aligned} \tag{7}$$

where $U \sim \text{Uniform}[0,1]$ by the probability integral transformation. Then, by the same reasoning, $F_{I|P}^{-1}(U \mid x)$ has the same distribution as X_I given $X_P = x$, so we finally have:

$$P(F_{I|P}^{-1}(U \mid x) \leq y \mid X_P = x) = P(X_I \leq y \mid X_P = x) = F_{I|P}(y \mid x), \tag{8}$$

and (4) holds as required.

[18] Extensive simulations showed that, under reasonable assumptions, this does not introduce any significant error if not true for some values of x.

In order to obtain the function g, we need to model the conditional distributions $F_{I|P}$ and $F_{R|P}$. Let $C_{P,I}$ denote the copula between X_P and X_I and let $C_{P,R}$ denote the copula between X_P and X_R. Then from (1) we have:

$$F_{P,I}(x,y) = C_{P,I}(F_P(x), F_I(y)), x \in \Re, y \in \Re. \tag{9}$$

$$F_{P,R}(x,y) = C_{P,R}(F_P(x), F_R(y)), x \in \Re, y \in \Re. \tag{10}$$

We can write the conditional distributions $F_{I|P}$ and $F_{R|P}$ as:

$$F_{I|P}(y \mid x) = \kappa_x^{P,I}(y), x \in \Re, y \in \Re, \text{ where } \kappa_x^{P,I}(y) = \frac{\partial C_{P,I}(u,v)}{\partial u}\Big|_{u=F_p(x), v=F_I(y)}$$

$$F_{R|P}(y \mid x) = \kappa_x^{P,R}(y), x \in \Re, y \in \Re \text{ , where } \kappa_x^{P,R}(y) = \frac{\partial C_{P,R}(u,v)}{\partial u}\Big|_{u=F_p(x), v=F_R(y)} .$$

So the cheapest function g in expression (6) can be rewritten as:

$$g(x,y) = \kappa_x^{(-1)P,I}\left(\kappa_x^{P,R}(y)\right), x \in \Re, y \in \Re . \tag{8}$$

We can now rewrite everything in terms of the end-of-month payoff to obtain the desired payoff function. The end-of-month replicated values from a monthly initial investment of 100 will be equal to:

$$S_g = g^*(s_P, s_R) = 100 \exp g\left(\log\left(\frac{s_P}{100}\right), \log\left(\frac{s_R}{100}\right)\right). \tag{9}$$

Theoretically, the vector (S_P, S_g) will have the same joint distribution as the vector (S_P, S_I), meaning that, as intended, we are not only replicating the end-of-month payoff of the fund, but also its dependence with the investor's existing portfolio.

Pricing and Generating the Desired Payoff Function

Having determined the desired payoff function, the next step is to price it. This is of course not a new problem. It is what arbitrage-based option pricing theory has concentrated on for the last 35 years. Following Harrison and Kreps (1979), the desired payoff function can be priced by calculating the discounted risk neutral expected payoff. The two most obvious methods to do so are either bivariate Monte Carlo simulation or a trinomial tree[19]. Once we are able to price the desired payoff function, we can work out the controls of the dynamic trading strategy generating it by straightforward partial differentiation of the value function.

Two points are worth noting at this stage. First, only after pricing the payoff function do we know what the expected return on the replicating strategy will be. The desired payoff function explicitly aims to replicate all aspects of the desired distribution, except the fund's expected return. The latter follows from the expected return on the investor's portfolio and the reserve asset, the desired payoff function, and the pricing environment for the latter, i.e. interest rates, expected dividends, volatilities, etc. In other words, it is the capital market that sets the expected return on the replicating strategy. Second, although determined in a much more flexible setting, the desired payoff function is priced in the standard model where asset returns are normally distributed. As long as we don't have access to a more sophisticated theoretical pricing model, we cannot escape this inconsistency.

4. Simulation Analysis

Given the desired and building block distributions, the above results allow us to derive, price and generate the cheapest payoff function that turns one into the other. The procedure is exact, so by itself it does not require any testing. Taking this procedure into the real world and using it to replicate fund returns, however, we are confronted with a number of problems. First, we do not know the true population distribution. The best we can do is estimating it from a small data sample. Second, the latter distribution may not be stationary over time. Third, due to market imperfections and insufficient information on the underlying price processes, we may not always be able to exactly generate the desired payoff function.

---

[19] See Jaeckel (2002) or Glasserman (2003) for an introduction to Monte Carlo methods. Details on the trinomial tree approach can be found in He (1990).

In this section, we use simulation methods to study the error resulting from determining the desired payoff function from a relatively small sample, instead of from the population distribution. In these simulations, we assume that the population distribution is stationary and that it is possible to generate the desired payoff function without any error. In the next section we perform a number of out-of-sample tests on real-life data to also include the error contributions of non-stationarity and sub-optimal dynamic trading. In the simulations, we study two different cases, selected to capture different distributional conditions. Throughout, we assume that the returns on the investor's portfolio and the reserve asset are both normally distributed with the parameter values given below. In addition, we assume they are related through a Gaussian copula with a correlation coefficient of 0.3.

Investor's portfolio

Log-returns - $X_P \sim N(0.01, 0.043301^2)$

Mean = 12% p.a.

Volatility = 15% p.a.

Reserve asset

Log-returns – $X_R \sim N(0.00833, 0.028867^2)$

Mean = 10% p.a.

Volatility = 10% p.a.

Case 1: Gaussian fund marginal, dependence higher in the lower tail.

Our first case assumes that the fund return is normally distributed, but that the relationship with the investor's portfolio is such that there is more dependence in the lower than in the upper tail. This could be the risk profile of a fund of funds with a bias towards risk arbitrage for example. The marginal distribution of the fund return and the relevant copula are specified as follows:

Fund

Log-returns – $X_I \sim N(0.015, 0.057735^2)$

Mean = 18% p.a.

Volatility = 20% p.a.

Copula (investor's portfolio, fund) = SJC (0.75, 0.10)

<< Insert Figure 2 and 3 Here >>

Give the desired and building block distributions, we derived the desired payoff function using the results of section 3. Figure 2 and 3 depict the latter graphically, as a contour plot as well as a 3D graph. From the graphs we see that the desired payoff is an increasing function of the reserve asset (by construction) as well as the investor's portfolio. The strategy's controls will therefore tell us to hold long positions in both assets. As is especially clear from the contour plot, the payoff function is quite curved. This of course serves to generate the required difference in dependence between the upper and lower tail.

To gain insight into the potential error when deriving the desired payoff function from only a small sample, instead of the population distribution, we took a sample of size N and derived a payoff function from it. Subsequently, we took 2000 observations from the building block distribution, and fed these observations through the latter payoff function to produce a joint distribution of replicated payoffs and the investor's portfolio. From the latter distribution we calculated the mean, standard deviation, skewness, and kurtosis of the replicated payoff as well as its correlation with the investor's portfolio. The above procedure was repeated 100 times, for different values of N (= 24, 48, 72, 96, 120, 240). Across each set of 100 runs, we subsequently calculated the mean, standard deviation and skewness of the replication errors, i.e. the differences between the above sample statistics and the true fund parameters. The results can be found in table 1.

<< Insert Table 1 Here >>

To be able to properly interpret the entries in the table, the first row in Table 1 shows the mean, standard deviation, skewness, kurtosis and correlation of the fund payoff, as implied by the assumed fund return distribution. The rows that follow show, for various sample sizes N (= 24, .., 240) and each over 100 runs, the average (Avg), standard deviation (SD) and skewness (SK) of the replication errors. Table 1 confirms that the larger the sample, the more accurate the desired payoff function will be. It also shows that even with a relatively small sample the procedure still works quite well and is unbiased. For all parameters and sample sizes, the average error is statistically insignificant at 5% (-1.96 SD, + 1.96 SD).

Case 2: Negatively skewed fund marginal, Gaussian copula.

Our second case is somewhat more extreme. It assumes that the fund return exhibits a high degree of negative skewness. To make up for that, however, it also has a relatively high mean and low correlation with the investor's portfolio. With a little imagination, this could be the risk profile of a fixed income arbitrage fund for example. The marginal distribution of the fund return and the relevant copula are specified as follows:

Fund

Log-returns – X_I ~ Johnson-SU (0.058604, 0.046978, 0.926426, 1.390468)

Mean = 18% p.a.

Volatility = 20% p.a.

Skewness = -2.0

Excess kurtosis = 10

Copula (investor's portfolio, fund) = Gaussian (0.2).

<< Insert Figure 4 and 5 Here >>

From the above population distributions we again derived the desired payoff function, which is graphically depicted in Figure 4 and 5. As required, the payoff is a positive function of the reserve asset. However, since the assumed correlation between the fund and the investor's portfolio is lower than the assumed correlation between the investor's portfolio and the reserve asset, the payoff is a negative function of the investor's portfolio. The strategy's controls will therefore want us to go long in the reserve asset, but short in the investor's portfolio. The slope of the payoff function increases as the investor's portfolio rises and the reserve asset drops, which serves to generate the required negative skewness.

<< Insert Table 2 Here >>

To gain insight into the potential error from deriving the desired payoff function from a small sample in this particular case, we repeated the procedure used earlier in case 1. The results can be found in Table 2. Not unexpectedly, given the much more extreme distributional assumptions, we see quite some variation for small sample sizes. Especially the replication of the assumed –2.0 skewness meets with some difficulty. Since, by definition, tail events only occur infrequently, many smaller samples will not contain enough information to estimate skewness accurately. This is reflected by the strong positive skew of the error distribution for small N.

The above two case studies suggest that, depending on the distributions involved, the error from working with a small sample may sometimes be quite substantial. It is important to note though, that when applying the procedure in practice, one will typically re-estimate the payoff function periodically as new fund return data becomes available. Through time therefore, these errors may diversify away to some extent.

5. Out-of-Sample Tests

We proceed with some out-of-sample tests. Taking the replication procedure into the real world introduces a new set of problems. Where the model assumes continuous trading at zero costs, we will necessarily have to trade discretely, pay commissions and possibly be confronted with significant market impact. In addition, where the model assumes all relevant parameters to be known, we are confronted with a significant degree of uncertainty about future parameter values. Fortunately, these problems are not new. They are characteristic to all model-based dynamic trading strategies. A number of authors have studied and suggested solutions to the above problems[20]. None of these, however, appears to be able to improve the efficiency of dynamic trading strategies to a very large extent. We therefore assume the simplest possible set-up. If our replication strategies do not work under these conditions, it is unlikely they will work in a more elaborate set-up.

The out-of-sample tests that follow are all structured in the same way. Given a fund, we take the first 24 months of its track record as given, assuming we do not know

[20] See for example, Boyle and Emanuel (1980), Leland (1985), Figlewski (1989), Kat (1996) or Clewlow and Hodges (1997).

anything about what is to come. If a fund's track record starts in January 1985 for example, we assume to be living on January 1st, 1987. Subsequently, we determine the desired payoff function from the available 24 monthly returns, calculate the accompanying strategy controls and set up the required positions. During the month, we adjust our portfolio on a daily basis, driven by the daily changes in the underlying index values. At the beginning of the next month, we include the hedge fund return over the previous month in our dataset and repeat the whole procedure, now using 25 monthly returns instead of 24[21]. The above is repeated until we arrive at the end of October 2004 (where our fund database ends).

Throughout we assume the investor's portfolio consists of 50% US equity, in the form of the S&P 500 tracking portfolio, and 50% long-dated US Treasury bonds. We use nearby Eurodollar futures as the reserve asset[22]. To minimize transaction costs, all trading is done in the futures markets[23]. Transaction costs for all futures contracts are assumed to be 1bp one-way. The necessary volatility and correlation inputs are obtained from historical estimates, using all available data at the time of determining the desired payoff function.

In what follows we discuss the out-of-sample replication results for three different hedge funds (of funds). We selected these funds because they are well known within the industry and among investors and because they have relatively long track records[24]. The latter requirement stems from the fact that when comparing the statistical properties of the fund and the replicated returns we are basically comparing two bivariate distributions, which is best done using as many data points as possible. All fund returns are net of fees and were taken from the TASS database, with all data

---

[21] In practice hedge funds typically take one or two weeks to report their end-of-month net asset value. For simplicity, we refrain from this complication here.

[22] The decision to use Eurodollar futures is primarily based on liquidity considerations. Research into the characteristics of the optimal reserve asset is ongoing, however, and may lead us to change the reserve asset in a later version of this paper.

[23] S&P 500 (SP) and Eurodollar futures (ED) are traded on the CME, while T-bond futures (US) are traded on the CBOT.

[24] More in particular, we did not select these funds because the replication procedure works especially well for them, nor do we mean to boost or damage these managers' business in any way.

series ending per October 2004. We do not charge any management fees in the replication strategy.

Several studies have shown that reported monthly hedge fund returns may exhibit highly significant levels of autocorrelation[25]. This primarily results from the fact that many hedge funds invest in illiquid securities, which are often hard to mark to market. When confronted with this problem, hedge fund administrators will either use the last reported transaction price or a conservative estimate of the current price, which creates artificial lags in the evolution of hedge funds' net asset values, resulting in artificial smoothing of the reported monthly returns. As a result, estimates of hedge fund volatility for example, can be biased downwards by 30-40% in some cases.

One possible way to correct for the above autocorrelation is found in the real estate finance literature. Due to smoothing in appraisals and infrequent valuations of properties, the returns on direct property investment indices suffer from similar problems as hedge fund returns. The approach employed in the literature has been to "unsmooth" the observed returns to create a new set of returns which are more volatile and whose characteristics are believed to more accurately capture the characteristics of the underlying property values. Nowadays, there are several unsmoothing methodologies available. We use the method originally proposed by Geltner (1991).

Leveraged Capital Holdings N.V.

Our first example concerns one of the first funds of hedge funds. Leveraged Capital Holdings (LCH) was started in 1969 (our return data, however, only start in 1985) by Georges Karlweis of Banque Privee Edmond de Rothchild in Geneva. Over the years, LCH has (been) invested in all well-known hedge fund managers, such as George Soros, Martin Zweig and Joseph DiMenna, and Michael Steinhardt. LCH is publicly listed on the Amsterdam Stock Exchange and currently has $1.32 billion under management (TASS, October 2005).

<< Insert Figure 6-7 Here >>

[25] See for example Brooks and Kat (2002) or Lo et al. (2004).

Figure 6 and 7 show the payoff function used for the replication of the LCH return per October 2004 (the last month for which we have fund return data available)[26]. Notice that the LCH payoff function is a lot more 'lively' than the payoff functions encountered in the previous section. This underlines the complexity of real-life hedge fund returns. The graphs show that the desired payoff is a positive function of the investor's portfolio as well as the reserve asset, implying that the replication strategy will be long in both assets. We also see quite some variation in the slope of the payoff surface. Since the controls of the replication strategy are nothing more than the slope coefficients of the payoff value function, this signals the presence of 'hot spots', where relatively small changes in the investor's portfolio and/or reserve asset will generate relatively large changes in the strategy's controls.

<< Insert Figure 8 Here >>

The left hand side of Figure 8 shows a scatter plot of the monthly returns on the investor's portfolio versus the LCH returns. The right hand side of Figure 8 shows a scatter plot of the monthly returns on the investor's portfolio versus the replicated returns. Comparing both plots, we see that they are very similar, which indicates that the replication strategy is indeed able to successfully replicate LCH's returns' statistical properties. We also see that the replication strategy is unable to replicate the three large losses that LCH reported in October 1987 (-22.52%), August 1998 (-11.45%) and April 2000 (-10.83%). Since these are clearly outliers, it is not surprising that the replication procedure was unable to capture them out-of-sample. Given the size of these losses, it is unlikely investors will consider this a real shortcoming though.

<< Insert Table 3 Here >>

Another indication of the accuracy of the replication strategy comes from comparing the actual mean, standard deviation, skewness and kurtosis of LCH's returns with those of the replicated returns. The latter statistics can be found in Table 3, together

[26] The jagged profile in the bottom right-hand corner of the contour plot is due to some numerical instability in the extremes. A similar phenomenon is observed in the two cases that follow.

with the correlation and Kendall's Tau with the investor's portfolio. Since LCH's returns exhibit some clear outliers, apart from the standard skewness and kurtosis measures we also report more robust skewness and kurtosis measures[27]. To test whether the marginal distribution of the replicated returns and the joint distribution of the replicated returns and the investor's portfolio are significantly different from the original distributions, we use the univariate and bivariate Kolmogorov-Smirnov (K-S) tests[28].

Comparing the entries in Table 3, it is clear that, despite the obvious limitations, the statistical properties of LCH's returns have been quite successfully replicated. The replication strategy has not only replicated the marginal distribution of LCH's returns but also its relationship with the investor's portfolio. This is also the conclusion from both the K-S tests. Although slightly higher (14.76% pa versus 12.48% pa), the mean of the replicated returns is similar to that of the LCH returns as well. This confirms the assumption underlying the replication procedure that in the longer run investors receive a return which is in line with the risk profile they take on, irrespective of how that risk profile is acquired.

<< Insert Figure 9 Here >>

It is interesting to delve a bit further into the workings of the replication strategy. The left hand side of Figure 9 shows a scatter plot of the reserve asset returns versus the replicated returns. The positive relationship confirms the efficiency of the replication strategy (see section 2). The right hand side of Figure 9 shows a scatter plot of the fund returns versus the replicated returns. The plot makes it clear that although the replicated returns have statistical properties, which are very similar to those of LCH, they come to the investor in a completely different order. It is exactly this feature of the replication process, i.e. giving up the requirement that returns need to be similar on a month-to-month basis as well, which allows us to do so much better than the standard factor model approach.

---

[27] See Hinkley (1975) and Crow and Siddiqui (1967). These measures are also discussed in Kim and White (2004).

[28] See Fasano and Franceschini (1987) for details. Since the mean is not explicitly replicated, we subtract the mean from both the fund and the replicated returns before performing these tests.

<< Insert Figure 10 Here >>

Figure 10 shows the evolution of the replication strategy's controls over the period Dec. 2002 – Oct. 2004[29]. The graph confirms that the replication strategy holds long positions in both the investor's portfolio and the reserve asset. It also shows that the number of units of the reserve asset held is much higher than for the investor's portfolio. This is because the volatility of the Eurodollar future is quite low compared to that of LCH and the investor's portfolio. It therefore requires substantial leveraging. The strategy is quite dynamic, with the strategy's controls exhibiting a number of peaks and troughs. The latter are the result of a combination of strong inter-month index movement, a steep payoff function and monthly strategy resetting. For example, during April 2004 the value of the investor's portfolio dropped by almost 4%. As a result, the number of units of the investor's portfolio to hold rose from 0.90 at the start to 1.54 at the end of the month. At the same time, the number of units of the reserve asset to hold rose from 9.34 to 10.63. At the beginning of May, however, the strategy was reset to its starting values, meaning that the allocation to the investor's portfolio dropped to 0.85 units and the allocation to the reserve asset to 9.58 units.

Calamos Multi-Strategy Fund L.P.

The second example is a convertible arbitrage fund. The Calamos Multi-Strategy Fund (CMSF) was established in 1989 by convertible bond experts John and Nick Calamos. For most of its life CMSF has pursued a convertible arbitrage strategy. Since 2004, however, CMSF has adopted a long/short equity strategy as well. Managed primarily for the personal accounts of the Calamos family and a small group of friends, the fund is relatively small with currently $14.1 million under management (TASS, October 2005)[30].

<< Insert Figure 11 -14 and Table 4 Here >>

---

[29] The period Dec. 2002 – Oct. 2004 is representative for the period 1987 – 2004. A graph covering the full 1987-2004 test period would be too condensed to provide any worthwhile insights.

[30] Although the fund is only small, we decided to include it because the Calamos family is very well known for their work on convertibles and convertible arbitrage. See for example Calamos (1998) and Calamos (2003).

The desired payoff function for CMSF as per October 1st, 2004 is shown in Fig 11 and 12. At first sight, it looks similar to that for LCH, but, as is easiest seen from the contour plot, there are some significant differences as well. Figure 13 shows the same scatter plots as in Figure 8. Comparing both plots, we again see that they are very similar, indicating the replication strategy performs quite well. This is confirmed by the entries in Table 4. As before, all parameters are very similar, including the means and the correlation with the investor's portfolio. Both the univariate and bivariate K-S test confirm that there is no significant difference between the original and replicated distributions. Figure 14 shows the same scatter plots as in Figure 9. We again see a positive relationship between the reserve asset returns and the replicated returns, confirming the efficiency of the replication strategy. The plot of the fund returns versus the replicated returns shows a random scatter, making it clear that although the replicated returns have similar statistical properties as CMSF, they come in a completely different order.

Rocker Partners L.P.

Most hedge funds' returns are positively correlated with the equity market. Our final example therefore concerns a dedicated short seller, the returns of which are likely to be negatively correlated with the stock market. Rocker Partners (RP) was started in 1985 by David Rocker. While RP maintains both long and short positions, the general focus is on short selling. The fund is therefore popular with investors as a hedge against their long biased investments. RP currently has $611.1 million under management (TASS, October 2005).

<< Insert Figure 15 - 18 and Table 5 Here >>

The desired payoff function for RP as per October 1st, 2004 can be found in Figure 15 and 16. From these graphs we see that the payoff function for RP is quite different from what we found for LCH and CMSF. Of course, the payoff is a positive function of the reserve asset. The RP payoff, however, is a negative function of the investor's portfolio. The replication strategy will therefore go long in the reserve asset, but short the investor's portfolio. This is of course what one would expect for a short seller, whose returns are likely to be negatively correlated with the market. From Figure 17 and Table 5 we see that the replication strategy performs in the same way as before. The replicated return statistics are again similar to those of the fund returns. Even the negative correlation with the investor's portfolio is closely replicated. Figure 18 paints a similar picture as Figure 9 and 14.

6. Conclusion

Much of investors' current interest in hedge funds derives from the fact that traditional asset classes seem to lack opportunity these days. Stock markets are hesitant, bond prices will come down when interest rates go up again and the yield curve is flattening. With fresh memories of double-digit returns, this has driven investors towards commodities, emerging markets, credit-based structures, and of course hedge funds. Having generated high returns in the early years, the average return on hedge funds over the last 10-15 years has been quite impressive and many investors seem more than happy to use this as a guide for future returns. Given today's low interest rates, low risk premiums across the board, as well as the current size of the hedge fund industry itself, a repeat of the last 10-15 years is extremely unlikely, however.

Investing in alternatives comes with many drawbacks, including due diligence, liquidity, capacity, transparency and style drift problems, and excessive management and incentive fees. As long as investors believe they will be rewarded with (close to) double-digit returns, they will take these problems for granted. However, when reality kicks in and investors realize that hedge funds are no longer the money machines they once were (thought to be), their attitude will undoubtedly change. The above drawbacks will become more and more important and may ultimately become a reason to say farewell to hedge funds altogether and migrate to other alternative asset classes like emerging markets for example, which has shown stellar performance over the last 3 years. In fact, according to HFR, during the third quarter of 2005 funds of hedge funds were confronted with their first net outflow of funds, in the amount of $1.2 billion[31].

---

[31] Financial Times, Oct. 31, 2005. According to HFR, from the first to the second quarter of 2005 asset flows into hedge funds dropped by 60%, from $27.3 billion to $10.9 billion.

Although we can't create something out of nothing, in this paper we have shown that it is possible to design dynamic trading strategies, which generate returns similar to those of individual hedge funds and funds of hedge funds. Since this is accomplished by trading (futures on) traditional assets only, these strategies avoid the typical drawbacks surrounding hedge fund and other alternative investments. As such, our synthetic hedge fund returns are clearly to be preferred over real hedge fund returns.

Finally, it should be noted that the applications of the technique introduced here are not limited to replication only. The same technique can also be used for performance evaluation for example. When the average replicated return is significantly higher than the average fund return, that fund cannot claim superiority. After all, superior returns can't be replicated. As it essentially allows one to design trading strategies that generate returns with predefined statistical properties, the technique can also be used for the creation of completely new risk-return profiles. This means that investors no longer have to go through the usual process of finding and combining assets and funds in a costly and often unsuccessful attempt to construct a portfolio with the risk-return characteristics they require. Given the proper trading strategy, investors can now generate directly whatever risk-return profile they are after. Based on this technique, a whole new industry could develop! We will investigate these possibilities in more detail in two forthcoming companion papers[32].

[32] See Kat and Palaro (2006a, 2006b).

References

Akaike, H., Information Theory and an Extension of the Maximum Likelihood Principle, in: B. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory*, Academiae Kiado, Budapest, 1973, pp. 267-281.

Agarwal, V. and N. Naik, Risks and Portfolio Decisions Involving Hedge Funds, *Review of Financial Studies*, Vol. 17, 2004, pp. 63-98.

Amin, G. and H. Kat, Hedge Fund Performance 1990-2000: Do the Money Machines Really Add Value?, *Journal of Financial and Quantitative Analysis*, Vol. 38, June 2003, pp. 1-24.

Ang, A. and J. Chen, Asymmetric Correlations of Equity Portfolios, *Journal of Financial Economics,* Vol. 63, 2002, pp. 443-494.

Bailey, W., H. Li and X. Zhang, Hedge Fund Performance Evaluation: A Stochastic Discount Factor Approach, Working Paper Johnson School of Management, Cornell University, 2004.

Black, F. and M. Scholes, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy,* Vol. 81, 1973, pp. 637-654.

Boyle, P. and D. Emanuel, Discretely Adjusted Option Hedges, *Journal of Financial Economics,* Vol. 8, 1980, pp. 259-282.

Brooks, C. and H. Kat, The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors, *Journal of Alternative Investments,* Fall 2002, pp. 26-44.

Calamos, J., *Convertible Securities: The Latest Instruments, Portfolio Strategies, and Valuation Analysis*, McGraw-Hill, 1998.

Calamos, N., *Convertible Arbitrage: Insights and Techniques for Successful Hedging*, John Wiley & Sons, 2003.

Cherubini, U., E. Luciano, and W. Vecchiato, *Copula Methods in Finance,* John Wiley & Sons, 2004.

Clewlow, L. and S. Hodges, Optimal Delta-Hedging Under Transaction Costs, *Journal of Economic Dynamics and Control*, Vol. 21, 1997, pp. 1353 – 1376.

Cox, J. and H. Leland, On Dynamic Investment Strategies, *Journal of Economic Dynamics and Control,* Vol. 24, 2000, pp. 1859-1880.

Crow, E. and M. Siddiqui, Robust Estimation of Location, *Journal of the American Statistical Association*, Vol. 62, 1967, pp. 353-389.

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

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

Dyment, J., J. Olstein and A. Jones, *2005 Alternative Investment Survey*, Deutsche Bank Hedge Fund Capital Group, 2005.

Fasano, G. and A. Franceschini, *Monthly Notices of the Royal Astronomical Society*, Vol. 225, 1987, pp. 155-170.

Figlewski, S., Options Arbitrage in Imperfect Markets, *Journal of Finance*, Vol. 44, 1989, pp. 1289-1311.

Fung, W., D. Hsieh, N. Naik and T. Ramadorai, Hedge Funds: Performance, Risk and Capital Formation, Working Paper London Business School, 2005.

Geltner, D., Smoothing in Appraisal-Based Returns, *Journal of Real Estate Finance and Economics*, Vol. 4, 1991, pp. 327-345.

Glasserman, P., *Monte Carlo Methods in Financial Engineering*, Springer Verlag, 2003.

Glosten, L. and R. Jagannathan, A Contingent Claim Approach to Performance Evaluation, *Journal of Empirical Finance,* Vol. 1, 1994, pp. 133-160.

Harrison, J. and D. Kreps, Martingales and Arbitrage in Multi-Period Securities Markets, *Journal of Economic Theory,* Vol. 20, 1979, pp. 381-408.

He, H., Convergence from Discrete to Continuous-Time Contingent Claims Price, *Review of Financial Studies*, Vol. 3, 1990, pp. 523-546.

Hinkley, D., On Power Transformations to Symmetry, *Biometrika*, Vol. 62, 1975, pp. 101-111.

Jaeckel, P., *Monte Carlo Methods in Finance*, John Wiley & Sons, 2002.

Joe, H., *Multivariate Models and Dependence Concepts,* Chapman and Hall, 1997.

Johnson, N., Systems of Frequency Curves Generated by Methods of Translation, *Biometrika,* Vol. 36, 1949, pp. 149-176.

Johnson, N., Tables to Facilitate Su Frequency Curves, *Biometrika,* Vol. 52, 1965, pp. 547-558.

Kat, H., Delta Hedging of S&P 500 Options: Cash versus Futures Market Execution, *Journal of Derivatives*, Spring 1996, pp. 6-25.

Kat, H., The Dangers of Using Correlation to Measure Dependence, *Journal of Alternative Investments,* Fall 2003, pp. 54-58.

Kat, H. and H. Palaro, Superstars or Average Joes? A Replication-Based Performance Evaluation of 3000 Hedge Funds and Funds of Funds, Working Paper Alternative Investment Research Centre, Cass Business School, City University London, forthcoming 2006a.

Kat, H. and H. Palaro, Tell Me What You Want, What You Really, Really Want! Tailor-Made Fund Creation Using Copulas, Working Paper Alternative Investment Research Centre, Cass Business School, City University London, forthcoming 2006b.

Kim, T.H, and H. White, On More Robust Estimation of Skewness and Kurtosis: Simulation and Application to the S&P500 Index, *Finance Research Letters*, Vol. 1, 2004, pp 56-73.

Leland, H., Option Pricing and Replication with Transaction Costs*, Journal of Finance*, Vol. 40, 1985, pp. 1283-1301.

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

Longin, F. and B. Solnik, Extreme Correlation of International Equity Markets, *Journal of Finance*, Vol. 56, 2001, pp. 649-676.

Nelsen, R., *Introduction to Copulas,* Springer Verlag, 1999.

Patton, A., Modelling Asymmetric Exchange Rate Dependence, Working Paper London School of Economics, forthcoming in the *International Economic Review,* 2005a.

Patton, A., Estimation of Multivariate Models for Time Series of Possibly Different Lengths, Working Paper London School of Economics, forthcoming in the *Journal of Applied Econometrics,* 2005b.

Schneeweis, T., H. Kazemi and V. Karavas, Eurex Derivative Products in Alternative Investments: The Case of Hedge Funds, Working Paper CISDM, University of Massachusetts, 2003.

Sharpe, W., Asset Allocation: Management Style and Performance Measurement, *Journal of Portfolio Management*, Winter 1992, pp. 7-19.

Xu, J., *Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data*, PhD Thesis, University of British Columbia, 1996.

Appendix I

A. Univariate Case (See also Dybvig (1988b))

Suppose we lived in the standard model with one stock, one bond, perfect markets, a positive equity risk premium, and where the stock price follows a binomial tree with n steps, with time increments Δt. In one step, for each node, the stock price can move up by a factor $u = 1 + \mu\Delta t + \sigma\sqrt{\Delta t}$ or down by a factor $d = 1 + \mu\Delta t - \sigma\sqrt{\Delta t}$ with the same probability. After n steps, the initial price S_0 will have evolved into one of $n+1$ possible values $S_0 u^n$, $S_0 u^{n-1} d^1$, $S_0 u^{n-2} d^2$, ..., $S_0 d^n$, which we label as states 1, 2, ..., $n+1$ respectively. The bond returns $r\Delta t$ over each period, with r denoting the riskless rate.

What is the state price for each of the above $n+1$ states? Suppose we are at time $t=0$ and want to replicate the payoff of some particular investment. The stock price is S_0 and the bond value is B. Suppose the value of the investment one step ahead is either v_1 if the stock price goes up, or v_2 if it goes down. In order to replicate the investment, we need an investment of v_S shares and v_B bonds:

$$\begin{cases} S_0(1+\mu\Delta t + \sigma\sqrt{\Delta t})v_S + B(1+r\Delta t)v_B = v_1 \\ S_0(1+\mu\Delta t - \sigma\sqrt{\Delta t})v_S + B(1+r\Delta t)v_B = v_2 \end{cases}$$

Solving and reordering this result, we have the following one-period pricing relationship:

$$S_0 v_S + B v_B = \frac{1}{2(1+r\Delta t)}\left[1 - \frac{(\mu - r)\Delta t}{\sigma\sqrt{\Delta t}}\right]v_1 + \frac{1}{2(1+r\Delta t)}\left[1 + \frac{(\mu - r)\Delta t}{\sigma\sqrt{\Delta t}}\right]v_2 = p_1 v_1 + p_2 v_2,$$

where p_1 and p_2 are the state prices. Dividing the state prices by the probability of each state, which is ½, we obtain the following state-price density:

$$\rho_1 = \frac{1}{(1+r\Delta t)}\left[1 - \frac{(\mu - r)\Delta t}{\sigma\sqrt{\Delta t}}\right] \text{ and } \rho_2 = \frac{1}{(1+r\Delta t)}\left[1 + \frac{(\mu - r)\Delta t}{\sigma\sqrt{\Delta t}}\right]$$

Note that, assuming a positive risk premium, $S_0(1+\mu\Delta t+\sigma\sqrt{t})$ $> S_0(1+\mu\Delta t-\sigma\sqrt{t})$ and $\rho_1 < \rho_2$. If we repeat this for all nodes in the tree, we find that for the last step the state-price density is inversely related to the terminal values of the stock. In state *1*, the stock has the highest value S_0u^n , but the state-price density assumes the lowest value, while in state *n+1* the stock has the lowest value $S_0\,d^n$ but the state-price density function assumes the highest value.

Now we are set to use Theorem 1 of Dybvig (1988a). By this theorem, the cheapest payoff function allocates terminal wealth as a non-increasing function of the state-price density. Combining this with the above, the cheapest payoff function should therefore *allocate terminal wealth as a non-decreasing function of the value of the stock.*

B. Bivariate Case

Now assume that instead of one, we have two risky assets we can trade, which we will refer to as “the investor’s portfolio” and “the reserve asset”. The prices of both are denoted as S_P and S_R respectively. Following He (1990), we assume a trinomial tree for the joint behaviour of both prices. We denote the mean and standard deviation of the terminal wealth provided by the investor’s portfolio as μ_P and σ_P Likewise, we denote the mean and standard deviation of the terminal wealth provided by the reserve asset as μ_R and σ_R . ρ denotes the correlation coefficient between the two assets.

From assumption 1 (section 2) we know that all investors are concerned about is the conditional distribution of S_R given S_P. For each value $S_P = x$ we can therefore study the question how to allocate terminal wealth between states as a univariate problem. In other words, we do not need the entire trinomial tree, but only the conditional distribution.

We know that the distribution of R given S_R = x is a normal distribution with mean $\mu_R+\rho\frac{\sigma_R}{\sigma_P}(x-\mu_P)$ and standard deviation $\sigma_R\sqrt{1-\rho^2}$. We can use a binomial model to approximate this distribution. Doing so, we can perform the same analysis as in the univariate case to obtain the state-price density for the first step of the tree. Suppose that the initial price of the reserve asset is $S_{R,0}$. Then the prices for the two nodes are

$$S_{R,1} = S_{R,0}\left\{1+\left(\mu_R + \rho\frac{\sigma_R}{\sigma_P}(x-\mu_P)\right)\Delta t + \sigma_R\sqrt{1-\rho^2}\sqrt{t}\right\}$$

and

$$S_{R,2} = S_{R,0}\left\{1+\left(\mu_R + \rho\frac{\sigma_R}{\sigma_P}(x-\mu_P)\right)\Delta t - \sigma_R\sqrt{1-\rho^2}\sqrt{t}\right\}.$$

The state-price density is therefore given by:

$$\rho_1 = \frac{1}{(1+r\Delta t)}\left[1-\frac{(\mu_R+\rho\frac{\sigma_R}{\sigma_P}(x-\mu_P)-r)\Delta t}{\sigma_R\sqrt{1-\rho^2}\sqrt{\Delta t}}\right]$$

and

$$\rho_2 = \frac{1}{(1+r\Delta t)}\left[1+\frac{(\mu_R+\rho\frac{\sigma_R}{\sigma_P}(x-\mu_P)-r)\Delta t}{\sigma_R\sqrt{1-\rho^2}\sqrt{\Delta t}}\right]$$

Clearly, much depends on the value of ρ. Suppose $\rho > 0$. In that case $\rho_1 < \rho_2$ if $x > \mu_P + (r-\mu_R)\frac{\sigma_P}{\rho\sigma_R} = x_{\min}$, and $\rho_1 > \rho_2$ otherwise. In other words, the allocation rule for the cheapest payoff function will depend on the value of the investor's portfolio. If $x > x_{min}$, the rule is to allocate terminal wealth as a non-decreasing function of the value of the reserve asset, just as in the univariate case. If $x < x_{min}$, however, the rule is to allocate terminal wealth as a non-increasing function of the value of the reserve asset.

When $\rho < 0$, we see a similar phenomenon. In that case $\rho_1 < \rho_2$ if $x < \mu_P + (r - \mu_R)\dfrac{\sigma_P}{\rho \sigma_R} = x_{\max}$, and $\rho_1 > \rho_2$ otherwise. This means that if $x < x_{max}$, the rule for the cheapest payoff function is to allocate terminal wealth as a non-decreasing function of the value of the reserve asset. When $x > x_{max}$, however, the cheapest payoff function allocates terminal wealth as a non-increasing function of the value of the reserve asset.

	Mean	St. Dev	Skewness	Excess Kurtosis	Corr. with Portfolio
Fund	101.6805	5.8267	0.0000	-0.0447	0.7400
Differences (replicated - fund)					
Avg 24	-0.0315	0.0416	0.0174	0.0952	-0.0171
Avg 48	0.1010	-0.0090	0.0156	0.0835	-0.0071
Avg 72	0.1390	0.0689	0.0052	0.0843	-0.0020
Avg 96	0.0151	0.0114	0.0187	0.0781	-0.0148
Avg 120	0.0142	-0.0079	0.0038	0.0935	-0.0011
Avg 240	-0.0209	-0.0118	0.0053	0.0595	0.0015
SD 24	1.0386	1.0253	0.2052	0.4181	0.1234
SD 48	0.8287	0.6701	0.1206	0.1641	0.0933
SD 72	0.5975	0.5496	0.1099	0.1613	0.0726
SD 96	0.5023	0.5620	0.1299	0.1712	0.0656
SD 120	0.4733	0.4794	0.1026	0.1437	0.0576
SD 240	0.3352	0.2918	0.0884	0.1465	0.0402
SK 24	0.3765	-0.0241	0.9582	4.4101	-1.1399
SK 48	0.3071	0.6894	0.3059	0.1667	-0.8320
SK 72	-0.2267	-0.3145	0.0991	0.1401	-0.5869
SK 96	-0.0114	0.2892	-0.1812	0.4813	-0.7360
SK 120	-0.0748	0.4211	-0.1757	0.1810	-0.3530
SK 240	-0.0991	0.3519	-0.1098	-0.0260	-0.4008

Table 1: Variation due to payoff construction from small sample case 1.

	Mean	St. Dev	Skewness	Excess Kurtosis	Corr. with Portfolio
Fund	101.6805	5.8604	-1.9851	9.8511	0.2000
Differences (replicated - fund)					
Avg 24	-0.0798	0.3568	1.1919	-0.0162	-0.0454
Avg 48	0.0221	0.5960	1.2455	4.5453	-0.0341
Avg 72	-0.1577	0.3589	0.7876	-2.4691	0.0450
Avg 96	-0.0201	0.1341	0.7929	-1.9950	-0.0152
Avg 120	0.0300	0.0985	0.6538	-1.6140	0.0090
Avg 240	-0.0762	0.1846	0.5238	-1.0256	0.0086
SD 24	1.6094	2.0242	1.1313	19.8187	0.2882
SD 48	1.3526	1.5985	1.7192	41.6658	0.2205
SD 72	0.8821	1.2115	0.6218	2.9127	0.1374
SD 96	0.7701	1.1864	0.6591	3.5340	0.1385
SD 120	0.7622	0.9344	0.5907	2.7828	0.1136
SD 240	0.3818	0.7366	0.6266	4.4289	0.0812
SK 24	-0.2246	0.9298	3.9961	8.2333	-0.1064
SK 48	0.0573	0.5893	5.6316	8.5348	0.1809
SK 72	0.0405	0.6589	0.1550	-0.2631	-0.0549
SK 96	0.0785	0.7394	0.9729	4.0508	-0.4042
SK 120	-0.1274	0.4989	0.0159	0.2908	-0.1725
SK 240	0.0587	0.6556	1.0541	5.0328	0.0376

Table 2: Variation due to payoff construction from small sample case 2.

	Mean	St. Dev	Skewness	Skewness (robust)	Excess Kurtosis	Ex. Kurt. (robust)	Corr. with Portfolio	Kendall's Tau
LCH	**0.0095**	**0.0419**	**-1.9675**	**-0.1641**	**13.4015**	**0.3156**	**0.704**	**0.536**
Replica	**0.0125**	**0.0355**	**-0.3541**	**-0.1681**	**0.7021**	**0.5736**	**0.728**	**0.571**
Univariate K-S Statistic = 0.056, (approximated) p-value = 0.884								
Bivariate K-S Statistic = 0.053, (approximated) p-value = 0.968								

Table 3: Monthly return statistics Leveraged Capital Holdings and replication strategy, 1987 - 2004.

	Mean	St. Dev	Skewness	Skewness (robust)	Excess Kurtosis	Ex. Kurt (robust)	Corr. with Portfolio	Kendall's Tau
CMSF	**0.080**	**0.0213**	**0.2357**	**0.0154**	**2.6296**	**1.6937**	**0.509**	**0.337**
Replica	**0.094**	**0.0170**	**0.6656**	**0.0582**	**2.2128**	**0.9525**	**0.506**	**0.388**
Univariate K-S Statistic = 0.103, (approximated) p-value = 0.322								
Bivariate K-S Statistic = 0.087, (approximated) p-value = 0.719								

Table 4: Monthly return statistics Calamos Multi-Strategy Fund and replication strategy, 1991 - 2004.

	Mean	St. Dev	Skewness	Skewness (robust)	Excess Kurtosis	Ex. Kurt. (robust)	Corr. with Portfolio	Kendall's Tau
RP	**0.0058**	**0.0684**	**-0.2456**	**-0.0992**	**1.5588**	**1.3862**	**-0.302**	**-0.179**
Replica	**0.0083**	**0.0430**	**0.8377**	**-0.0385**	**5.0043**	**1.5521**	**-0.346**	**-0.196**
Univariate K-S Statistic = 0.117, (approximated) p-value = 0.101								
Bivariate K-S Statistic = 0.111, (approximated) p-value = 0.295								

Table 5: Monthly return statistics Rocker Partners and replication strategy, 1987 - 2004.

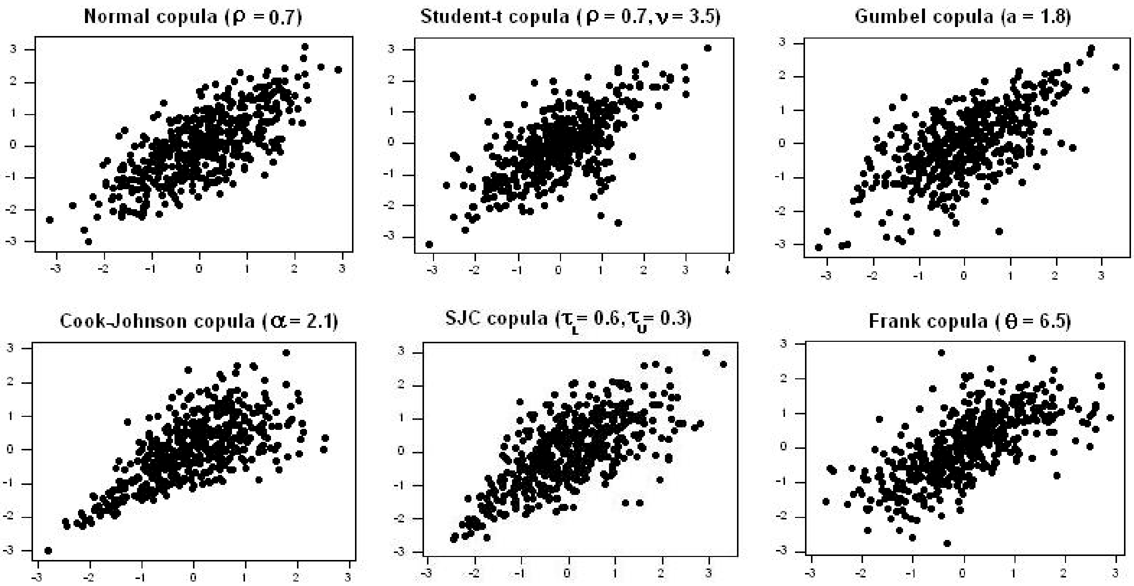

Figure 1. Random drawings from various copulas, assuming standard normal marginals and a linear correlation coefficient of 0.7.

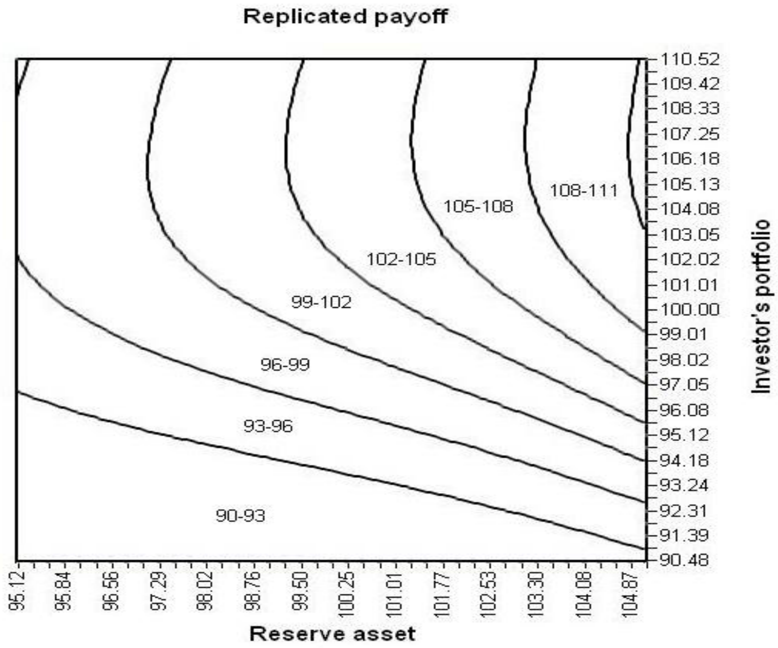

Figure 2. Contour plot payoff function from population distribution case 1.

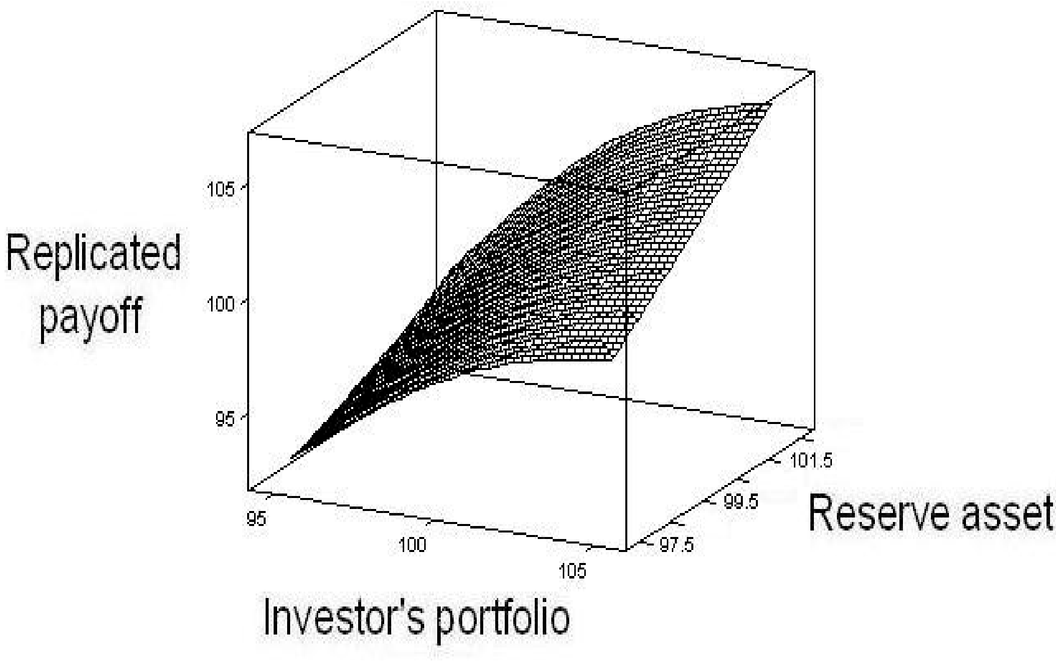

Figure 3. 3D plot payoff function from population distribution case 1.

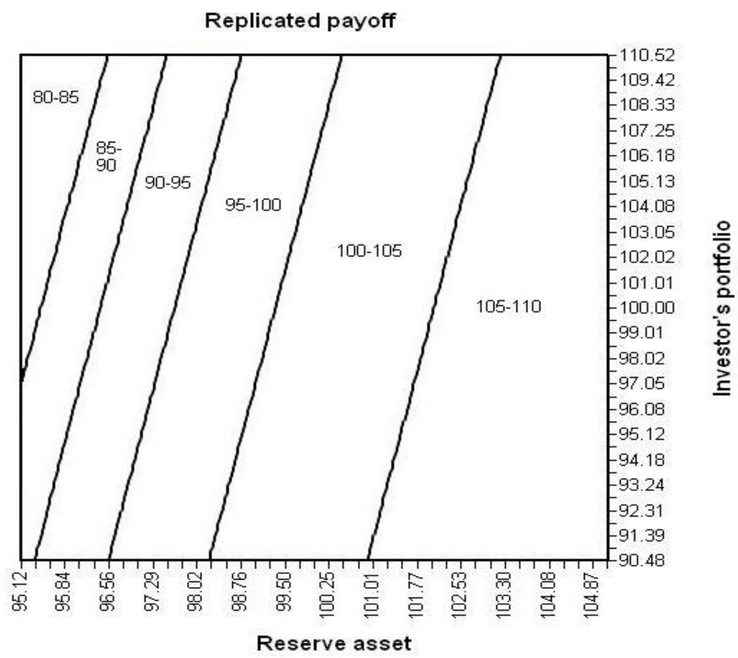

Figure 4. Contour plot payoff function from population distribution case 2.

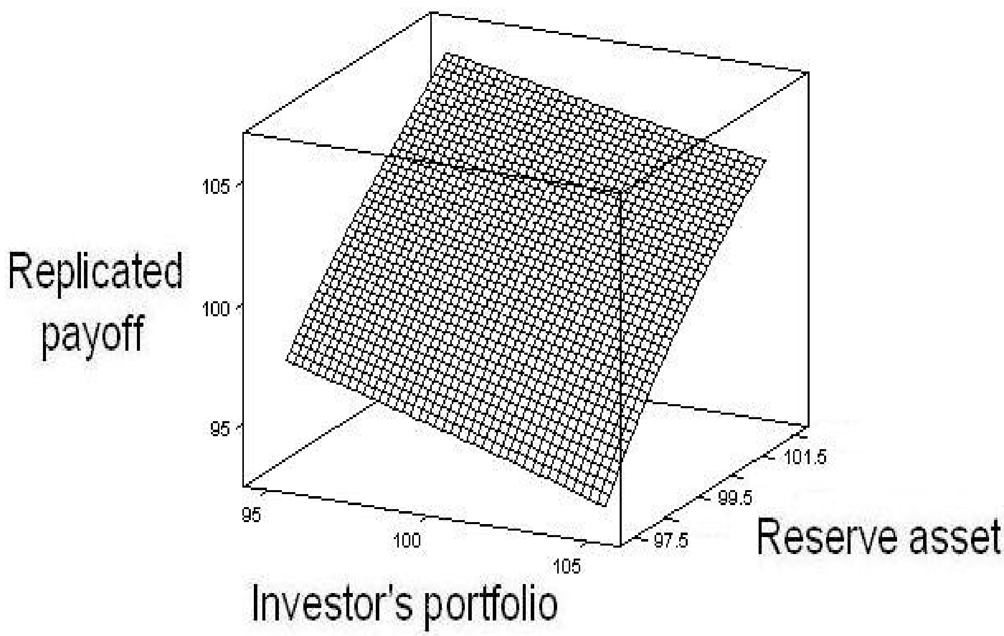

Figure 5. 3D plot payoff function from population distribution case 2.

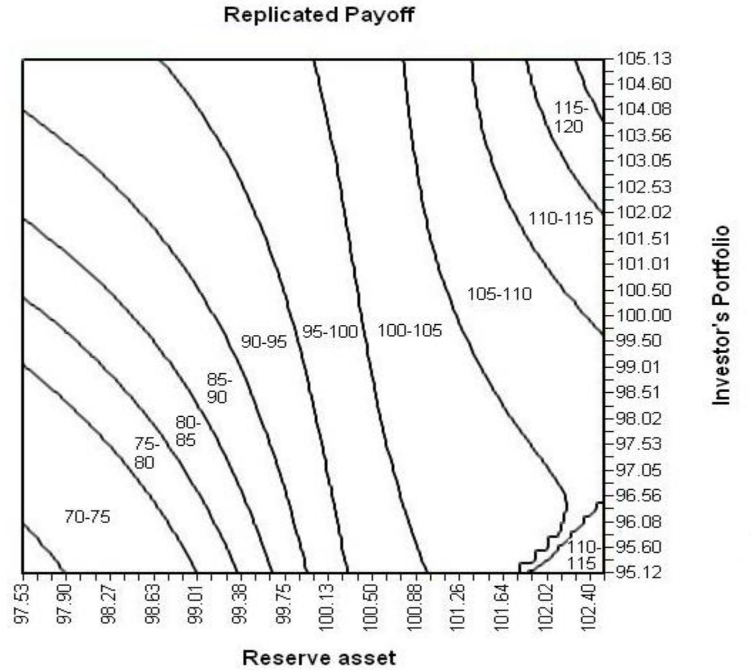

Figure 6. Contour plot payoff function Leveraged Capital Holdings.

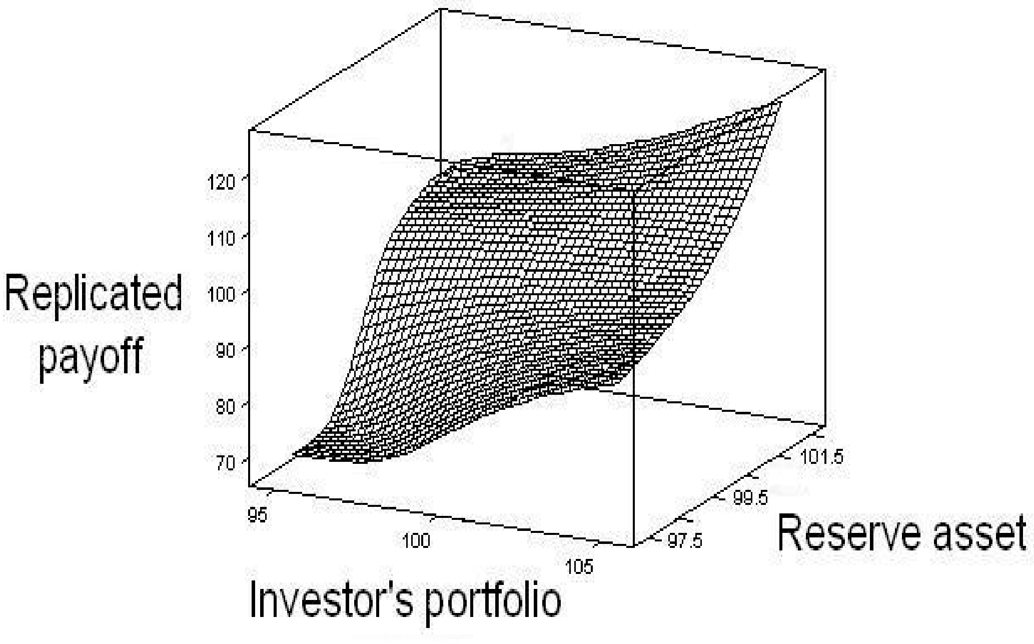

Figure 7. 3D plot payoff function Leveraged Capital Holdings.

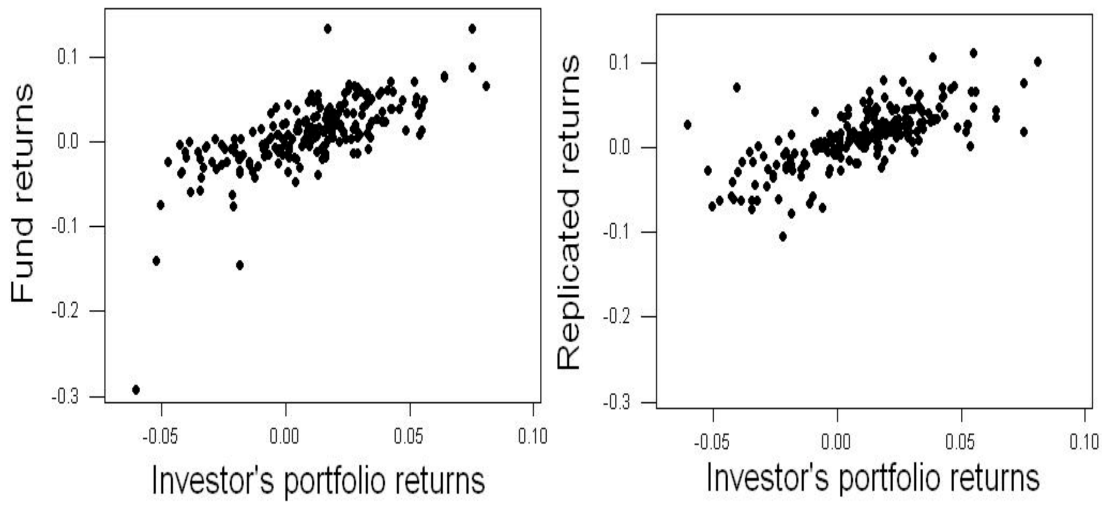

Figure 8. Scatter plot investor's portfolio returns versus Leveraged Capital Holdings returns (left) and replicated returns (right), 1987-2004.

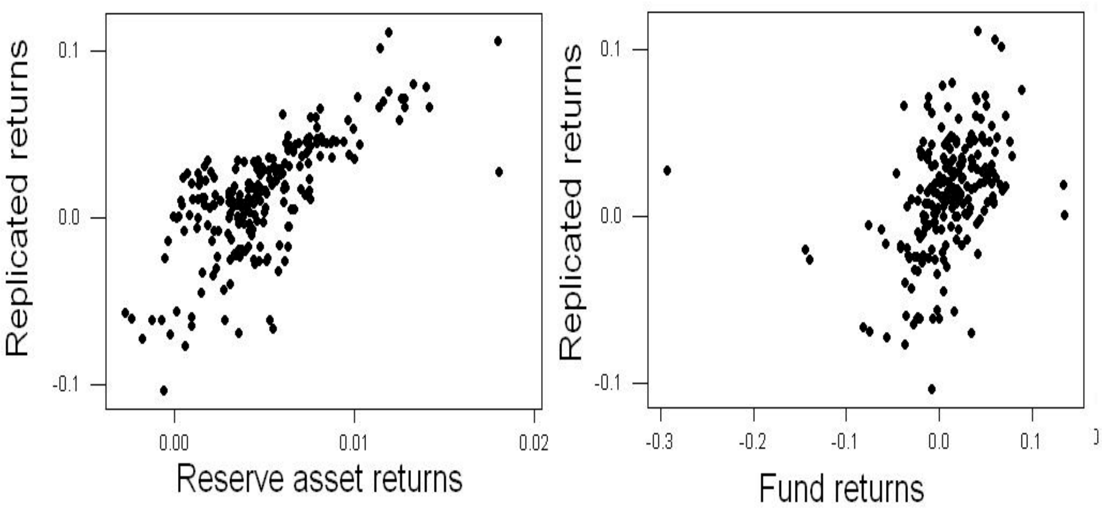

Figure 9. Scatter plot reserve asset returns versus replicated returns (left) and Leveraged Capital Holdings returns versus replicated returns (right), 1987 - 2004.

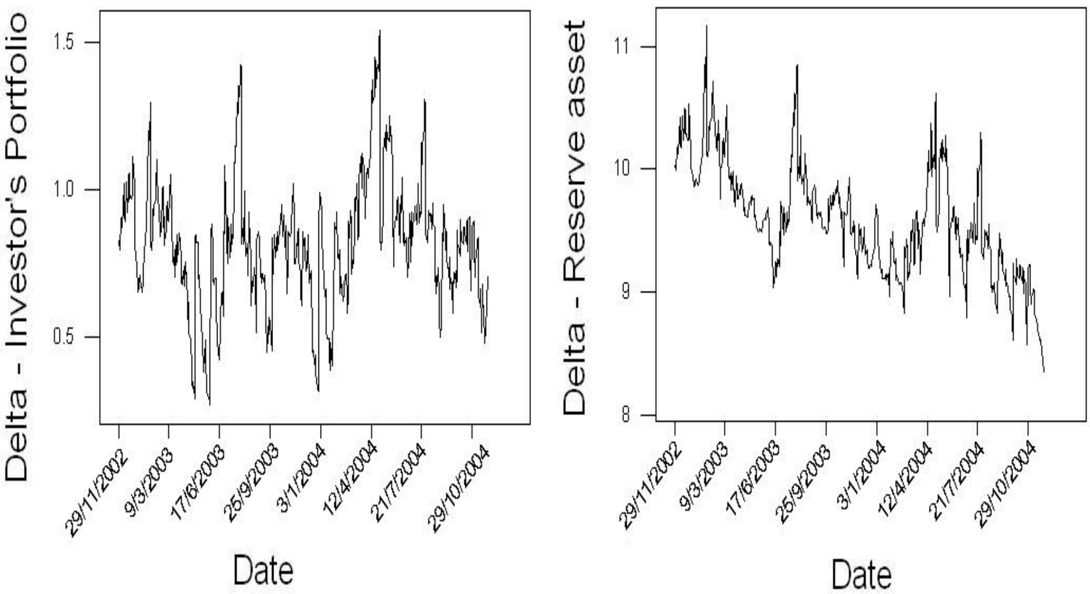

Figure 10. Evolution of controls Leveraged Capital Holdings return replication strategy, Dec. 2002 – Oct. 2004.

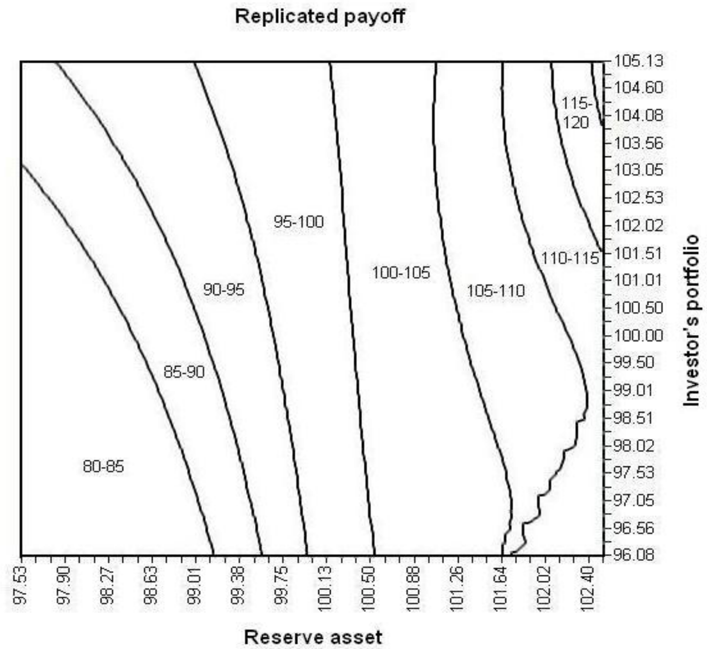

Figure 11. Contour plot payoff function Calamos Multi-Strategy Fund.

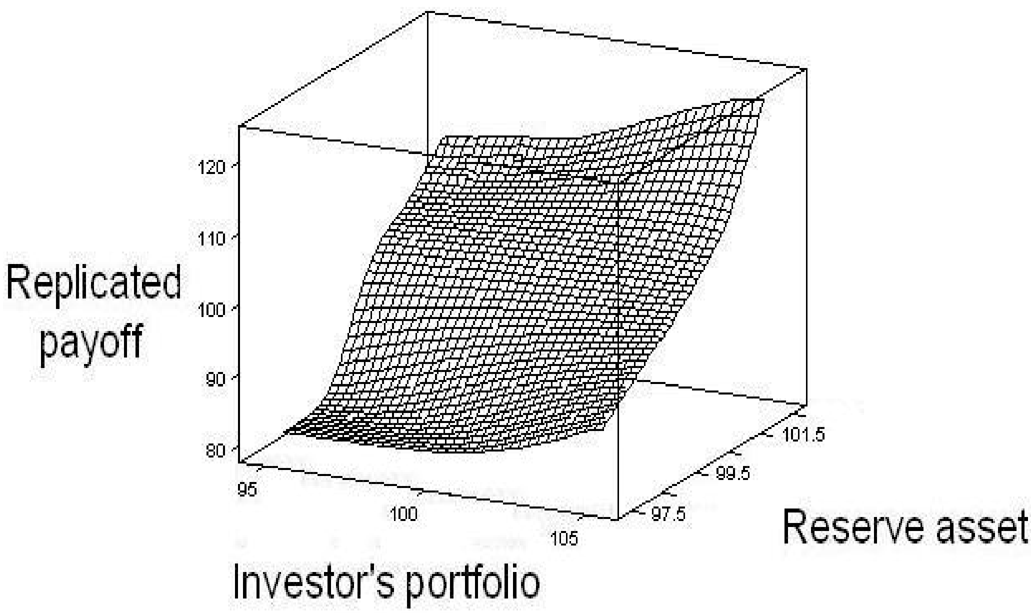

Figure 12. 3D plot payoff function Calamos Multi-Strategy Fund.

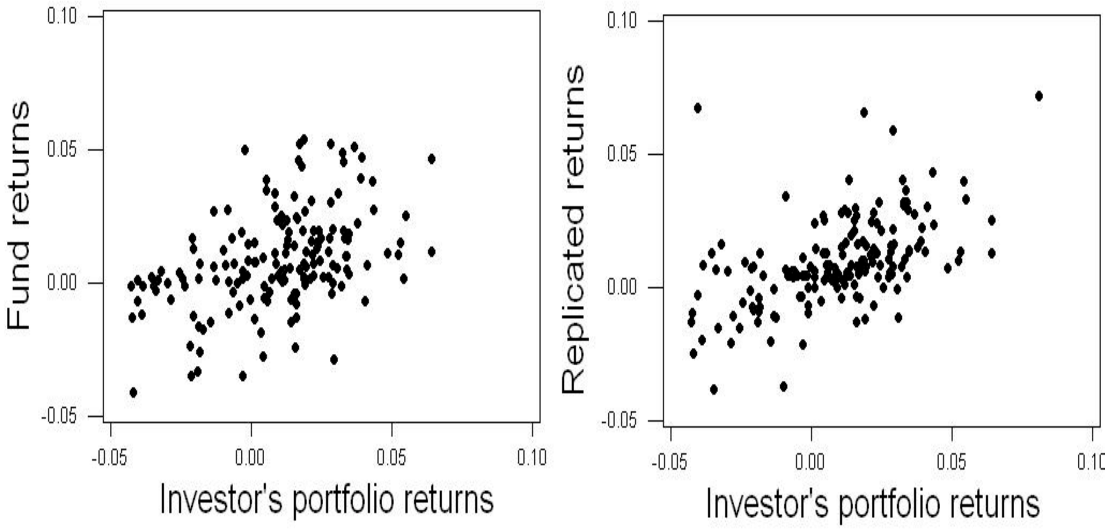

Figure 13. Scatter plot investor's portfolio returns versus Calamos Multi-Strategy Fund returns (left) and replicated returns (right), 1991 - 2004.

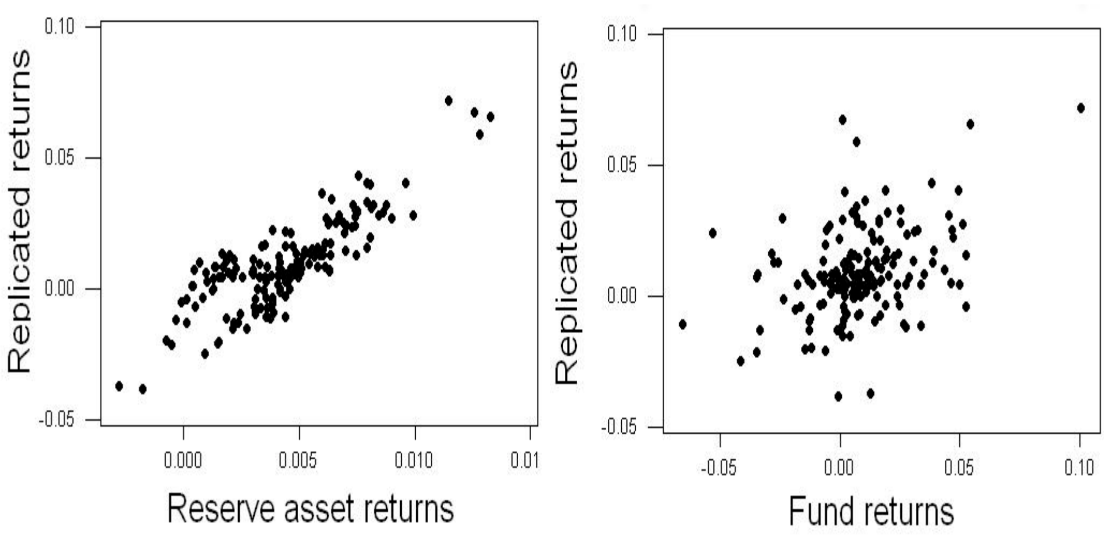

Figure 14. Scatter plot reserve asset returns versus replicated returns (left) and Calamos Multi-Strategy Fund returns versus replicated returns (right), 1991 - 2004.

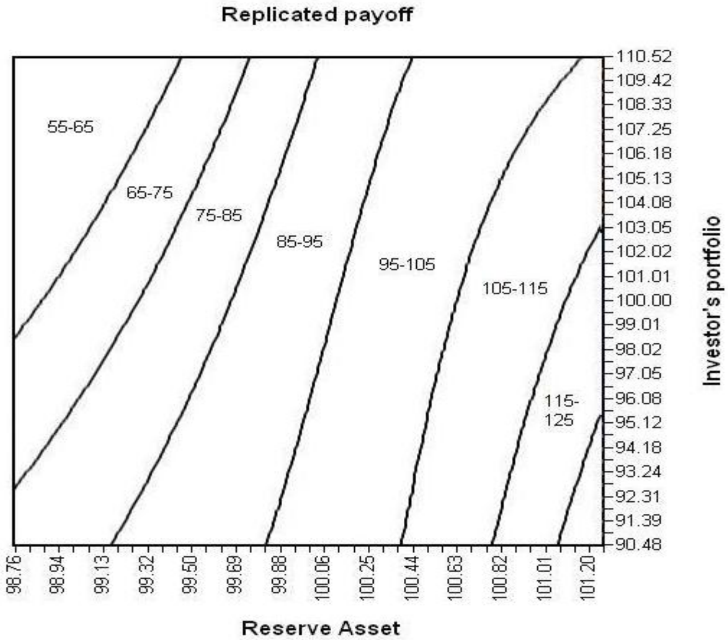

Figure 15. Contour plot payoff function Rocker Partners.

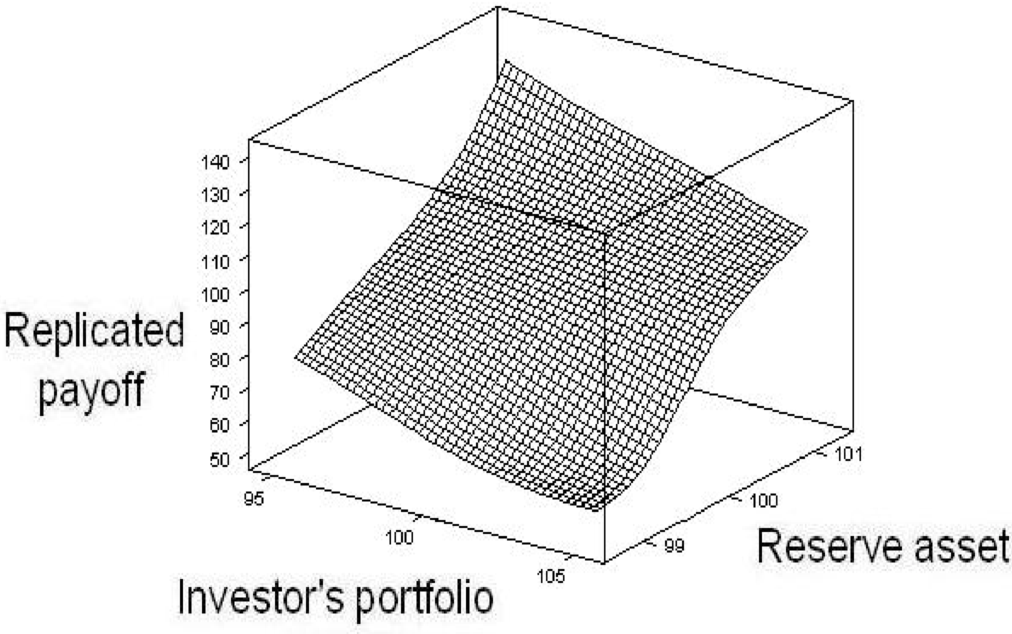

Figure 16. 3D plot payoff function Rocker Partners.

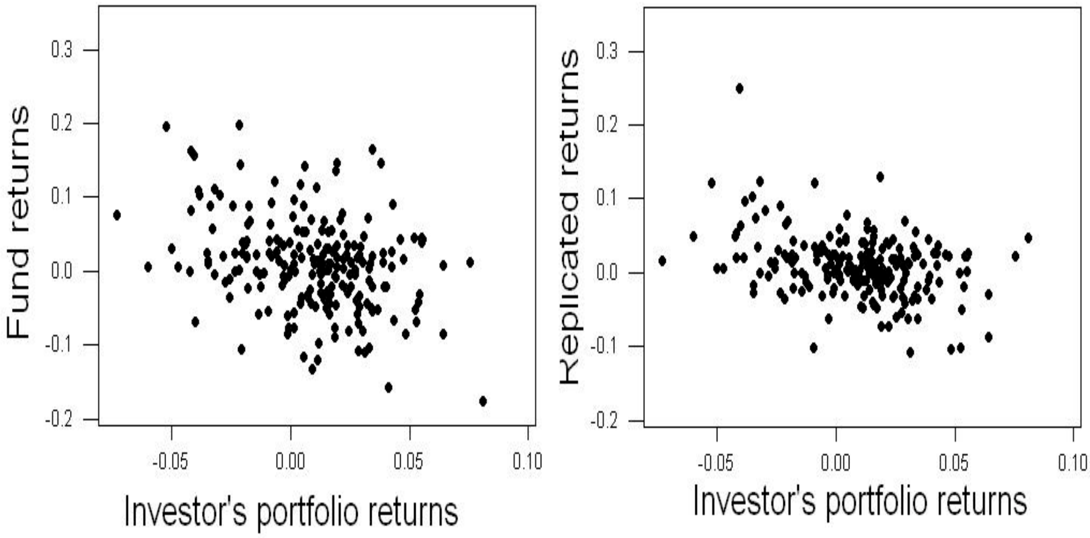

Figure 17. Scatter plot investor's portfolio returns versus Rocker Partners returns (left) and replicated returns (right), 1987 - 2004.

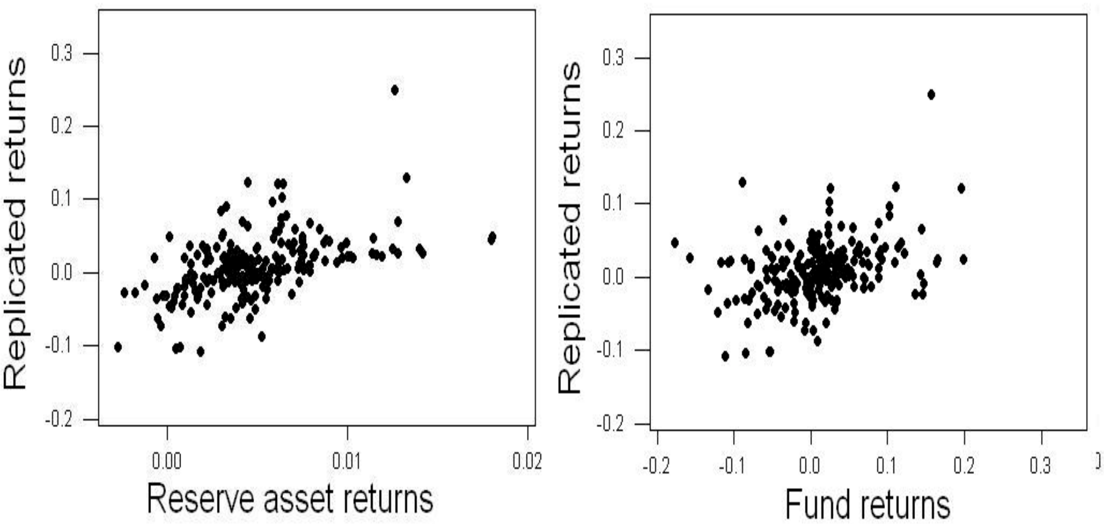

Figure 18. Scatter plot reserve asset returns versus replicated returns (left) and Rocker Partners returns versus replicated returns (right), 1987 – 2004.