Replicating the properties of hedge fund returns *

Nicolas Papageorgiou, Bruno Rémillard, Alexandre Hocquard
HEC Montréal

October 3, 2007

Replicating the properties of hedge fund returns

Abstract

In this paper, we implement a multi-variate extension of Dybvig (1988) Payoff Distribution Model that can be used to replicate not only the marginal distribution of most hedge fund returns but also their dependence with other asset classes. In addition to proposing ways to overcome the hedging and compatibility inconsistencies in Kat and Palaro (2005), we extend the results of Schweizer (1995) and adapt American options pricing techniques to evaluate the model and also derive an optimal dynamic trading (hedging) strategy. The proposed methodology can be used as a benchmark for evaluating fund performance, as well as to replicate hedge funds or generate synthetic funds.

Key Words: Hedge Funds, Hedging, Replication, Copula, Gaussian mixtures.

J.E.L. classification: G10, G20, G28, C16

*Corresponding author: Nicolas Papageorgiou, Finance Department, HEC Montréal, 3000 Cote Sainte-Catherine, Montreal,QC, H3T 2A7, Canada. All the authors are at HEC Montréal and can be reached at firstname.lastname@hec.ca. This work is supported in part by the Fonds pour la formation de chercheurs et l'aide à la recherche du Gouvernement du Québec, by the Fonds québécois de la recherche sur la société et la culture, by the Natural Sciences and Engineering Research Council of Canada and by Desjardins Global Asset Management. The authors would also like to thank Frank Leclerc, Aziz Soré, Gauthier Webanck and Hugues Langlois.

1 Introduction

The impressive growth of the hedge fund industry has naturally led to an increased scrutiny of the fund managers and of their investment strategies. Given the often exorbitant management and performance fees charged by hedge fund managers, it is not surprising that investors are starting to question what they are actually getting for their money. Shrewd investors and institutional fund of funds are becoming increasingly careful about paying alpha fees for beta returns. The challenge that investors and researchers are therefore confronted with is how to reliably separate the funds that are generating alpha returns from the ones that are simply repackaging beta.

The approach that has generally been favored by academics and practitioners in order to extract information about hedge fund returns is the factor model approach. The underlying idea is to try and separate the returns that are due to systematic exposure to risk factors (beta returns) from those that are due to managerial skill (alpha returns). Once the relevant risk factors have been identified, one can evaluate whether the funds exhibit abnormal returns based the intercept of a linear regression of the fund returns against the factor returns. A further advantage of this methodology is that if the linear model is well-specified, one can attempt to replicate the returns of the hedge fund by investing in the appropriate portfolio of factors. A recent paper by Hasanhodzic and Lo (2007) provides some evidence that linear replication can be successful for certain strategies whilst offering certain advantages to hedge fund investing. These include more transparency, increased liquidity and fewer capacity constraints. However the authors warn that the heterogeneous risk profile of hedge funds and the non-linear risk exposures greatly reduce the ability of these models to consistently replicate hedge fund returns. Over the last few months, several banks including Goldman Sachs, JP Morgan and Merril Lynch have launched linear replication funds.

Certain generic hedge fund characteristics help explain some of the difficulty in identifying a well specified linear model. The use of financial derivatives, the use of dynamic leverage, the use of dynamic trading strategies and the asymmetric performance fee structures are some of the most obvious sources of non-linearities in hedge fund returns. Several recent papers, such as Mitchell and Pulvino (2001), Fung and Hsieh (2001), Agarwal and Naik (2004), and Chen and Liang (2006) have dealt with the inclusion of risk premia that attempt to account for these non-linearities. The inclusion of the above option-based factors significantly improves the explanatory power of factor models, however, most of these factors are not tradable and therefore cannot be used to construct a replicating portfolio.

In order to circumvent the issue of identifying tradable risk factors, an interesting alternative approach was proposed by Amin and Kat (2003) and more recently extended by Kat and Palaro (2005). Based on earlier work by Dybvig (1988), the authors evaluate hedge fund performance not by identifying the return generating betas, but rather by attempting to replicate the distribution of the hedge fund returns. The underlying idea is based on the hypothesis that much of the trading activity undertaken by hedge funds is not creating value, just altering the timing of the returns available from traditional assets. In effect, many hedge funds are simply distorting readily available asset distributions. So the real challenge is whether or not we can find a more efficient method to distort these distributions than by investing in hedge funds. Armed with their new efficiency measure, Kat and Palaro (2005) show that hedge fund returns are by no means exceptional and that for the majority of funds an alternative dynamic strategy would have provided investors with superior returns. This methodology not only provides a model free benchmark for evaluating hedge funds, it can also be used to create synthetic funds with predetermined distributional properties.

The efficiency measure as presented by Kat and Palaro (2005) is however subject to several shortcomings and inconsistencies. The most significant of these relates to the way that the daily trading strategies are derived from the distribution of monthly returns. The properties of the estimated monthly distributions and copula functions proposed by the authors are not infinitely divisible and therefore the true properties of the daily returns are not known. As a result, the replicating strategy will not be precise. A further weakness pertains to the fact that although the hedge fund returns and traded assets are clearly non-normal, the efficiency measure is calculated within the confines of the Black-Scholes-Merton world, hence ignoring the higher moments of the distributions.

In this paper, we will implement a multi-variate extension of Dybvig (1988) Payoff Distribution Model that can be used to replicate not only the marginal distribution of most hedge fund returns but also their dependence with other asset classes. In addition to proposing ways to overcome the hedging and compatibility inconsistencies in previous papers, we extend the results of Schweizer (1995) and adapt American options pricing techniques to evaluate the model and also derive an optimal dynamic trading (hedging) strategy. The proposed methodology can be used as a benchmark for evaluating fund performance, as well as to replicate hedge funds or generate synthetic funds.

The rest of the paper will be structured as follows. Section 2 will explain the intuition behind our multi-variate extension of Dybvig's Payoff Distribution model. Section 3 presents the technical details relating to the modeling and estimation of the distributions. Section 4 presents the payoff function. Section 5 presents the replication issues and presents the optimal dynamic trading strategy. Section 6 presents some numerical results. Section 7 concludes.

2 The Multivariate Payoff Distribution Model

In Kat and Palaro (2005), the authors show that given two risky assets $S^{(1)}$ and $S^{(2)}$, it is possible to "reproduce" the statistical properties of the joint return distribution of asset $S^{(1)}$

and a third asset $S^{(3)}$. Let's assume asset $S^{(1)}$ is the investor portfolio, asset $S^{(2)}$ is a tradable security and asset $S^{(3)}$ is a hedge fund, this result implies that we can generate the distribution of the hedge fund and its dependence with the investor portfolio, by only investing in the tradable security $S^{(2)}$ and the investor portfolio $S^{(1)}$. Note that we do not replicate the month by month returns of the hedge fund, but instead we replicate its distributional properties (i.e. expectation, volatility, skewness and kurtosis) as well as dependence measures with respect to the returns of the investor portfolio (i.e. Pearson, Spearman correlations...).

Essentially, there exist a payoff function that will allow us to transform the joint distribution of assets $S^{(1)}$ and $S^{(2)}$ into the bivariate distributions of $S^{(1)}$ and $S^{(3)}$. This payoff function is easily shown to be calculable using the marginal distribution functions F_1, F_2 and F_3 of $S_T^{(1)}, S_T^{(2)}, S_T^{(3)}$, and the copulas $\mathcal{C}_{1,2}$ and $\mathcal{C}_{1,3}$ associated respectively with the joints returns $\left(R_{0,T}^{(1)}, R_{0,T}^{(2)}\right)$ and $\left(R_{0,T}^{(1)}, R_{0,T}^{(3)}\right)$. The exact expression for the payoff function is given in section 4.

The challenge that we are confronted with is how to best evaluate this function, and this is by no means a trivial problem. The problem can however be broken down into three separate components. The first part relates to the proper modeling of the distributions and copula functions. The second part consists in calculating the payoff function. The third part consists in selecting an approach that will allow us to generate a dynamic trading strategy that provides us with the best possible approximation of the payoff function.

3 Modeling the returns

In order to provide a robust solution in this framework, we propose the following steps. First, we will model the joint daily distribution of $S^{(1)}$ and $S^{(2)}$ using bivariate Gaussian mixtures. Since we will be trading these assets on a daily basis, it is imperative that the distribution of the monthly returns for both the investor portfolio and the reserve asset are consistent with the distribution of the daily returns. We need to be sure that by dynamically trading the assets based on the joint daily distributions we will be able to generate the desired monthly properties. We will therefore estimate the parameters of the bivariate Gaussian mixtures of R_t, (investor portfolio and reserve asset) using the historical daily returns of $S^{(1)}$ and $S^{(2)}$. We can then solve for the law of the monthly returns that is compatible with the law of daily returns. Furthermore, the daily dependence which is modeled with the bivariate mixtures will allow us to obtain the desired monthly dependence. This would not have been possible if we used univariate laws to model the marginal distributions and a copula to model the dependence structure. Although copula provide us with much flexibility in terms of modeling the dependence, there is however no proof to this day that the statistical properties of copula functions are divisible. Finally, we need to estimate the monthly distribution of the hedge fund returns as well as the dependence between the hedge fund and the investor portfolio.

There are no particular restrictions regarding the choice of the distribution of $S^{(3)}$ and the copula $\mathcal{C}_{1,3}$. We have developed statistical tests that allow us to select the most appropriate marginal distribution and copula function. We now consider each of these steps in detail.

3.1 Mixtures of Gaussian distributions

The choice of Gaussian mixtures to model the bivariate distribution of investor portfolio and the reserve asset is due to both the flexibility of the mixtures in capturing high levels of skewness as well as the fact that the bivariate distribution is infinitely divisible. In this section, we will first provide a brief description of bivariate Gaussian mixtures and discuss their statistical properties. Finally we will present a goodness-of-fit test that we developed in order to estimate the mixtures and select the optimal number of regimes.

3.1.1 Definition of mixtures of Gaussian bivariate vectors

A bivariate random vector X is a Gaussian mixture with m regimes and parameters $(\pi_k)_{k=1}^m$, $(\mu_k)_{k=1}^m$ and $(A_k)_{k=1}^m$, if its density is given by

$$f(x) = \sum_{k=1}^{m} \pi_k \phi_2(x; \mu_k, A_k)$$

where $\phi_2(x;\mu,A) = \frac{e^{-\frac{1}{2}(x-\mu)^\top A^{-1}(x-\mu)}}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}}$ is the density of a bivariate Gaussian vector with mean vector $\mu = (\mu_1, \mu_2)^\top$ and covariance matrix $A = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. Its distribution function is

$$F(x_1, x_2) = \sum_{k=1}^{m} \pi_k \Phi_2\left(\frac{x_1 - \mu_{k1}}{\sigma_{k1}}, \frac{x_2 - \mu_{k2}}{\sigma_{k2}}; \rho_k\right),$$

where $\Phi_2(\cdot, \cdot; \rho)$ is the bivariate standard Gaussian distribution function with correlation ρ.

3.1.2 Some properties of mixtures of bivariate Gaussian variables

One property that is quite important in our setting is the fact that a sum of independent Gaussian mixtures is still a Gaussian mixture. In fact, if $X_1, \ldots, X_n$ are independent and identically Gaussian mixtures with parameter θ, then $X = X_1 + \cdots + X_n$ is also a Gaussian mixture. To describe the associated parameters, let

$$\mathcal{A} = \{\alpha = (\alpha_1, \ldots, \alpha_m); \alpha_j \geq 0 \text{ and } \alpha_1 + \ldots + \alpha_m = n\}.$$

Then $\text{card}(\mathcal{A}) = \binom{n+m-1}{m-1}$ so there are $\binom{n+m-1}{m-1}$ regimes. The parameters of the mixture are $(\pi_\alpha)_{\alpha\in\mathcal{A}}$, $(\mu_\alpha)_{\alpha\in\mathcal{A}}$, $(A_\alpha)_{\alpha\in\mathcal{A}}$, where for each $\alpha \in \mathcal{A}$, π_α is the multinomial probability

$$\pi_\alpha = \pi_{(\alpha_1,\cdots,\alpha_m)} = \frac{n!}{\alpha_1!\cdots\alpha_m!}\prod_{k=1}^m \pi_k^{\alpha_k},$$

and the mean vectors μ_α and covariances A_α are respectively given by

$$\mu_\alpha = \sum_{k=1}^n \alpha_k \mu_k, \qquad A_\alpha = \sum_{k=1}^n \alpha_k A_k.$$

Remark 3.1 *If n is moderately large, then m^n is huge and it is computationally impossible to calculate the new parameters. In fact, most probabilities could be very small so in fact, the sum could be a mixture of fewer terms. Therefore, one has to estimate again the joint law of $\left(R_{0,T}^{(1)}, R_{0,T}^{(2)}\right)$ by a Gaussian mixture, using the monthly returns this time. As a result, the marginal distributions F_1 and F_2 are (univariate) Gaussian mixtures and $\mathcal{C}_{1,2}$ is the copula deduced from the bivariate Gaussian mixture.*

Finally, consider the conditional distribution of a bivariate Gaussian mixture $X = (X^{(1)}, X^{(2)})$. Set $\beta_k = \rho_k \frac{\sigma_{k2}}{\sigma_{k1}}$ and $\alpha_k = \mu_{k2} - \beta_k \mu_{k1}$, $k = 1, \ldots, m$. Then it is easy to check that the conditional distribution of $X^{(2)}$ given $X^{(1)} = x_1$ is a Gaussian mixture with parameters $\{\tilde\pi_k(x_1)\}_{k=1}^m$, $\{\tilde\mu_k(x_1)\}_{k=1}^m$, $\{\tilde\sigma_k^2\}_{k=1}^m$, where

$$\tilde\pi_k(x_1) = \frac{\pi_k \phi(x_1; \mu_{k1}, \sigma_{k1}^2)}{\sum_{j=1}^m \pi_j \phi(x_1; \mu_{j1}, \sigma_{j1}^2)} \tag{1}$$

and

$$\tilde\mu_k(x_1) = \alpha_k + \beta_k x_1, \qquad \tilde\sigma_k^2 = \sigma_k^2(1-\rho_k^2). \tag{2}$$

3.1.3 Estimation and goodness-of-fit

In order to choose the optimal number of regimes, we need to first estimate the parameters of the model, and then provide a goodness-of-fit test to evaluate whether a greater number of regimes is required. The estimation method is based on the EM algorithm of (Dempster et al., 1977). It is presented in Appendix A.1 for the bivariate case.

A new goodness-of-fit test, described in Appendix A.4, can be performed to assess the suitability as well as to select the number of mixture regimes m. The proposed test, based on the work in Genest et al. (2007), uses the Rosenblatt's transform.

For the selection of the number m of regimes, the following two steps procedure is suggested:

(a) Find the first m_0 for which the P-value of the tests described in A.4 is larger than 5%.

(b) Estimate parameters for $m_0 + 1$ regimes and apply the likelihood ratio test to check if the null hypothesis $H_0 : m = m_0$ vs $H_1 : m = m_0 + 1$. If H_0 is rejected at the 5% level, repeat steps (a) and (b) starting at $m = m_0 + 1$. However, if the parameters under H_1 yield a degenerate density (e.g., $|\rho_k| = 1$), stop and set $m = m_0$.

3.2 Choice/estimation of the marginal distribution F_3

There are no restrictions on the choice of F_3, which is the distribution of the hedge fund that we seek to replicate (or the desired distribution in the case of a synthetic fund). Unlike the reserve asset and investor portfolio that require divisible laws, we are only interested in monthly return distribution and hence can introduce any distribution. In the case of the replication of an existing hedge fund, goodness-of-fit is important and therefore we test using a Durbin type test, as described in Appendix A.3.

3.3 Choice/estimation of the copula $\mathcal{C}_{1,3}$

Again, there are no restrictions on the choice of copula function $\mathcal{C}_{1,3}$, between the monthly returns of the hedge fund and the investor portfolio. Suppose that we have historical monthly returns $(Y_1, Z_1), \ldots, (Y_n, Z_n)$ belong to a copula family C_θ. To estimate θ, one often uses the so-called IFM method. However, we do not recommend it as the parameters of the copula function rely on the estimated marginal distributions. Any mis-specification of the marginal distributions will bias the choice of copula. For reasons of robustness, it is therefore preferable to use normalized ranks, i.e. if R_{i1} represents the rank of Y_i among $Y_1, \ldots, Y_n$ and if R_{i2} represents the rank of Z_i among $Z_1, \ldots, Z_n$, with $R_{ij} = 1$ for the smallest observations, then set

$$U_i = \frac{R_{i1}}{n+1}, \quad V_i = \frac{R_{i2}}{n+1}, \quad i = 1, \ldots, n.$$

To estimate θ one could try to maximize the pseudo-log-likelihood

$$\sum_{i=1} \log c_\theta(U_i, V_i),$$

as suggested in Genest et al. (1995). For example, if the copula is the Gaussian copula with correlation ρ, the pseudo-likelihood estimator for ρ yields the famous van der Waerden coefficient defined to be the correlation between the pairs $\{\Phi^{-1}(U_i), \Phi^{-1}(V_i); i = 1, \ldots, n\}$. For other families that can be indexed by Kentall's tau, e.g., Clayton, Frank and Gumbel families, one could estimate the parameter by inversion of the sample Kendall's tau. See, e.g., Genest et al. (2006).

Finally, to test for goodness-of-fit, one can use Cramér-von Mises type statistics for the empirical copula or for the Rosenblatt's transform. The latter could be the best choice given that $\frac{\partial}{\partial u}\mathcal{C}_{1,3}(u,v)$ needed to be calculated for the evaluation of the payoff function. These tests are described in Genest et al. (2007) and in view of their results, we recommend to use the test statistic $S_n^{(B)}$.

4 The payoff function

Having estimated the necessary distributions and copula function, one must now calculate the payoff's return function g. As deduced by Kat and Palaro (2005), its formula is given by

$$g(x,y) = Q\left\{x, P\left(R_{0,T}^{(2)} \le y | R_{0,T}^{(1)} = x\right)\right\},$$

where $Q(x,\alpha)$ is the order α quantile of the conditional law of $R_{0,T}^{(3)}$ given $R_{0,T}^{(1)} = x$, i.e., for any $\alpha \in (0,1)$, $q(x,\alpha)$ satisfies

$$P\left\{R_{0,T}^{(3)} \le Q(x,\alpha) | R_{0,T}^{(1)} = x\right\} = \alpha.$$

Using properties of copulas, e.g. Nelsen (1999), the conditional distributions can be expressed in terms of the margins and the associated copulas.

$$P\left(R_{0,T}^{(2)} \le y | R_{0,T}^{(1)} = x\right) = \left.\frac{\partial}{\partial u}\mathcal{C}_{1,2}(u,v)\right|_{u=F_1(x), v=F_2(y)}.$$

Note that $\dfrac{\partial}{\partial u}\mathcal{C}_{1,2}(u,v) = P\left\{F_2(R_{0,T}^{(2)}) \le v | F_1(R_{0,T}^{(1)}) = u\right\}$. In addition, if $\mathcal{Q}(u,\alpha)$ is the order α quantile of the distribution function $\dfrac{\partial}{\partial u}\mathcal{C}_{1,3}(u,v)$, then one obtains

$$Q(x,\alpha) = F_2^{-1} \circ \mathcal{Q}(F_1(x), \alpha).$$

In our methodology, since the monthly returns $\left(R_{0,T}^{(1)}, R_{0,T}^{(2)}\right)$ are modeled by a Gaussian mixtures with parameters $(\pi_k)_{k=1}^m$, $(\mu_k)_{k=1}^m$ and $(A_k)_{k=1}^m$, the conditional distributions can be expressed as follows

$$P(R_{0,T}^{(2)} \le y | R_{0,T}^{(1)} = x) = \sum_{k=1}^m \tilde{\pi}_k(x)\phi\{y; \tilde{\mu}_k(x), \tilde{\sigma}^2\}$$

where $\tilde{\pi}_k(x)$, $\tilde{\mu}_k(x)$ and $\tilde{\sigma}^2$ are given by formulas (1) and (2).

5 Dynamic replication

Having solved for g, we need to find an optimal dynamic trading strategy that will replicate the payoff function. We do so by selecting the portfolio (V_0, φ) such as to minimize the expected square hedging error

$$E\left[\beta_T^2 \left\{V_T(V_0,\varphi) - C_T\right\}^2\right],$$

where β_T is the discount factor.

In order to achieve this, we develop extensions of the results of Schweizer (1995). Note that there is no "risk-neutral" evaluation involved in our approach and that all calculations are carried out under the objective probability measure.

5.1 Optimal hedging

Suppose that $(\Omega, P, \mathcal{F})$ is a probability space with filtration $\mathbb{F} = \{\mathcal{F}_0, \ldots, \mathcal{F}_T\}$, under which the stochastic processes are defined. For the moment, assume that the price process S_t is d-dimensional, i.e. $S_t = \left(S_t^{(1)}, \ldots, S_t^{(d)}\right)$. In the next section, one will come back with the case $d = 2$.

Before defining what is meant by a dynamic replicating strategy, let β_t denote the discount factor, i.e. β_t is the value at period 0 to be invested in the non risky asset so that it has a value of 1\$ at period t. By definition, $\beta_0 = 1$. It is assumed that the process β is predictable, i.e. β_t is $\mathcal{F}_{t-1}$-measurable for all $t = 1, \ldots, T$.

A dynamic replicating strategy can be described by a (deterministic) initial value V_0 and a sequence of random weight vectors $\varphi = (\varphi_t)_{t=0}^T$, where for any $j = 1, \ldots, d$, $\varphi_t^{(j)}$ denotes the number of parts of assets $S^{(j)}$ invested during period $(t-1, t]$. Because φ_t may depend only on the values values $S_0, \ldots, S_{t-1}$, the stochastic process φ_t is assumed to be predictable. Initially, $\varphi_0 = \varphi_1$, and the portfolio initial value is V_0. It follows that the amount initially invested in the non risky asset is

$$V_0 - \sum_{j=1}^d \varphi_1^{(j)} S_0^{(j)} = V_0 - \varphi_1^\top S_0.$$

Since the hedging strategy must be self-financing, it follows that for all $t = 1, \ldots, T$,

$$\beta_t V_t(V_0,\varphi) - \beta_{t-1} V_{t-1}(V_0,\varphi) = \varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1}). \tag{3}$$

Using the self-financing condition (3), it follows that

$$\beta_T V_T = \beta_T V_T(V_0, \varphi) = V_0 + \sum_{t=1}^{T} \varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1}). \tag{4}$$

The replication strategy problem for a given payoff C is thus equivalent to finding the strategy (V_0, φ) so that the hedging error

$$G_T(V_0, \varphi) = \beta_T V_T(V_0, \varphi) - \beta_T C \tag{5}$$

is as small as possible. In this paper, we choose the expected square hedging error as a measure of quality of replication. It is therefore natural to suppose that the prices $S_t^{(j)}$ have finite second moments. We further assume that the hedging strategy φ satisfies a similar property, namely that for any $t = 1, \ldots, T$, $\varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1})$ have finite second moments. Note that these two technical conditions were also made by Schweizer (1995).

For simplicity, set

$$\Delta_t = S_t - E(S_t|\mathcal{F}_{t-1}), \qquad t = 1, \ldots, T.$$

Under the above moment conditions, the conditional covariance matrix Σ_t of Δ_t exists and is given by

$$\Sigma_t = E\left\{\Delta_t \Delta_t^\top | \mathcal{F}_{t-1}\right\}, \; 1 \le t \le T.$$

In Schweizer (1995), the author treats the case $d = 1$ and assumes a restrictive boundedness condition. Here, in contrast, we treat the general d-dimensional case and we only suppose that Σ_t is invertible for all $t = 1, \ldots, T$. This was implicitly part of the boundedness condition of Schweizer (1995).

If Σ_t is not invertible for some t, there would exists a $\varphi_t \in \mathcal{F}_{t-1}$ such that $\varphi_t^\top S_t = \varphi_t^\top E(S_t|\mathcal{F}_{t-1})$, that is, $\varphi_t^\top S_t$ is predictable. Our assumption can be interpreted as saying that the genuine dimension of the assets is d. One may now state the main result whose proof is given in Appendix D.1.

Theorem 1 *Suppose that Σ_t is invertible for all $t = 1, \ldots, T$. Then the risk $E\{G^2(V_0, \varphi)\}$ is minimized by choosing recursively $\varphi_T, \ldots, \varphi_1$ satisfying*

$$\varphi_t = (\Sigma_t)^{-1} E\left(\{S_t - E(S_t|\mathcal{F}_{t-1})\} C_t | \mathcal{F}_{t-1}\right), \; t = T, \ldots, 1, \tag{6}$$

where $C_T, \ldots, C_0$ are defined recursively by setting $C_T = C$ and

$$\beta_{t-1} C_{t-1} = \beta_t E(C_t|\mathcal{F}_{t-1}) - \varphi_t{}^\top E(\beta_t S_t - \beta_{t-1} S_{t-1}|\mathcal{F}_{t-1}), \tag{7}$$

for $t = T, \ldots, 1$.

Moreover the optimal value of V_0 *is* C_0*, and*

$$E(G^2) = \sum_{t=1}^{T} E\left(\beta_t^2 G_t^2\right),$$

where $G_t = \varphi_t{}^\top \{S_t - E(S_t|\mathcal{F}_{t-1})\} - \{C_t - E(C_t|\mathcal{F}_{t-1})\}$*,* $1 \le t \le T$*.*

Having found the optimal hedging strategy, according to the mean square error criterion, one might ask what the link is between the price given by C_0, as in Theorem 1, and the price suggested by the martingale measure method. The answer is given by the following result proven in Appendix D.2.

Corollary 1 *For any* $t = 1 \ldots, T$*, set*

$$U_t = 1 - \Delta_t^\top (\Sigma_t)^{-1} E\left(S_t - \beta_{t-1} S_{t-1}/\beta_t | \mathcal{F}_{t-1}\right). \tag{8}$$

Further set $M_0 = 1$ *and* $M_t = U_t M_{t-1}$*,* $1 \le k \le n$*. Then* $(M_t, \mathcal{F}_t)_{t=0}^T$ *is a (not necessarily positive) martingale and*

$$\beta_{t-1} C_{t-1} = E(\beta_t C_t U_t | \mathcal{F}_{t-1}).$$

In particular $\beta C_t M_t$ *is a martingale and* $C_0 = E(\beta_T C_T M_T | \mathcal{F}_0)$*. Moreover* $E(\beta_t S_t U_t | \mathcal{F}_{t-1}) = \beta_{t-1} S_{t-1}$*, so* $\beta_t S_t M_t$ *is a martingale.*[1]

5.1.1 The Markovian case

If the price process S is Markovian, i.e., the law of S_t given $\mathcal{F}_{t-1}$ is $\nu_t(S_{t-1}, dx)$, and if the terminal payoff $C_T = C$ only depends on the terminal prices, that is $C = f_T(S_T)$, then the

[1] When the market is complete, there is a unique martingale measure Q and every claim is attainable, so the risk associated with the optimal strategy is zero. Therefore M_t, as defined in Corollary 1 is positive, and as a by-product of our method, we have an explicit representation of the density of Q with respect to P.

Markov property, together with Theorem 1, yield that $C_t = f_t(S_t)$ and $\varphi_t = \psi_t(S_{t-1})$, where

$$\begin{aligned}
L_{1t}(s) &= E(S_t|S_{t-1}=s) = \int x\nu_t(s,dx), \\
L_{2t}(s) &= E(S_tS_t^\top|S_{t-1}=s) = \int xx^\top\nu_t(s,dx), \\
A_t(s) &= L_{2t}(s) - L_{1t}(s)L_{1t}(s)^\top, \\
\psi_t(s) &= A_t(s)^{-1}E\left[\{S_t - L_{1t}(s)\}f_t(S_t)|S_{t-1}=s\right] \\
&= A_t(s)^{-1}\int (x - L_{1t}(s))f_t(x)\nu_t(s,dx), \\
U_t(s,x) &= 1 - (L_{1t}(s) - \beta_{t-1}s/\beta_t)^\top A_t(s)^{-1}(x - L_{1t}(s)), \\
f_{t-1}(s) &= \frac{\beta_t}{\beta_{t-1}}E\{U_t(s,S_t)f_t(S_t)|S_{t-1}=s\} \\
&= \frac{\beta_t}{\beta_{t-1}}\int U_t(s,x)f_t(x)\nu_t(s,dx).
\end{aligned}$$

Note that $E(S_t|\mathcal{F}_{t-1}) = L_{1t}(S_{t-1})$ and $\Sigma_t = A_t(S_{t-1})$. Explicit calculations can be done when the returns are assumed to be a finite Markov chain. In most models, one can write $S_t = \omega_t(S_{t-1},\xi_t)$ where ξ_t is independent of $\mathcal{F}_{t-1}$ and has law P_t. When μ_t has an infinite support, there are ways to approximate ψ_t and f_t.

The importance of Theorem 1 to the replication problem of hedge funds is obvious, particularly under the Markovian setting. All that is needed is a way to calculate or approximate the value of f_0 and of the deterministic functions $\psi_t(s), f_t(s)$, $t = 1,\ldots$. In particular $V_0 = f_0$ and $\varphi_t = \psi_t(s)$ gives the optimal hedging strategy when $S_{t-1} = s$.

5.1.2 The dynamic trading strategy

In the Markovian case, one can use the methodology developed by Del Moral et al. (2006) to calculate both the φ_t's and the C_t's. The algorithm for implementing the dynamic trading strategy, based on Monte Carlo simulations and linear interpolation, is described in more details in Appendix B.

5.2 A comparison between optimal hedging and hedging under Black-Scholes setting

To compare the two methods, simply take $T = 1$ and $r = 0$ and $d = 1$. In this case, the solution for optimal hedging yields $\varphi^\star = \mathrm{Cov}\{\Delta S_1, C(S_1)\}/\mathrm{Var}(\Delta S_1)$, where $\Delta S_1 = S_1 - S_0$, and $V_0^\star = E\{C(S_1)\} - \varphi^\star E(\Delta S_1)$.

For the Black-Scholes setting, we have

$$V_0^{BS} = E\left\{C\left(S_0 e^{\sigma Z - \sigma^2/2}\right)\right\} \quad \text{and} \quad \varphi^{BS} = E\left\{e^{\sigma Z - \sigma^2/2} C'\left(S_0 e^{\sigma Z - \sigma^2/2}\right)\right\},$$

with $\sigma^2 = \text{Var}\{\log(S_1/S_0)\}$, where $Z \sim N(0,1)$, provided C is differentiable. See, e.g., Broadie and Glasserman (1996). In general, $\varphi^\star \neq \varphi^{BS}$ and $V_0^\star \neq V_0^{BS}$, so

$$E\left[\{V_1(V_0^\star, \varphi^\star) - C(S_1)\}^2\right] < E\left[\left\{V_1(V_0^{BS}, \varphi^{BS}) - C(S_1)\right\}^2\right].$$

For an analysis of the (discrete) hedging error in a Black-Scholes setting, see, e.g., Wilmott (2006). To illustrate the difference in an hedge funds context, we performed a numerical experiment in which we tried (10 000 times) to reproduce a synthetic fund with centered Gaussian distribution with volatility 12%, independent of the portfolio. The distribution of the daily returns of the (portfolio, reserve) pair are modeled by a a mixture of 4 regimes for the daily returns distribution with parameters given in Table 1. With this choice of parameters, it turns out that the associated monthly returns are best modeled by a bivariate Gaussian with parameters are given in Table 2.

Table 1: Parameters for the Gaussian mixture with 4 regimes used for modeling daily returns

π_k	μ_{k1}	μ_{k2}	σ_{1k}	σ_{2k}	ρ_k
0.0956	0.0016	0.0008	0.0039	0.0016	0.9754
0.4673	0.0000	0.0002	0.0069	0.0032	0.7981
0.0763	-0.0003	-0.0005	0.0115	0.0054	0.6964
0.3607	0.0006	0.0005	0.0037	0.0027	0.4613

Table 2: Estimation of the parameters of the Gaussian model compatible with the daily returns

μ_1	μ_2	σ_1	σ_2	ρ
0.007892797	0.0068086	0.029334999	0.034641016	0.700295314

As said previously, we simulated 10 000 values of $g\left(R_{0,T}^{(1)}, R_{0,T}^{(2)}\right)$, $\log(V_T^\star/100)$ (under optimal hedging) and $\log(V_T^{BS}/100)$ (under delta hedging). Some sample characteristics of these three variables are given in Table 3, together with the corresponding true values, while for each dynamic trading method, the estimated mean hedging error and square root mean square error are given in Table 4.

By construction, optimal hedging always produces an hedging error with zero mean. However, this is not the case in general for delta hedging. Note how far the delta hedging method is off the goal of a zero mean of the replicating portfolio, while the optimal hedging error is much smaller.

As our proposed method is optimal for minimizing the square hedging error, it is not surprising that it dominates delta hedging. However, since the theoretical setting is very close to the Black-Scholes setting, all monthly returns being Gaussian, it is worth noting that the square root Mean Square Error of the optimal hedging is 150% less than the one of the delta hedging.

Finally, the distribution of the respective hedging errors is illustrated in Figure 1. From that graph, it appears that the values of the replication portfolio with the methodology proposed in Kat and Palaro (2005) are almost always smaller than the target values.

Table 3: Replication results based on 10 000 trajectories for $g\left(R_{0,T}^{(1)}, R_{0,T}^{(2)}\right) = \log(C_T/100)$ and $\log(V_T/100)$ under optimal hedging and delta hedging.

Parameter	True value	g	Optimal hedging	Delta hedging
Mean	0	3.957E-07	3.574E-07	-0.000422735
Std. dev.	0.034641016	0.034957842	0.034961135	0.034985553
Skewness	0	-0.058910418	-0.064053039	-0.063978046
Kurtosis	0	0.029916203	0.032479236	0.032374552
ρ	0.3	0.30283895	0.30279462	0.30288552

Table 4: Replication results based on 10 000 trajectories for the payoff $\tilde{g}$ and $\log(V_T/100)$ under optimal hedging and delta hedging.

Parameter	Optimal hedging	Delta hedging	$\|OH/BS\|$
Mean hedging error	0.000004009	-0.042061101	10491.66889
Square root MSE	0.017861376	0.045665732	2.556674977

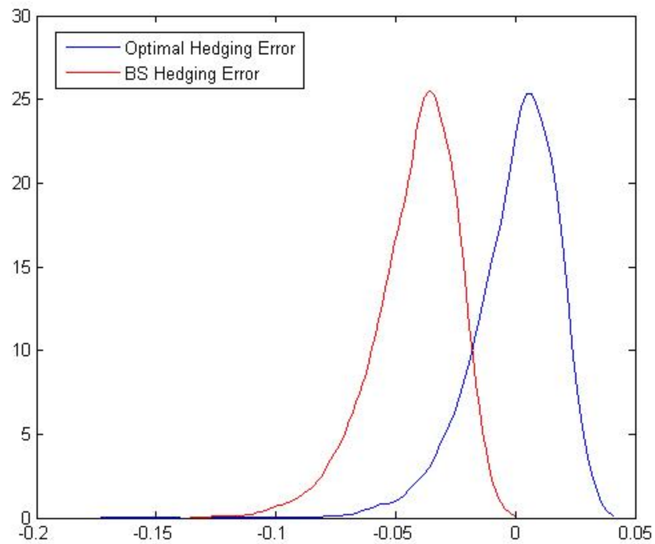

Figure 1: Kernel density estimation of hedging errors for optimal hedging and delta hedging.

6 Replication of hedge funds

In this section we will provide some empirical evidence regarding the ability of the model to replicate hedge fund returns. For the sake of parsimony, we will present results for the (in-sample) replication of the EDHEC indices and HFR indices. We will look at the models ability to replicate the statistical properties of the monthly returns of the different indices over the ten year period from 01/30/1997 to 12/29/2006 (120 months), as well as for 2 subperiods ranging respectively from 01/30/1997 to 12/29/2001 (59 months) and from 12/30/2001 - 12/29/2006 (61 months).

6.1 Portfolio and Reserve assets

The first step is top select the assets that will make up the investor portfolio, $S^{(1)}$, and the reserve asset, $S^{(2)}$. Because these two portfolios are dynamically traded on a daily basis, we seek very liquid instruments with low transaction costs. We therefore restrict the components of these two assets to be either Futures contracts or Exchange Traded Funds (ETF).

All futures data comes from CRB Trader database. The cash rate is the BBA Libor 1 month rate. Log-returns on futures are calculated from the reinvestment of a rolling strategy in the front contract. The front contract is the nearest to maturity, on the March/June/-

September/December schedule and is rolled on the first business day of the maturity month at previous close prices. Each future contract is fully collateralized, so that, the total return is the sum the rolling strategy's return and the cash rate. The ETF data is obtained from Bloomberg.

The investor portfolio, which is meant to be a proxy for a typical institutional portfolio, will be an equal-weighted portfolio of S&P500 futures contracts and 30 year US Treasury Bond futures contracts. In order to illustrate the sensitivity of the methodology to the choice of reserve asset, we will perform the study using two very different reserve assets. The first asset (Reserve 1) is made up of 50% PowerShares Dynamic Small Cap Value Portfolio, 25% iShares Lehman 20 Year Treasury Bond Fund and 25% Citigroup Treasury 10 Year Bond Fund. The second asset (Reserve 2) is an equally weighted portfolio Two Year Treasury Notes, Ten Year Treasury Notes, S&P500, and Goldman Sachs Commodity Index future contracts.

Table 5 presents some of the statistical properties of our investor portfolio and the two reserve assets for the entire ten year period and the two sub-periods. We report the mean, standard deviation, skewness, robust skewness[2], kurtosis, robust kurtosis[3]

As explained in Section 3.1, we have chosen to model the daily returns of the pairs (portfolio, reserve) by bivariate Gaussian mixtures with m regimes, denoted by BGM(m).

In Table 6, the distributions of the daily and monthly returns for the (portfolio, reserve) pairs are given, over the three time periods. These results were obtained by using the estimation and goodness-of-it procedures described in Section 3.1.3.

It may seems odd at first that the model for the joint monthly returns is a (bivariate) Gaussian mixture with fewer regimes than for the daily returns. However, as explained in Remark 3.1, it is quite normal. In fact, in view of the central limit theorem, the number of regimes would possibly be 1 if we were to consider returns over a two months period.

6.2 Hedge fund indices

For the sake of comparison, we chose to replicate the 13 EDHEC indices and the 22 HFRI indices. According to the procedures described in Sections 3.2 and 3.3, the marginal distribution F_3 and the copula $C_{1,3}$ were estimated for each hedge fund index.

For the marginal distributions, we considered (univariate) Gaussian mixtures with m regimes, denoted GM(m) and Johnson distribution. For the copula families, we selected the Gaussian, Student, Clayton, Frank and Gumbel. In each case, we estimated Kendall's tau, which measures the dependence between the hedge fund returns and the portfolio returns.

[2] Defined by $\{E(X) - Q(1/2)\} \Big/ E\{|X - Q(1/2)|\}$, where Q_α is the α-quantile.

[3] Defined by $0.09 + \{Q(.975) - Q(.025)\} \Big/ \{Q(.75) - Q(.25)\}$.

Table 5: Summary statistics for the portfolio and the reserve assets over the three time periods.

Asset	Statistics	Period 1 (97–06)	Period 2 (97–01)	Period 3 (02–06)
Portfolio	Mean	0.0035	0.0047	0.0024
	S.Dev	0.0244	0.0289	0.0192
	Skew	-0.2150	-0.2697	-0.2482
	R. Sk	-0.0813	-0.2665	-0.1097
	Kurt	3.2109	2.6942	3.6637
	R. Kurt	3.2467	2.7483	3.6386
Reserve 1	Mean	0.0094	0.0095	0.0093
	S.Dev	0.0225	0.0260	0.0187
	Skew	0.3006	0.5346	-0.3480
	R. Sk	0.0362	0.0552	0.0159
	Kurt	5.0025	5.0399	3.2161
	R. Kurt	3.2419	4.0561	2.9244
	Corr. with Port.	0.6749	0.7054	0.6206
Reserve 2	Mean	0.0031	0.0016	0.0047
	S.Dev	0.0195	0.0219	0.0168
	Skew	0.0338	0.3193	-0.3886
	R. Sk	-0.0891	-0.0161	-0.2345
	Kurt	3.4509	3.3083	3.7213
	R. Kurt	3.3207	3.3894	3.4959
	Corr. with Port.	0.6040	0.7231	0.3989

Table 6: Distribution of the daily and monthly returns for the two pairs (portfolio, reserve), over the three time periods.

Returns	Period 1 (97–06)		Period 2 (97–01)		Period 3 (02–06)	
	Reserve 1	Reserve 2	Reserve 1	Reserve 2	Reserve 1	Reserve 2
Daily	BGM(5)	BGM(5)	BGM(5)	BGM(5)	BGM(3)	BGM(4)
Monthly	BGM(2)	BGM(2)	BGM(4)	BGM(2)	BGM(2)	BGM(3)

Except for the Student copula, which is dependent on two parameters, the other families only depend on one parameter.

The best fitting models are displayed in Tables 8–10.

6.3 Performance of the replication

There are two important issues that need to be addressed when analyzing the models ability to replicate hedge fund returns. The first issue concerns the models ability to effectively replicate hedge fund indices. The second issue pertains to the choice of the reserve asset and it's impact on the models performance.

To study the effectiveness of the replication strategies, there are two main factors to consider: the initial investment V_0 that is required to replicate each index as well as the actual quality of the replication. In order to obtain the payoff distribution of the hedge fund indices, we follow the approach used by Kat and Palaro (2005)- we calculate the monthly returns assuming an investment of 100 at the beginning of each month. Therefore, if the value V_0 of the replicating strategy is below 100, this would lead us to conclude that the replicating strategy offers a cheaper alternative to the hedge fund index, and therefore is the better investment choice. This analysis can however be misleading if we do not also examine the precision of the replication strategy. Before dismissing the hedge fund indices as poor-performers, we need to properly evaluate whether the properties of the replication strategies and hedge fund indices are truly the same. A proper examination of both the cost and the precision of the replication strategy is fundamental before any strong conclusion can be drawn about the model's ability to replicate hedge fund indices.

Then arises the question of the reserve asset. Does the reserve asset impact the performance of the model, and if so does it affect only V_0 or also the ability of the model to replicate the statistical properties of the hedge fund indices? In other words, does the choice of reserve asset impact the performance measure and/or for the quality of the replication?

Tables 11–13 present the values of V_0 for the HFRI and EDHEC hedge fund indices. It is quite clear that even without correcting for the well documented biases in hedge fund indices, the replicating strategies still out-perform a large number of the hedge fund indices over the entire period as well as over the two sub-periods. In order to show that the replication strategies are effectively reproducing the statistical properties of the hedge fund indices, Figures 2–6 present the target mean, volatility, Kendall's tau, skewness and kurtosis of the indices as well as those for the replication strategies. It is quite clear that independently of the period that is considered, the volatility and Kendall's tau are reproduced with great precision. It is important to note that the only moment that is sensitive to the choice of reserve asset is the return of the replication strategy - the other moments as well as the dependence coefficient appear to be insensitive to the choice of reserve asset. Our results clearly indicate that the reserve asset plays a role in the measure of performance, V_0, but it has almost no effect on the quality of the replication.

Table 7: Regression of EDHEC and HFRI indices returns with the replication returns (for reserve assets 1–2) for the following target parameters: volatility, skewness, robust skewness, kurtosis, robust kurtosis, Kendall's tau and Pearson's rho.

Period: (1997–2006)	Reserve 1			Reserve 2		
Target	Intercept	Slope	$R^2(\%)$	Intercept	Slope	$R^2(\%)$
Volatility	0.000624738	0.997485421	99.38	0.000132117	1.034882095	99.38
Skewness	-1.21833672	1.135438624	63.48	-0.660897414	1.017065756	78.82
Robust Skew	0.005285212	0.591422785	38.74	0.049971694	0.845539485	68.79
Kurtosis	1.427089662	1.320048662	26.05	1.738641971	1.116543294	79.34
Robust Kurt	2.057766094	0.48321167	36.19	1.800169291	0.491547026	34.31
Kendall's Tau	0.040820382	1.009779392	98.80	0.034979443	1.024409652	99.36
Pearson's Rho	0.031939885	1.046644073	95.80	0.030103056	1.064383569	96.32

Period: (1997–2001)	Reserve 1			Reserve 2		
Target	Intercept	Slope	$R^2(\%)$	Intercept	Slope	$R^2(\%)$
Volatility	0.000246245	0.999597258	98.15	0.000303651	1.026011825	98.27
Skewness	-0.58025733	0.917232282	32.10	-0.86585092	1.542889847	65.17
Robust Skew	0.044192227	0.916761936	56.14	-0.00965482	0.729453739	54.08
Kurtosis	5.581125649	0.675401646	15.27	2.081736175	1.733323643	20.62
Robust Kurt	-0.85346256	1.451214328	67.35	-0.58104505	1.267471264	63.20
Kendall's Tau	0.0254162	1.019450292	98.52	0.020171502	1.016297547	99.18
Pearson's Rho	0.056163582	1.022429795	91.53	0.021180004	1.06516375	94.76

Period: (2002–2006)	Reserve 1			Reserve 2		
Target	Intercept	Slope	$R^2(\%)$	Intercept	Slope	$R^2(\%)$
Volatility	-0.00015482	0.987878677	99.84	0.000145601	0.984626992	99.59
Skewness	-0.07264573	1.035201227	83.03	0.045240922	1.104989802	80.44
Robust Skew	0.004977508	0.816341527	53.79	0.068954198	1.033664472	61.45
Kurtosis	1.26397635	0.768109088	35.44	0.536774019	0.93664374	67.14
Robust Kurt	1.471815049	0.540835024	45.90	1.421018259	0.566191474	26.63
Kendall's Tau	-0.00043248	1.069855948	98.96	0.019777659	1.034958883	98.97
Pearson's Rho	0.059853575	1.027479829	91.93	0.119157269	1.054127939	91.77

In order to further examine the model's ability to replicate the statistical properties of the hedge fund indices, Table 7 presents the results obtained by regressing the statistical properties of the replication portfolios against the estimated parameters of both EDHEC and HFRI indices for the three samples periods. If the replications were perfect, the slope would then be 1 and the intercept would be 0. As one can see, the fit is very impressive for both reserve assets. The volatility and dependence measures (Kendal's tau and Spearman's Rho) are perfectly replicated, and the regression coefficients for the higher moments, although not perfect, support the model's ability to replicate the statistical properties of hedge fund returns.

The final stage of the analysis consists of breaking down the costs and other potential sources of error associated with the dynamic replicating strategy. We quantify three potential costs/errors associated with our methodology. The first is the transaction costs related to the dynamic trading; the second is the rounding error that results from not being able to trade fractions of futures contracts; the third, and most significant, is the profit/loss that is due to the hedging error of the discrete hedging strategy.

The transaction costs are assumed to be 1 basis point for the sale/purchase of all futures contracts. Obviously, the amount of trading required to replicate the different indices can vary substantially. In table 14 we present the average monthly transaction costs (in terms of basis points) incurred for each replicating portfolio over the whole sample period. Note that the average monthly transaction costs for the replication strategies is approximately 5 basis points.

The rounding error that results from the inability to buy or sell fractions of futures contracts depends very much on the size of the replication portfolio and this error tends to zero as the portfolio increases in size. For a replicating strategy with $100 Million invested, the average monthly rounding error is approximately 1 basis point.

Finally, we calculate replicating errors, that is the average difference between the value of the replicating strategy and the value of the hedge fund index. The results are presented in Tables 15–17. Note that the average monthly hedging error on all replications as defined in Equation 5 is around 3 basis points.

7 Conclusion

In this paper, we implement a multi-variate extension of Dybvig (1988) Payoff Distribution Model that can be used to replicate not only the marginal distribution of hedge fund returns but also their dependence with other asset classes. In addition to proposing ways to overcome the hedging and compatibility inconsistencies in Kat and Palaro (2005) we extend the results of Schweizer (1995) and adapt American options pricing techniques to evaluate the model and also derive an optimal dynamic trading (hedging) strategy. In section 5.2 we demonstrate the superiority of the hedging algorithm that is used to generate the dynamic replicating strategy. We successfully replicate the statistical properties of the HFRI and EDHEC indices over the period from 1997-2006, as well as for two 60 month sub-periods. Even without correcting for the well-documented biases in hedge fund index returns, the indices can be readily replicated using this methodology. The volatility and the dependence coefficients are replicated with great precision; the skewness and kurtosis are also captured by the model, however with slightly less accuracy.

Contrary to the conclusions put forth by recent studies at EDHEC and Northwater (2007), the choice of reserve asset does not impact the model's ability to replicate the statistical properties of the indices. The choice of reserve asset only impacts the initial cost of investing in the replicating portfolio (and hence only impacts the return of the replicating strategy). This is not to say that the return generated by the model is not important, however it is not a measure of the model's success. One must dissociate the technical issues of the replicating methodology (i.e how to best model the returns and solve for the optimal trading strategy) from the choice of the reserve asset. Our contribution is to provide a robust framework for the replication methodology, and address the technical shortcomings of the much publicized research of Kat and Palaro.

As is the case with any investment strategy, the returns depend on the choice of assets. The results in this paper indicate, however, that it is not necessary to select the best performing assets over the sample period in order to replicate and outperform the hedge fund indices. In fact, we show that by using run-of-the-mill exposures in our reserve asset we can nonetheless outperform the majority of hedge fund indices. We purposely selected two reserve assets that have exposures to different yet common market premia over the sample period, and we find that both reserve assets outperform a large percentage of the indices. (reserve 1 being the better of the two). We also find that the EDHEC indices, which are subject to less significant biases, are more easy to replicate that the HFRI indices. It is important to remember that we are comparing an investable trading strategy to non-investable indices- the actual return we would anticipate from investing in a hedge fund index would be considerably lower than the "non-investable" index returns used in this study. Our results reinforce the notion that on aggregate, hedge funds are on aggregate simply repackaging beta returns.

References

Agarwal, V. and Naik, N. (2004). Risks and portfolio decisions involving hedge funds. *Review of Financial Studies*, 17:63–98.

Amin, G. and Kat, H. (2003). Hedge fund performanc 1990-2000: Do the "money machines" rellay add value. *Journal of Financial and Quantitative Analysis*, 38(2):251–275.

Broadie, M. and Glasserman, P. (1996). Estimating security price derivatives using simulation. *Management Science*, 42:260–285.

Chen, Y. and Liang, B. (2006). Do market timing hedge funds time the market? Technical report, Carroll School of Management, Boston College.

Del Moral, P., Rémillard, B., and Rubenthaler, S. (2006). Monte Carlo approximations of American options. Technical report, GERAD.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 39:1–38.

Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.*, 1(2):279–290.

Dybvig, P. (1988). Distributional analysis of portfolio choice. *The Journal of Business*, 61(3):369–393.

Fung, W. and Hsieh, D. (2001). The risk in hedge fund stategies: Theory and evidence from trend followers. *Review of Financial Studies*, 14:313–341.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82:543–552.

Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the integral probability transformation. *Scand. J. Statist.*, 33:337–366.

Genest, C. and Rémillard, B. (2005). Validity of the parametric bootsrap for goodness-of-fit testing in semiparametric models. Technical Report G-2005-51, GERAD.

Genest, C., Rémillard, B., and Beaudoin, D. (2007). Omnibus goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.*, 40:in press.

Hasanhodzic, J. and Lo, A. W. (2007). Can hedge-fund returns be replicated?: The linear case. *Journal of Investment Management*, 5(2):5–45.

Kat, H. and Palaro, H. (2005). Who needs hedge funds? A copula-based approach to hedge fund return replication. Technical report, Cass Business School, City University.

Mitchell, M. and Pulvino, T. (2001). Characteristics of risk and return in risk arbitrage. *Journal of Finance*, LVI:2135–2175.

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

Northwater (2007). Northwater capital managements thoughts on hedge fund replication. *Technical report*.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.*, 23:470–472.

Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Math. Oper. Res.*, 20(1):1–32.

Stute, W., Gonzáles Manteiga, W., and Presedo Quindimil, M. (1993). Bootstrap based goodness-of-fit tests. *Metrika*, 40:243–256.

Wilmott, P. (2006). *Paul Wilmott on Quantitative Finance*, volume 3. John Wiley & Sons.

A Estimation and goodness-of-fit

In this section, we describe the estimation procedure and the goodness-of-fit tests.

A.1 EM algorithm for bivariate Gaussian mixtures

Let $y_1, \ldots, y_n$ be a random sample from a bivariate Gaussian mixture with parameters $\pi = (\pi_k)_{k=1}^m$, $\mu = (\mu_k)_{k=1}^m$ and $A = (A_k)_{k=1}^m$. Start with an initial estimator $\theta^{(0)}$. Given an estimator $\theta^{(\ell)} = \left(\pi^{(\ell)}, \mu^{(\ell)}, A^{(\ell)}\right)$ of the parameters $\theta = (\pi, \mu, A)$, set

$$\pi_k\left(y_i, \theta^{(\ell)}\right) = \frac{\pi_k^{(\ell)} \phi_2\left(y_i; \mu_k^{(\ell)}, A_k^{(\ell)}\right)}{\sum_{j=1}^m \pi_j^{(\ell)} \phi_2\left(y_i; \mu_j^{(\ell)}, A_j^{(\ell)}\right)}, \quad i = 1, \ldots, n,$$

and define the new estimator $\theta^{(\ell+1)} = \left(\pi^{(\ell+1)}, \mu^{(\ell+1)}, A^{(\ell+1)}\right)$ viz.

$$\pi_k^{(\ell+1)} = \frac{1}{n} \sum_{i=1}^n \pi_k\left(y_i, \theta^{(\ell)}\right),$$

$$\mu_k^{(\ell+1)} = \frac{1}{n} \sum_{i=1}^n y_i \pi_k\left(y_i, \theta^{(\ell)}\right) \Big/ \pi_k^{(\ell+1)},$$

and

$$A_k^{(\ell+1)} = \frac{1}{n} \sum_{i=1}^n \left(y_i - \mu_k^{(\ell+1)}\right)\left(y_i - \mu_k^{(\ell+1)}\right)^\top \pi_k\left(y_i, \theta^{(\ell)}\right) \Big/ \pi_k^{(\ell+1)},$$

for $k = 1, \ldots, m$. As ℓ increases, the numbers $\left\{\pi_k\left(y_i, \theta^{(\ell)}\right); k = 1, \ldots, i = 1, \ldots, n\right\}$ stabilize and the estimators converge.

A.2 Tests of goodness-of-fit

Testing goodness-of-fit is an essential step for modelling data. There are many tests available but to our knowledge, the best ones are based on empirical processes (Genest and Rémillard, 2005, Genest et al., 2007). Here, we only consider two tests based on the so-called Rosenblatt‘s transform. The first one is due to Durbin (1973) but the calculation of P-values is recent (Stute et al., 1993). For the second test designed for testing goodness-of-fit for bivariate data, the validity of the algorithm for calculating P-values follows from Genest and Rémillard (2005).

A.3 Tests of goodness-of-fit for a univariate parametric distribution

Let $X_1, \ldots, X_n$ be a sample of size n from a (continuous) distribution F on $\mathbb{R}$. Suppose that the hypotheses to be tested are

$$\mathcal{H}_0 : F \in \mathcal{F} = \{F_\theta; \theta \in \Theta\} \qquad \text{vs} \qquad \mathcal{H}_1 : F \notin \mathcal{F}$$

For example, the parametric family $\mathcal{F}$ could be the family of univariate Gaussian mixtures with m regimes.

The proposed test statistic is based on Durbin (1973). Let $\theta_n = T_n(X_1, \ldots, X_n)$ be a regular estimator of θ, in the sense of Genest and Rémillard (2005) and set

$$D_n(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(U_i \le u), \quad u \in [0, 1],$$

where $U_i = F_{\theta_n}(X_i)$, $i = 1, \ldots, n$. To test $\mathcal{H}_0$ against $\mathcal{H}_1$, one may use the Cramér-von Mises type statistic

$$\begin{aligned} S_n &= n \int_0^1 \{D_n(u) - u\}^2 du \\ &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ \frac{U_i^2 + U_j^2 - 2\max(U_i, U_j)}{2} + \frac{1}{3} \right\}. \end{aligned}$$

Since the U_i's are "almost uniformly distributed on $[0, 1]$" under the null hypothesis, large values of S_n should lead to rejection of the null hypothesis. However, in general the limiting distribution of S_n depend on the unknown parameter θ. To calculate the P-value of S_n, one can use a parametric bootstrap approach as described below.

a) Calculate θ_n and S_n.

b) For some large integer N (say 1000), repeat the following steps for every $k \in \{1, \ldots, N\}$:

   (i) Generate a random sample $X_{1,k}, \ldots, X_{n,k}$ from distribution F_{θ_n}.

   (ii) Calculate

$$\begin{aligned} \theta_{n,k} &= T_n\left(X_{1,k}, \ldots, X_{n,k}\right), \\ U_{i,k} &= F_{\theta_{n,k}}(X_{i,k}), \quad i = 1, \ldots, n, \\ S_{n,k} &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ \frac{U_{i,k}^2 + U_{j,k}^2 - 2\max(U_{i,k}, U_{j,k})}{2} + \frac{1}{3} \right\}. \end{aligned}$$

An approximate P-value for the test based on the Cramér–von Mises statistic S_n is then given by

$$\frac{1}{N}\sum_{k=1}^{N}\mathbb{I}(S_{n,k} > S_n).$$

A.4 Tests of goodness-of-fit for a bivariate parametric distribution

Let $(X_1, Y_1)\ldots,(X_n, Y_n)$ be a sample of size n from a (continuous) distribution F on $\mathbb{R}^2$. Suppose that the hypotheses to be tested are

$$\mathcal{H}_0 : F \in \mathcal{F} = \{F_\theta; \theta \in \Theta\} \qquad \text{vs} \qquad \mathcal{H}_1 : F \notin \mathcal{F}$$

For example, the parametric family $\mathcal{F}$ could be the family of bivariate Gaussian mixtures with m regimes. Denote by G_θ the distribution function of X_i and let H_θ be the conditional distribution function of Y_i given X_i, i.e., $H_\theta(x, y) = P(Y_i \le y|X_i = x)$.

The proposed test statistic is based on Durbin (1973) and the Rosenblatt's transform (Rosenblatt, 1952).

Suppose that $\theta_n = T_n(X_1, Y_1, \ldots, X_n, Y_n)$ is a regular estimator of θ, in the sense of Genest and Rémillard (2005) and set

$$D_n(u, v) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(U_i \le u, V_i \le v), \quad u, v \in [0, 1],$$

where $U_i = G_{\theta_n}(X_i)$, $V_i = H_{\theta_n}(X_i, Y_i)$, $i = 1, \ldots, n$. To test $\mathcal{H}_0$ against $\mathcal{H}_1$, one may use the Cramér-von Mises type statistic

$$\begin{aligned} S_n &= n\int_0^1\int_0^1 \{D_n(u, v) - uv\}^2 du dv \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{1}{9} - \frac{1}{4}(1 - U_i^2)(1 - V_i^2) - \frac{1}{4}(1 - U_j^2)(1 - V_j^2)\right. \\ &\qquad \left. + \{1 - \max(U_i, U_j)\}\{1 - \max(V_i, V_j)\}\right]. \end{aligned}$$

Since the pairs (U_i, V_i)'s are "almost uniformly distributed on $[0, 1]^2$" under the null hypothesis, large values of S_n should lead to rejection of the null hypothesis. However, in general the limiting distribution of S_n depend on the unknown parameter θ. To calculate the P-value of S_n, one can use a parametric bootstrap approach as described below.

a) Calculate θ_n and S_n.

b) For some large integer N (say 1000), repeat the following steps for every $k \in \{1, \ldots, N\}$:

(i) Generate a random sample $(X_{1,k}, Y_{1,k}), \ldots, (X_{n,k}, Y_{n,k})$ from distribution F_{θ_n}.

(ii) Calculate

$$\begin{aligned}
\theta^*_{n,k} &= T_n\left(X_{1,k}, Y_{1,k}, \ldots, X_{n,k}, Y_{n,k}\right), \\
U_{i,k} &= G_{\theta_{n,k}}(X_{i,k}), \quad V_{i,k} = H_{\theta_{n,k}}(X_{i,k}, Y_{i,k}), \quad i = 1, \ldots, n \\
S_{n,k} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{1}{9} - \frac{1}{4}(1-U_{i,k}^2)(1-V_{i,k}^2) - \frac{1}{4}(1-U_{j,k}^2)(1-V_{j,k}^2)\right. \\
&\qquad \left. + \{1-\max(U_{i,k}, U_{j,k})\}\{1-\max(V_{i,k}, V_{j,k})\}\right].
\end{aligned}$$

An approximate P-value for the test based on the Cramér–von Mises statistic S_n is then given by

$$\frac{1}{N}\sum_{k=1}^{N}\mathbb{I}(S_{n,k} > S_n).$$

B Implementation of the dynamic trading strategy

Before describing the algorithm, it is important to define what is meant by a partition. Here we assume that $S_t = \omega_t(S_{t-1}, \xi_t)$, $\xi_t \sim \mu_t$ being independent of $\mathcal{F}_{t-1}$, $t = 1, \ldots, T$.

Definition B.1 *A partition $\mathcal{P}$ of a compact convex set K, is any finite set $\mathcal{P} = \{S_1, \ldots, S_m\}$ of simplexes with disjoint non empty interiors, so that $K = \bigcup_{j=1}^m S_j$. The set of vertices of the partition $\mathcal{P}$ is denoted by $\mathcal{V}(\mathcal{P})$.*

Note that K is then the convex hull generated by $\mathcal{V}(\mathcal{P})$.

The algorithm is based on Monte Carlo simulations, combined with a sequence of approximations on compact sets $K_0, \ldots, K_{T-1}$, determined by partitions $\mathcal{P}_0, \ldots, \mathcal{P}_{T-1}$. The idea behind the algorithm is quite simple: Given approximations $\tilde{f}_t$, of f_t, one first get $\hat{L}_{1t}, \hat{L}_{2t}, \hat{A}_t, \hat{\Delta}_t, \hat{U}_t$ and $\hat{f}_{t-1}$, by estimating these functions at every vertices $x \in \mathcal{V}(\mathcal{P}_{t-1})$, using Monte Carlo simulations, and then, one uses a linear interpolation to extend them at any point $x \in K_{t-1}$. More precisely, one may proceed through the following steps.

B.1 Algorithm

- Set $\tilde{f}_T = f_T$;

- For each $t = T, \ldots, 1$
  - Generate $\xi_{1,t}, \ldots, \xi_{N_t,t}$ according to μ_t;
  - For every $s \in \mathcal{V}(\mathcal{P}_{t-1})$, calculate

$$\begin{aligned}
\hat{L}_{1t}(s) &= \frac{1}{N_t}\sum_{i=1}^{N_t} \omega_t(s, \xi_{i,t}) \\
\hat{L}_{2t}(s) &= \frac{1}{N_t}\sum_{i=1}^{N_t} \omega_t(s, \xi_{i,t})\omega_t(s, \xi_{i,t})^\top \\
\hat{A}_t(s) &= \hat{L}_{2t}(s) - \hat{L}_{1t}(s)\hat{L}_{1t}(s)^\top \\
\hat{\psi}_t(s) &= \hat{A}_t(s)^{-1}\frac{1}{N_t}\sum_{i=1}^{N_t}\{\omega_t(s, \xi_{i,t}) - \hat{L}_{1t}(s)\}\tilde{f}_t\{\omega_t(s, \xi_{i,t})\} \\
\hat{U}_t(s, x) &= 1 - \{\hat{L}_{1t}(s) - \beta_{t-1}s/\beta_t\}^\top \hat{A}_t(s)^{-1}\{x - \hat{L}_{1t}(s)\} \\
\hat{f}_{t-1}(s) &= \frac{\beta_t}{\beta_{t-1}}\frac{1}{N_t}\sum_{i=1}^{N_t}\hat{U}_t\{s, \omega_t(s, \xi_{i,t})\}\tilde{f}_t\{\omega_t(s, \xi_{i,t})\}.
\end{aligned}$$

  - Interpolate linearly $\hat{\Delta}_t$ and $\hat{f}_{t-1}$ over K_{t-1} and extend it to all of $\mathfrak{X}$.

A detailed description of the linear interpolation implementation techniques is given below, but first, the following result adapted from Del Moral et al. (2006), confirms that the algorithm produces good approximations.

Theorem 2 *Suppose that f_T is continuous and that for all $1 \le t \le T$, $\omega_t(\cdot, \xi)$ are continuous for a fixed ξ. Let K_0 be a given compact convex subset of $\mathfrak{X}$. Let $\epsilon > 0$ be given. Then one can find compact convex sets $K_1, \ldots, K_{n-1} \subset \mathfrak{X}$, partitions $\mathcal{P}_0, \ldots \mathcal{P}_{n-1}$ generating respectively $K_0, \ldots, K_{n-1}$, and integers $N_{10}, \ldots, N_{n0}$, so that for the simple interpolation method,*

$$\max_{1\le k\le n} \|\psi_t - \tilde{\psi}_t\|_{K_{t-1}} < \epsilon,$$

and

$$\max_{0\le k\le n-1} \|f_t - \tilde{f}_t\|_{K_t} < \epsilon,$$

whenever $N_1 \ge N_{10}, \ldots, N_n \ge N_{n0}$.

B.2 Linear interpolations

Definition B.2 *Given a function h and a partition $\mathcal{P}$ of K, a linear interpolation of h over $\mathcal{P}$ is the (unique) function $\tilde{g}$ defined in the following way:*

If $S \in \mathcal{P}$ is a simplex with vertices $x_1, \ldots, x_{d+1}$, then set

$$\tilde{h}(x) = \sum_{i=1}^{d+1} \lambda_i h(x_i),$$

where the barycenters $\{\lambda_1, \ldots, \lambda_{d+1}\}$ are the unique solution of

$$x = \sum_{i=1}^{d+1} \lambda_i x_i, \quad \sum_{i=1}^{d+1} \lambda_i = 1, \quad \lambda_i \in [0,1], i = 1, \ldots d+1.$$

If $x \notin K$, let x_K be the (unique) closest point to x that belongs to K, and set $\tilde{h}(x) = \tilde{h}(x_K)$. Uniqueness follows from the convexity of K and the strict convexity of the Euclidean norm.

Remark B.1 *Note that since each x_i is extreme in S, the unique solution of*

$$x_i = \sum_{j=1}^{d+1} \lambda_j x_j, \quad \sum_{j=1}^{d+1} \lambda_j = 1, \quad \lambda_j \in [0,1], j = 1, \ldots d+1,$$

is $\lambda_i = 1$ and $\lambda_j = 0$ for all $j \neq i$, yielding $\tilde{g}(x_i) = g(x_i)$ for all $1 \leq i \leq m$. Moreover, $\tilde{g}$ is affine on each simplex, justifying the term "linear interpolation".

Finally, $\tilde{g}$ is continuous and bounded on $\mathfrak{X}$ and

$$\sup_{x \in K} |g(x) - \tilde{g}(x)| \leq \omega(g, K, \mathrm{mesh}(\mathcal{P})),$$

where

$$\mathrm{mesh}(\mathcal{P}) = \max_{\mathrm{S} \in \mathcal{P}} \sup_{\mathrm{x,z} \in \mathrm{S}} \|\mathrm{x} - \mathrm{z}\|$$

and $\omega(g, K, \delta)$ is the modulus of continuity of g over K, i.e.

$$\omega(g, K, \delta) = \sup_{x,z \in K,\, \|x-z\| \leq \delta} |g(x) - g(z)|.$$

Example B.1 *Suppose $d = 1$. Then the linear interpolation $\tilde{g}$ of a monotone (respectively convex) function g on $K = [a,b]$ is monotone (respectively convex). To see that, set $a_i = a + i(b-a)/m$, $i = 0, \ldots, m$ and let $\mathcal{P}$ be the partition given by $\mathcal{P} = \{[a_{i-1}, a_i]; i = 1, \ldots, m\}$. Set $\Delta_i = \frac{g(a_i) - g(a_{i-1})}{a_i - a_{i-1}}$, $1 \leq i \leq m$. Then the linear interpolation of g over K is given by*

$$\tilde{h}(x) = \begin{cases} h(a), & x \leq a, \\ h(a_i) + (x - a_i)\Delta_{i+1}, & x \in [a_i, a_{i+1}],\ i = 0, \ldots, m-1, \\ h(b) & x \geq b. \end{cases}$$

If h is monotone, the slopes Δ_i all have the same sign, so $\tilde{h}$ has the same monotonicity. If h is convex, the slopes Δ_i are non decreasing, so $\tilde{h}$ is also convex.

Example B.2 *Suppose $d = 2$. First define interpolation on $[0,1]^2$. Suppose that h is known at points $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$. If one wants to linearly interpolate h, as in Definition B.2, a convenient choice for the partition $\mathcal{P}$ of $[0,1]^2$ is $\mathcal{P} = \{S_1, S_2\}$ where*

$$S_1 = \{(x_1,x_2) \in [0,1]^2; x_1 \le x_2\} \quad S_1 = \{(x_1,x_2) \in [0,1]^2; x_1 \ge x_2\}.$$

Any $x \in S_1$ can be uniquely written as

$$x = \lambda_1(0,1) + \lambda_2(1,1) + \lambda_3(0,0),$$

with $\lambda_2 = x_1$, $\lambda_1 = x_2 - x_1$, and $\lambda_3 = 1 - x_2$, so one can define

$$\begin{aligned} \tilde{h}(x) &= \lambda_1 h(0,1) + \lambda_2 h(1,1) + \lambda_3 h(0,0) \\ &= h(0,0) + x_1\{h(1,1) - h(0,1)\} + x_2\{h(0,1) - h(0,0)\}. \end{aligned}$$

Similarly, for any $x \in S_2$, one obtains

$$\begin{aligned} \tilde{h}(x) &= \lambda_1 h(0,1) + \lambda_2 h(1,1) + \lambda_3 h(0,0) \\ &= h(0,0) + x_1\{h(1,0) - h(0,0)\} + x_2\{h(1,1) - h(1,0)\}. \end{aligned}$$

Suppose now that $K = [a_1,b_1] \times [a_2,b_2]$ is partition into smaller rectangles. On each of these sub-rectangles $R = [y_1,y_2] \times [z_1,z_2]$, just use the linear interpolation on $[0,1]^2$ by transforming $x \in R$ into $x' = (x_1', x_2') \in [0,1]^2$ through the mapping $x_1' = \frac{x_1-y_1}{y_2-y_1}, x_2' = \frac{x_2-z_1}{z_2-z_1}$.

Outside K, $\tilde{h}$ is defined as follows:

$$\tilde{h}(x) = \begin{cases} \tilde{h}(x_1,a_2) & \text{if} \quad x \in [a_1,b_1] \times (-\infty, a_2) \\ \tilde{h}(x_1,b_2) & \text{if} \quad x \in [a_1,b_1] \times (b_2, \infty) \\ \tilde{h}(a_1,x_2) & \text{if} \quad x \in (-\infty, a_1) \times [a_2,b_2] \\ \tilde{h}(b_1,x_2) & \text{if} \quad x \in (b_1, \infty) \times [a_2,b_2] \\ \tilde{h}(a_1,a_2) & \text{if} \quad x \in (-\infty, a_1) \times (-\infty, a_2) \\ \tilde{h}(b_1,a_2) & \text{if} \quad x \in (b_1, \infty) \times (-\infty, a_2) \\ \tilde{h}(a_1,b_2) & \text{if} \quad x \in (-\infty, a_1) \times (b_2, \infty) \\ \tilde{h}(b_1,b_2) & \text{if} \quad x \in (b_1, \infty) \times (b_2, \infty) \end{cases}.$$

C Auxiliary results

Throughout this appendix, $L^2 = L^2(\Omega, \mathcal{F}, P)$ is the set of all random variables on $(\Omega, \mathcal{F})$ which are square integrable.

Proposition 1 *Suppose that X is non negative random variable on $(\Omega, \mathcal{F}, P)$ such that $E(X) < \infty$. Suppose $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$ and let $Z = E(X|\mathcal{G}) \geq 0$, P almost surely. Then for any non negative $\mathcal{G}$-measurable random variable ξ, the following equality holds*

$$E(\xi X) = E(\xi Z).$$

Proof In the case of bounded random variable ξ, the result follows from the very definition of the conditional expectation. In particular it is true for $\xi_n = \min(n, \xi) \geq 0$, for any $n \geq 1$. Since $\xi_n \uparrow \xi$, it follows from Beppo-Levy theorem that

$$E(\xi X) = \lim_{n\to\infty} E(\xi_n X) = \lim_{n\to\infty} E(\xi_n Z) = E(\xi Z).$$

Proposition 2 *Suppose that $\xi \in \mathbb{R}^d$ and $\eta \in \mathbb{R}$ are L^2 random variables in $(\Omega, \mathcal{F})$ and suppose that $A = E(\xi\xi^\top|\mathcal{G})$ is invertible, where $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$. Then $\varphi \in \mathbb{R}^d$ minimizes $E\{(\varphi^\top\xi - \eta)^2\}$ over all $\varphi \in \mathcal{G}$ such that $\varphi^\top\xi \in L^2$ if and only if $\varphi = A^{-1}b$, where $b = E(\xi\eta|\mathcal{G})$. In particular $\varphi^\top\xi$ is square integrable.*

Proof Set $\varphi = A^{-1}b$. To prove that $\varphi^\top\xi \in L^2$, note that it follows from Proposition 1 that

$$\begin{aligned} E\left\{(\varphi^\top\xi)^2\right\} &= \sum_{i=1}^d E(\varphi_i^2\xi_i^2) \\ &= \sum_{i=1}^d E\{\varphi_i^2 E(\xi_i^2|\mathcal{G})\} \\ &= \sum_{i=1}^d E(\varphi_i^2 A_{ii}) \\ &= E(b^\top A^{-1} b). \end{aligned}$$

Since A is symmetric and positive definite, there exist a $d \times d$ matrix $M \in \mathcal{G}$ such that $M^{-1} = M^\top$ and a $d \times d$ diagonal matrix $\Delta \in \mathcal{G}$ such that $A = M\Delta M^\top$. Set $\tilde{\xi} = M^\top\xi$ and $\tilde{b} = M^\top b$. Then $\Delta = E(\tilde{\xi}\tilde{\xi}^\top|\mathcal{G})$, $\tilde{b} = E(\tilde{\xi}\eta|\mathcal{G})$, $E(\tilde{\xi}_i^2|\mathcal{G}) = \Delta_{ii} > 0$ by hypothesis, and

$$\begin{aligned} b^\top A^{-1} b &= \tilde{b}^\top \Delta^{-1} \tilde{b} \\ &= \sum_{i=1}^d \frac{E^2(\tilde{\xi}_i\eta|\mathcal{G})}{E(\tilde{\xi}_i^2|\mathcal{G})} \\ &\leq dE(\eta^2|\mathcal{G}) \text{ a.s. }, \end{aligned}$$

from Cauchy-Schwarz inequality. Hence

$$E\{(\varphi^\top \xi)^2\} \le pE(\eta^2) < \infty.$$

Next, let ψ be any random vector in $\mathcal{G}$ such that $\psi^\top \xi \in L^2$. Then

$$E\{(\psi^\top \xi - \eta)^2\} = E\left[E\{(\psi^\top \xi - \eta)^2 | \mathcal{G}\}\right],$$

and it is easy to check that

$$\begin{aligned} E\{(\psi^\top \xi - \eta)^2 | \mathcal{G}\} &= \psi^\top A \psi - 2\psi^\top b + c \\ &= (\psi - \varphi)^\top A(\psi - \varphi) + \varphi^\top A \varphi - 2\varphi^\top b + c \\ &= (\psi - \varphi)^\top A(\psi - \varphi) + E\{(\varphi^\top \xi - \eta)^2 | \mathcal{G}\}. \end{aligned}$$

Hence the result.

D Proof of the main results

In this section, we will prove the two main results, using the propositions proved in Appendix C.

D.1 Proof of Theorem 1

Recall that the process $\varphi = (\varphi_t)_{t=0}^T$ is predictable. For any $1 \le t \le T$, set $\Delta_t = S_t - E(S_t|\mathcal{F}_{t-1})$ and

$$G_t = {\varphi_t}^\top \Delta_t - \{C_t - E(C_t|\mathcal{F}_{t-1})\}, \tag{9}$$

where $C_T = C$ and

$$\beta_{t-1} C_{t-1} = E(\beta_t C_t | \mathcal{F}_{t-1}) - {\varphi_t}^\top E(\beta_t S_t - \beta_{t-1} S_{t-1} | \mathcal{F}_{t-1}). \tag{10}$$

It follows from equations (9)-(10) that

$$\beta_t G_t = \beta_{t-1} C_{t-1} - \beta_t C_t + {\varphi_t}^\top (\beta_t S_t - \beta_{t-1} S_{t-1}), \qquad 1 \le t \le T. \tag{11}$$

Note that the $G_t \in \mathcal{F}_t$ and $E(G_t|\mathcal{F}_{t-1}) = 0$, for all $1 \le t \le T$. Moreover, using (4)–(5) and (11), one gets

$$\sum_{t=1}^T \beta_t G_t = C_0 - \beta_T C + \sum_{t=1}^T \varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1}) = G + C_0 - V_0$$

and $E(G) = E(G|\mathcal{F}_0) = C_0 - V_0$, since $E(G_t|\mathcal{F}_{t-1}) = 0$ for all $t = 1, \ldots, T$. It also follows from well known properties of conditional expectations that

$$\begin{aligned} E(G^2) &= E(G^2|\mathcal{F}_0) = (C_0 - V_0)^2 + \sum_{t=1}^{T} E\left(\beta_t^2 G_t^2 |\mathcal{F}_0\right) \\ &= (C_0 - V_0)^2 + \sum_{t=1}^{T} E\left\{\beta_t^2 \, E\left(G_t^2|\mathcal{F}_{t-1}\right) \big| \, \mathcal{F}_0\right\}. \end{aligned} \tag{12}$$

Because G_t depends only on $\varphi_t, \ldots, \varphi_T$ through C_t, to minimize $E(G^2)$, it suffices to find φ_T minimizing $E\left(G_T^2|\mathcal{F}_0\right)$, then to find φ_{T-1} minimizing $E\left(G_{T-1}^2|\mathcal{F}_0\right)$ and so on. Doing so, we will find the minimum since each term is non negative. Having found the optimal φ, one obtains that the optimal choice for V_0 is C_0.

First, note that $G_T = \xi_T^\top \varphi_T - \eta_T$, where $\xi_T = \Delta_T = S_T - E(S_T|\mathcal{F}_{T-1})$ and $\eta_T = C - E(C|\mathcal{F}_{T-1}) = C_T - E(C_T|\mathcal{F}_{T-1})$.

Using Proposition 2, one can conclude that

$$\varphi_T = (\Sigma_T)^{-1} E\left(\xi_T \eta_T|\mathcal{F}_{T-1}\right) = (\Sigma_T)^{-1} E\left(\xi_T C_T|\mathcal{F}_{T-1}\right)$$

minimizes $E(G_T^2|\mathcal{F}_0)$. Having found the optimal φ_T, one can define C_{T-1} as in (10).

Suppose now that $\varphi_T, \ldots, \varphi_t$ have been defined and define G_{t-1} and C_{t-1} according to (9) and (10). Then one can use again Proposition (2) to conclude φ_{t-1} given by (6) minimizes $E(G_{t-1}^2|\mathcal{F}_0)$.

Therefore the risk $E(G^2|\mathcal{F}_0)$ is minimized by choosing the φ_t's according to (6). Finally, using (12), the optimal value of V_0 is C_0. This completes the proof.

D.2 Proof of Corollary 1

The proof of the representation $C_{t-1} = E(C_t U_t|\mathcal{F}_{t-1})$ follows directly from Theorem 1. In fact, using equations (6) and (7), one obtains

$$\begin{aligned} \beta_{t-1} C_{t-1} &= E(\beta_t C_t|\mathcal{F}_{t-1}) - \varphi_t^\top E(\beta_t S_t - \beta_{t-1} S_{t-1}|\mathcal{F}_{t-1}) \\ &= E(\beta_t C_t|\mathcal{F}_{t-1}) \\ &\quad - E\left\{ C_t \Delta_t^\top (\Sigma_t)^{-1} E(\beta_t S_t - \beta_{t-1} S_{t-1}|\mathcal{F}_{t-1}) \big| \, \mathcal{F}_{t-1}\right\} \\ &= E(C_t U_t|\mathcal{F}_{t-1}), \end{aligned}$$

where U_t is defined by (8). One can easily see that $E(U_t|\mathcal{F}_{t-1}) = 1$, so $(M_t)_{t=0}^T$ is a martingale.

It only remains to prove that $\beta_t S_t M_t$ is a martingale. All is needed is to prove that $E(\beta_t S_t U_t|\mathcal{F}_{t-1}) = \beta_{t-1} S_{t-1}$. To this end, let $t \in \{1, \ldots, T\}$ be given and set $\xi_t =$

$E\left(\beta_t S_t - \beta_{t-1} S_{t-1} | \mathcal{F}_{t-1}\right)$. Note that

$$\beta_t S_t U_t = \beta_t S_t - \left\{\Delta_t + E(S_t|\mathcal{F}_{t-1})\right\} \Delta_t^\top \left(\Sigma_t\right)^{-1} \xi_t.$$

Next, since $E(\Delta_t|\mathcal{F}_{t-1}) = 0$, one has

$$\begin{aligned}
E(\beta_t S_t U_t|\mathcal{F}_{t-1}) &= E(\beta_t S_t|\mathcal{F}_{t-1}) - E(\Delta_t \Delta_t^\top|\mathcal{F}_{t-1}) \left(\Sigma_t\right)^{-1} \xi_t \\
&\quad - E(S_t|\mathcal{F}_{t-1}) E(\Delta_t^\top|\mathcal{F}_{t-1}) \left(\Sigma_t\right)^{-1} \xi_t \\
&= E(\beta_t S_t|\mathcal{F}_{t-1}) - \Sigma_t \left(\Sigma_t\right)^{-1} \xi_t - 0 \\
&= E(\beta_t S_t|\mathcal{F}_{t-1}) - \xi_t = \beta_{t-1} S_{t-1}.
\end{aligned}$$

Hence the result.

Table 8: Marginal distribution, copula and Kendall's tau for entire period (1997–2006).

Fund	Marginal	Copula	Kendall's tau
EDHEC-Convertible Arbitrage	GM(3)	Frank	0.0927
EDHEC-CTA Global	GM(2)	Gumbel	0.0552
EDHEC-Distressed Securities	GM(2)	Clayton	0.2311
EDHEC-Emerging Markets	Johnson	Frank	0.3394
EDHEC-Equity Market Neutral	GM(2)	Frank	0.2302
EDHEC-Event Driven	GM(3)	Frank	0.3724
EDHEC-Fixed Income Arbitrage	GM(3)	Frank	0.0997
EDHEC-Global Macro	GM(3)	Frank	0.3316
EDHEC-Long/Short Equity	GM(2)	Student	0.4529
EDHEC-Merger Arbitrage	GM(2)	Frank	0.2956
EDHEC-Relative Value	GM(3)	Gaussian	0.3324
EDHEC-Short Selling	GM(2)	Frank	-0.4636
EDHEC-Funds of Funds	GM(4)	Gaussian	0.3536
HFRI Convertible Arbitrage Index	GM(3)	Frank	0.1048
HFRI Distressed Securities Index	GM(3)	Clayton	0.2160
HFRI Emerging Markets (Total)	Johnson	Student	0.3269
HFRI Equity Hedge Index	GM(2)	Clayton	0.4530
HFRI Equity Market Neutral Index	GM(3)	Frank	0.1345
HFRI Equity Non-Hedge Index	GM(3)	Student	0.4770
HFRI Event-Driven Index	GM(3)	Clayton	0.3700
HFRI Fixed Income (Total)	GM(3)	Frank	0.3168
HFRI Fixed Income: Arbitrage Index	GM(3)	Ind.	0
HFRI Fixed Income: High Yield Index	GM(2)	Student	0.2036
HFRI FOF: Conservative Index	Johnson	Frank	0.3021
HFRI FOF: Diversified Index	GM(3)	Frank	0.2945
HFRI FOF: Market Defensive Index	GM(2)	Frank	0.1020
HFRI FOF: Strategic Index	GM(3)	Frank	0.3555
HFRI FOF Composite Index	GM(3)	Frank	0.3327
HFRI FOF Composite Index (Off.)	GM(3)	Frank	0.3180
HFRI Fund Weighted Composite Index	GM(3)	Clayton	0.4403
HFRI Macro Index	GM(2)	Clayton	0.2364
HFRI Merger Arbitrage Index	GM(3)	Frank	0.2568
HFRI Regulation D Index	GM(3)	Gaussian	0.2210
HFRI Relative Value Arbitrage Index	GM(3)	Gaussian	0.2567
HFRI Short Selling Index	GM(3)	Frank	-0.4520

Table 9: Marginal distribution, copula and Kendall's tau for first sub-period (1997–2001).

Fund	Marginal	Copula	Kendall's tau
EDHEC-Convertible Arbitrage	GM(3)	Gumbel	0.0777
EDHEC-CTA Global	GM(2)	Ind.	0
EDHEC-Distressed Securities	GM(3)	Clayton	0.2309
EDHEC-Emerging Markets	GM(3)	Frank	0.3241
EDHEC-Equity Market Neutral	GM(2)	Gaussian	0.3691
EDHEC-Event Driven	Johnson	Clayton	0.3793
EDHEC-Fixed Income Arbitrage	GM(3)	Frank	0.1268
EDHEC-Global Macro	GM(3)	Frank	0.4198
EDHEC-Long/Short Equity	GM(2)	Frank	0.4868
EDHEC-Merger Arbitrage	GM(4)	Gumbel	0.2951
EDHEC-Relative Value	GM(3)	Clayton	0.3454
EDHEC-Short Selling	GM(2)	Frank	-0.4695
EDHEC-Funds of Funds	GM(2)	Frank	0.3934
HFRI Convertible Arbitrage Index	GM(3)	Frank	0.1011
HFRI Distressed Securities Index	GM(3)	Gaussian	0.1939
HFRI Emerging Markets (Total)	GM(3)	Frank	0.3148
HFRI Equity Hedge Index	GM(2)	Frank	0.4880
HFRI Equity Market Neutral Index	GM(2)	Frank	0.1607
HFRI Equity Non-Hedge Index	Johnson	Frank	0.4962
HFRI Event-Driven Index	GM(3)	Frank	0.3461
HFRI Fixed Income (Total)	GM(3)	Frank	0.3078
HFRI Fixed Income: Arbitrage Index	GM(3)	Ind.	0
HFRI Fixed Income: High Yield Index	GM(3)	Frank	0.2367
HFRI FOF: Conservative Index	GM(3)	Frank	0.3310
HFRI FOF: Diversified Index	Johnson	Frank	0.2915
HFRI FOF: Market Defensive Index	GM(3)	Frank	0.1257
HFRI FOF: Strategic Index	GM(3)	Frank	0.3600
HFRI FOF Composite Index	GM(2)	Frank	0.3427
HFRI FOF Composite Index (Off.)	GM(2)	Frank	0.3276
HFRI Fund Weighted Composite Index	GM(3)	Frank	0.4567
HFRI Macro Index	GM(2)	Clayton	0.2975
HFRI Merger Arbitrage Index	Johnson	Gumbel	0.2285
HFRI Regulation D Index	GM(3)	Gaussian	0.2736
HFRI Relative Value Arbitrage Index	GM(3)	Frank	0.2705
HFRI Short Selling Index	GM(2)	Frank	-0.4402

Table 10: Marginal distribution, copula and Kendall's tau for second sub-period (2002–2006).

Fund	Marginal	Copula	Kendall's tau
EDHEC-Convertible Arbitrage	GM(3)	Gaussian	0.0885
EDHEC-CTA Global	GM(2)	Frank	0.0743
EDHEC-Distressed Securities	GM(2)	Gaussian	0.2224
EDHEC-Emerging Markets	GM(3)	Frank	0.2710
EDHEC-Equity Market Neutral	Johnson	Frank	0.0896
EDHEC-Event Driven	Johnson	Gaussian	0.3052
EDHEC-Fixed Income Arbitrage	GM(3)	Ind.	0
EDHEC-Global Macro	GM(2)	Gaussian	0.1987
EDHEC-Long/Short Equity	GM(2)	Clayton	0.3377
EDHEC-Merger Arbitrage	GM(3)	Clayton	0.3126
EDHEC-Relative Value	GM(2)	Clayton	0.2973
EDHEC-Short Selling	GM(2)	Frank	-0.4266
EDHEC-Funds of Funds	Johnson	Clayton	0.2470
HFRI Convertible Arbitrage Index	GM(3)	Gumbel	0.0743
HFRI Distressed Securities Index	GM(2)	Clayton	0.2109
HFRI Emerging Markets (Total)	GM(3)	Frank	0.2797
HFRI Equity Hedge Index	GM(3)	Frank	0.2993
HFRI Equity Market Neutral Index	GM(2)	Frank	0.0874
HFRI Equity Non-Hedge Index	GM(2)	Frank	0.3687
HFRI Event-Driven Index	GM(3)	Gaussian	0.3377
HFRI Fixed Income (Total)	GM(3)	Gaussian	0.2303
HFRI Fixed Income: Arbitrage Index	GM(3)	Ind.	0
HFRI Fixed Income: High Yield Index	GM(2)	Gumbel	0.1311
HFRI FOF: Conservative Index	GM(2)	Frank	0.2164
HFRI FOF: Diversified Index	GM(2)	Clayton	0.2437
HFRI FOF: Market Defensive Index	GM(2)	Frank	0.0831
HFRI FOF: Strategic Index	GM(3)	Clayton	0.2885
HFRI FOF Composite Index	GM(2)	Clayton	0.2383
HFRI FOF Composite Index (Off.)	GM(2)	Clayton	0.2164
HFRI Fund Weighted Composite Index	GM(2)	Frank	0.3243
HFRI Macro Index	GM(2)	Gumbel	0.0787
HFRI Merger Arbitrage Index	GM(2)	Clayton	0.2984
HFRI Regulation D Index	Johnson	Clayton	0.1552
HFRI Relative Value Arbitrage Index	GM(2)	Clayton	0.2328
HFRI Short Selling Index	GM(2)	Frank	-0.4319

Table 11: Initial investment V_0 in the replication of EDHEC and HFRI indices for both reserve assets over the entire period (1997–2006).

Fund	V_0	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	99.88746927	100.3546058
EDHEC-CTA Global	99.22395238	100.2822217
EDHEC-Distressed Securities	100.0433158	100.5343205
EDHEC-Emerging Markets	99.20994993	100.5118262
EDHEC-Equity Market Neutral	100.0923959	100.3305248
EDHEC-Event Driven	99.99904541	100.5027729
EDHEC-Fixed Income Arbitrage	99.68524183	100.0620038
EDHEC-Global Macro	99.83012861	100.4453958
EDHEC-Long/Short Equity	99.91948345	100.5253251
EDHEC-Merger Arbitrage	99.94738788	100.3347095
EDHEC-Relative Value	100.044295	100.3582369
EDHEC-Short Selling	97.91881695	99.96879961
EDHEC-Funds of Funds	99.88679097	100.4167799
Percentage of V_0 under 100$	76.92%	7.69%
HFRI Convertible Arbitrage Index	99.9104685	100.321649
HFRI Distressed Securities Index	99.9100765	100.446987
HFRI Emerging Markets (Total)	99.1617091	100.497154
HFRI Equity Hedge Index	99.760536	100.537810
HFRI Equity Market Neutral Index	99.8160615	100.178244
HFRI Equity Non-Hedge Index	99.2694693	100.529065
HFRI Event-Driven Index	99.8678282	100.443743
HFRI Fixed Income (Total)	99.8533463	100.180401
HFRI Fixed Income: Arbitrage Index	99.4744962	99.9612590
HFRI Fixed Income: High Yield Index	99.4606113	100.118320
HFRI FOF: Conservative Index	99.8019766	100.171418
HFRI FOF: Diversified Index	99.5428340	100.224120
HFRI FOF: Market Defensive Index	99.6295097	100.290348
HFRI FOF: Strategic Index	99.3496291	100.310468
HFRI FOF Composite Index	99.6186407	100.240115
HFRI FOF Composite Index (Off.)	99.4353982	100.150926
HFRI Fund Weighted Composite Index	99.7328707	100.309632
HFRI Macro Index	99.6917718	100.369990
HFRI Merger Arbitrage Index	99.8584340	100.285088
HFRI Regulation D Index	99.9386375	100.681884
HFRI Relative Value Arbitrage Index	100.055301	100.346992
HFRI Short Selling Index	97.5229297	99.8979799
Percentage of V_0 under 100$	95.45%	9.09%

Table 12: Initial investment V_0 in the replication of EDHEC and HFRI indices for both reserve assets for first sub-period (1997–2001).

Fund	V_0	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	100.2944853	100.8987467
EDHEC-CTA Global	99.46788172	100.9472588
EDHEC-Distressed Securities	99.90059626	100.720508
EDHEC-Emerging Markets	98.69192451	100.9835661
EDHEC-Equity Market Neutral	100.3954179	100.7210641
EDHEC-Event Driven	100.0609365	100.868357
EDHEC-Fixed Income Arbitrage	99.59077798	100.2017969
EDHEC-Global Macro	99.97142407	100.979203
EDHEC-Long/Short Equity	100.1375749	101.127196
EDHEC-Merger Arbitrage	100.2331299	100.7861365
EDHEC-Relative Value	100.2203665	100.6965085
EDHEC-Short Selling	99.03421453	102.1095181
EDHEC-Funds of Funds	99.96160577	100.9516279
Percentage of V_0 under 100$	53.84%	0.00%
HFRI Convertible Arbitrage Index	100.2829484	100.8055676
HFRI Distressed Securities Index	99.72936197	100.6646377
HFRI Emerging Markets (Total)	98.09524276	100.9525596
HFRI Equity Hedge Index	100.056951	101.5042088
HFRI Equity Market Neutral Index	100.038409	100.6734399
HFRI Equity Non-Hedge Index	99.05531596	101.2392224
HFRI Event-Driven Index	99.97242706	100.980233
HFRI Fixed Income (Total)	99.75412401	100.3504572
HFRI Fixed Income: Arbitrage Index	99.3254573	100.0324407
HFRI Fixed Income: High Yield Index	99.31890751	100.1544936
HFRI FOF: Conservative Index	99.86644524	100.4768867
HFRI FOF: Diversified Index	99.52279888	100.9361689
HFRI FOF: Market Defensive Index	99.80508973	100.7630553
HFRI FOF: Strategic Index	99.28992499	100.9862717
HFRI FOF Composite Index	99.60846434	100.7634087
HFRI FOF Composite Index (Off.)	99.36188049	100.7094413
HFRI Fund Weighted Composite Index	99.75155852	100.9238131
HFRI Macro Index	99.76812518	100.8842713
HFRI Merger Arbitrage Index	100.1469401	100.7111258
HFRI Regulation D Index	100.5815208	101.5412257
HFRI Relative Value Arbitrage Index	100.1412334	100.6153432
HFRI Short Selling Index	98.51962283	100.8070637
Percentage of V_0 under 100$	72.72%	0.00%

Table 13: Initial investment V_0 in the replication of EDHEC and HFRI indices for both reserve assets for second sub-period (2002–2006).

Fund	V_0	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	99.54307232	99.91754606
EDHEC-CTA Global	99.00591261	99.85286557
EDHEC-Distressed Securities	100.2752537	100.645535
EDHEC-Emerging Markets	100.0757102	100.4515871
EDHEC-Equity Market Neutral	99.85680498	99.99383759
EDHEC-Event Driven	99.87363087	100.3162115
EDHEC-Fixed Income Arbitrage	99.88868645	100.0907229
EDHEC-Global Macro	99.84474995	100.2384539
EDHEC-Long/Short Equity	99.60337666	100.1087833
EDHEC-Merger Arbitrage	99.70200103	99.99705359
EDHEC-Relative Value	99.81967336	100.109444
EDHEC-Short Selling	98.05685558	99.04396197
EDHEC-Funds of Funds	99.74332198	100.0559835
Percentage of V_0 under 100$	84.62%	38.46%
HFRI Convertible Arbitrage Index	99.60821174	99.93483497
HFRI Distressed Securities Index	100.2391759	100.6380069
HFRI Emerging Markets (Total)	100.0572944	100.8595669
HFRI Equity Hedge Index	99.58364075	100.014334
HFRI Equity Market Neutral Index	99.66759956	99.85722405
HFRI Equity Non-Hedge Index	99.37314042	100.2792862
HFRI Event-Driven Index	99.80612072	100.3402519
HFRI Fixed Income (Total)	99.95688427	100.1391919
HFRI Fixed Income: Arbitrage Index	100.0072767	100.1695353
HFRI Fixed Income: High Yield Index	100.0771647	100.3417642
HFRI FOF: Conservative Index	99.82149692	100.0377755
HFRI FOF: Diversified Index	99.7547993	100.0216789
HFRI FOF: Market Defensive Index	99.56207483	99.97381601
HFRI FOF: Strategic Index	99.62610152	99.96801828
HFRI FOF Composite Index	99.73892366	100.0563079
HFRI FOF Composite Index (Off.)	99.68519975	100.0475484
HFRI Fund Weighted Composite Index	99.78329249	100.2000232
HFRI Macro Index	99.73199639	100.3030235
HFRI Merger Arbitrage Index	99.66510204	100.0050475
HFRI Regulation D Index	99.47794411	100.3049513
HFRI Relative Value Arbitrage Index	99.94588108	100.1510614
HFRI Short Selling Index	98.37750341	99.15058551
Percentage of V_0 under 100$	81.81%	22.73%

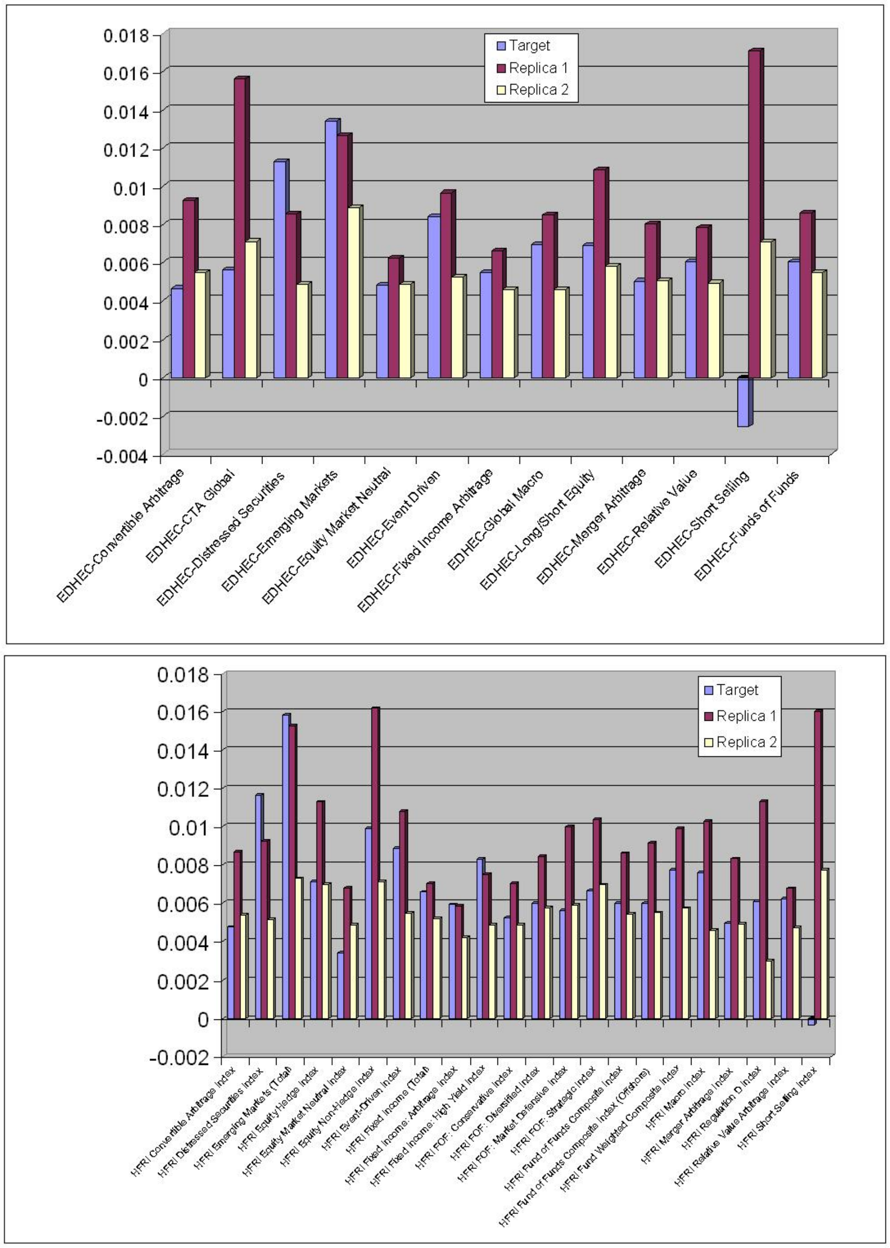

Figure 2: Mean return of replication for both reserve assets vs mean return for EDHEC (top) and HFRI (bottom) indices (2002–2006)

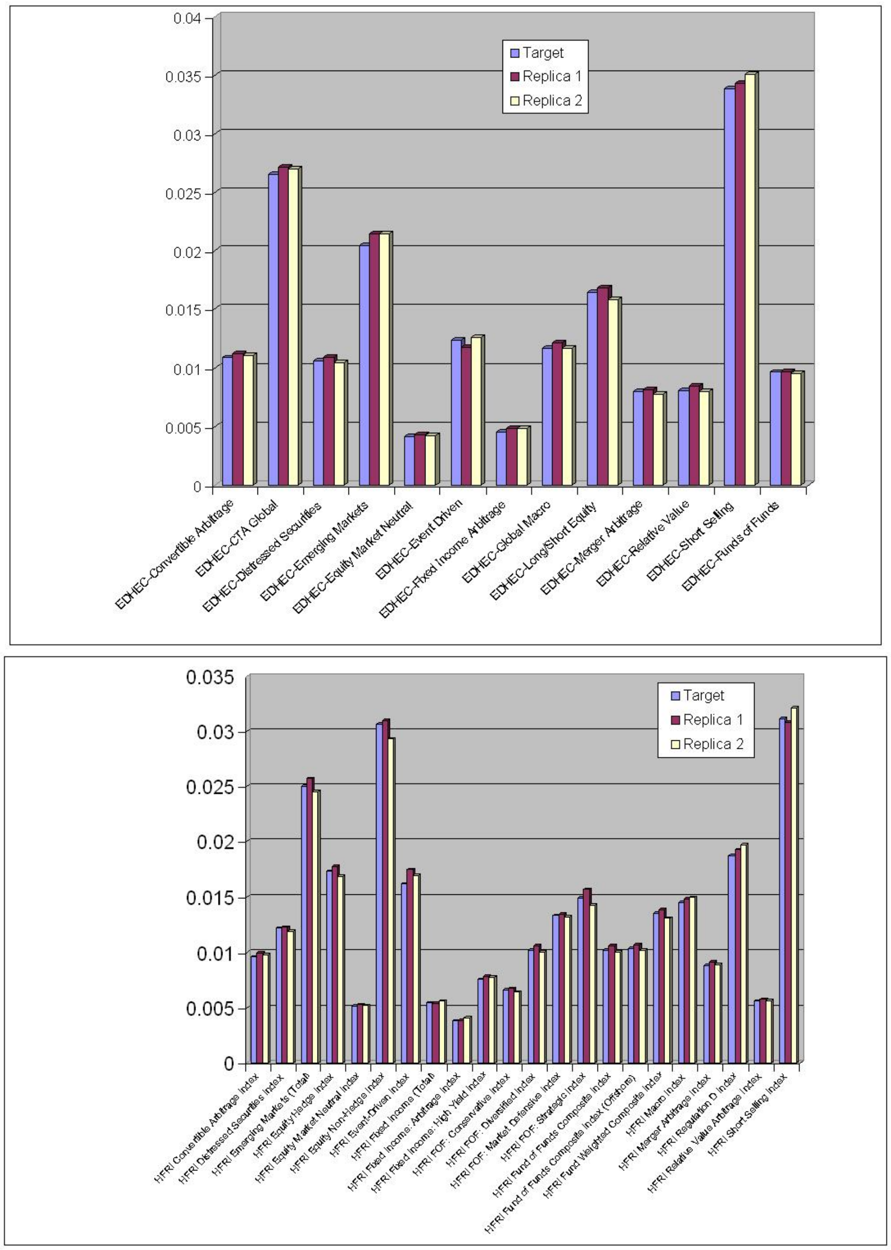

Figure 3: Volatility of the replication with each reserve asset vs target volatility for EDHEC (top) and HFRI (bottom) indices (2002–2006)

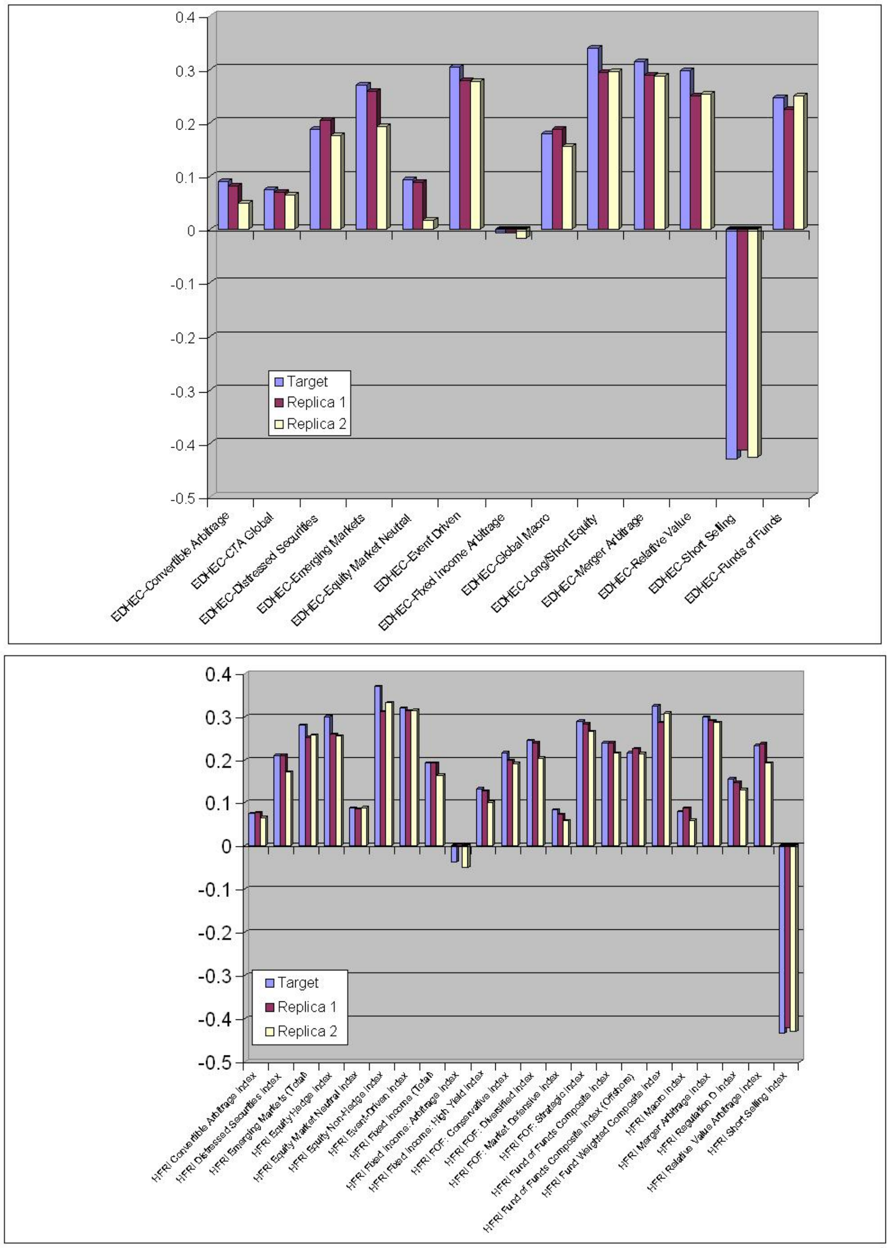

Figure 4: Kendall's tau of the replication with each reserve asset vs target Kendall's tau for EDHEC (top) and HFRI (bottom) indices (2002–2006)

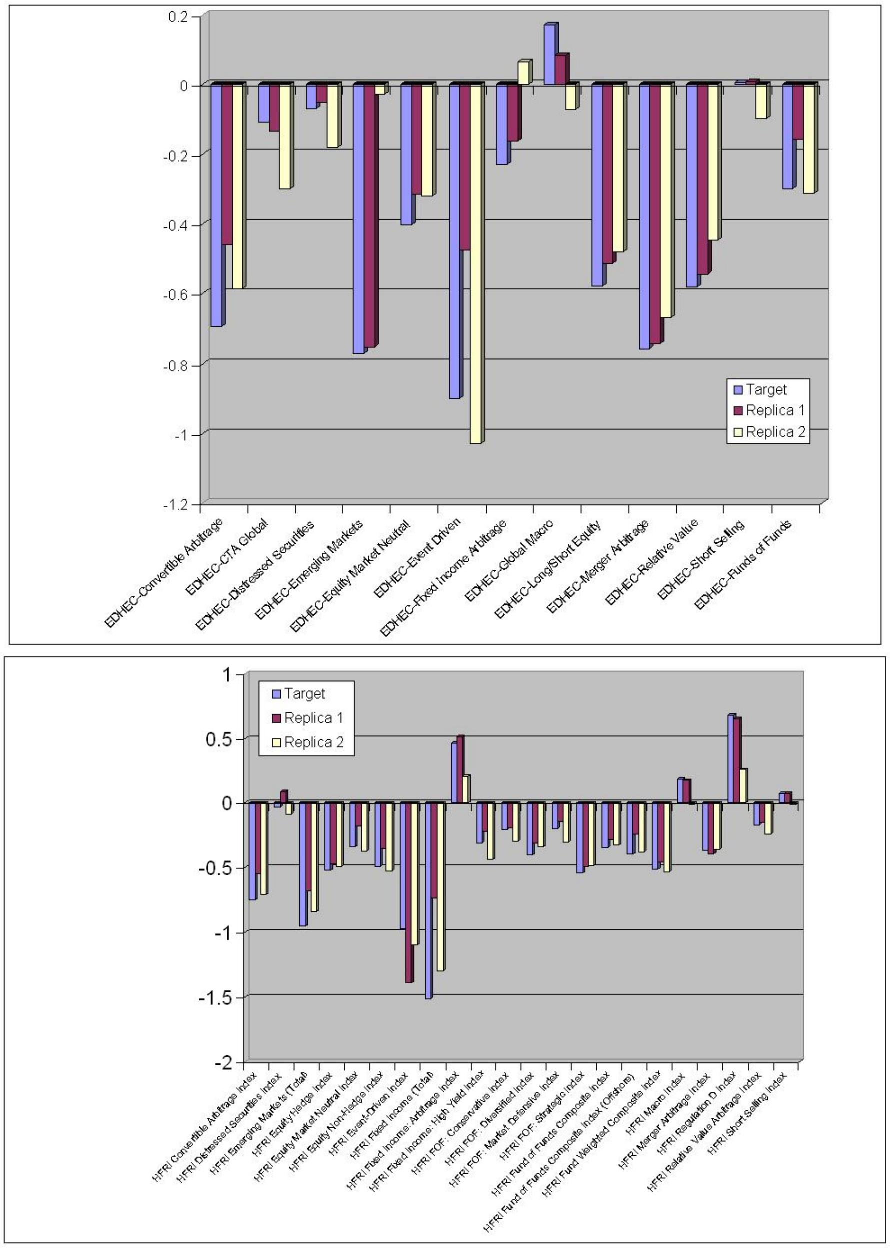

Figure 5: Skewness of the replication with each reserve asset vs target skewness for EDHEC (top) and HFRI (bottom) indices (2002–2006)

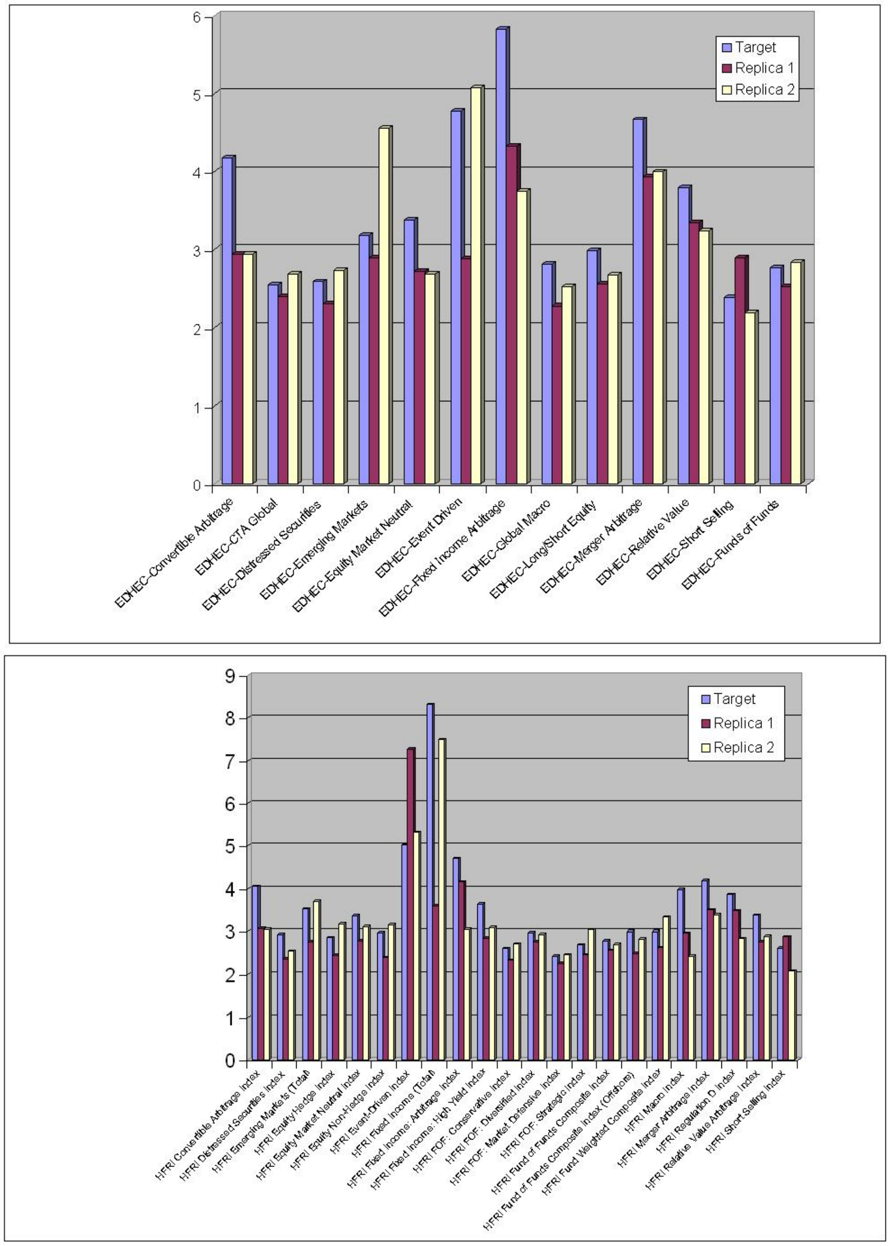

Figure 6: Kurtosis of the replication with each reserve asset vs target kurtosis for EDHEC (top) and HFRI (bottom) indices (2002–2006)

Table 14: Transaction costs (basis points) of the EDHEC and HFRI indices for each of two reserve assets over the entire period (1997–2006).

Fund	Transaction costs	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	-3.5760	-2.6937
EDHEC-CTA Global	-5.1209	-3.6392
EDHEC-Distressed Securities	-3.1461	-2.9916
EDHEC-Emerging Markets	-10.436	-8.5692
EDHEC-Equity Market Neutral	-1.1785	-1.2782
EDHEC-Event Driven	-4.9894	-3.7833
EDHEC-Fixed Income Arbitrage	-5.6955	-3.5177
EDHEC-Global Macro	-3.1539	-3.5487
EDHEC-Long/Short Equity	-3.5405	-3.5815
EDHEC-Merger Arbitrage	-3.5994	-2.7794
EDHEC-Relative Value	-2.1994	-1.8390
EDHEC-Short Selling	-14.472	-12.690
EDHEC-Funds of Funds	-2.5685	-2.7680
Average of the transaction costs over the indices	-4.8982	-4.1292
HFRI Convertible Arbitrage Index	-2.9748	-2.3503
HFRI Distressed Securities Index	-3.7409	-3.1175
HFRI Emerging Markets (Total)	-10.409	-11.231
HFRI Equity Hedge Index	-5.2928	-5.5529
HFRI Equity Market Neutral Index	-1.9814	-1.8804
HFRI Equity Non-Hedge Index	-7.6039	-7.7172
HFRI Event-Driven Index	-3.7228	-3.3989
HFRI Fixed Income (Total)	-2.8376	-2.2500
HFRI Fixed Income: Arbitrage Index	-6.1764	-4.3318
HFRI Fixed Income: High Yield Index	-6.4438	-3.6841
HFRI FOF: Conservative Index	-2.4110	-2.1042
HFRI FOF: Diversified Index	-4.7314	-4.0279
HFRI FOF: Market Defensive Index	-3.6750	-2.8050
HFRI FOF: Strategic Index	-6.2475	-6.1420
HFRI FOF Composite Index	-3.7260	-3.9430
HFRI FOF Composite Index (Off.)	-4.6198	-4.6972
HFRI Fund Weighted Composite Index	-4.2733	-4.2082
HFRI Macro Index	-3.3393	-3.5459
HFRI Merger Arbitrage Index	-3.9681	-2.8237
HFRI Regulation D Index	-3.5011	-3.7099
HFRI Relative Value Arbitrage Index	-2.4469	-1.7284
HFRI Short Selling Index	-19.302	-17.595
Average of the transaction costs over the indices	-5.1557	-4.6747

Table 15: Hedging errors (basis per points) of the EDHEC and HFRI indices for each of two reserve assets over the entire period (1997–2006).

Fund	Hedging error	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	-5.022966343	2.689724779
EDHEC-CTA Global	-8.058744042	5.645421806
EDHEC-Distressed Securities	4.124754378	19.47155871
EDHEC-Emerging Markets	-11.21163859	13.259774
EDHEC-Equity Market Neutral	-1.471590683	1.56198415
EDHEC-Event Driven	-3.020763751	6.22406221
EDHEC-Fixed Income Arbitrage	-5.177575949	3.189905767
EDHEC-Global Macro	-4.053867497	4.395207
EDHEC-Long/Short Equity	4.47809413	3.734220311
EDHEC-Merger Arbitrage	-3.442046302	2.242736202
EDHEC-Relative Value	-1.10554998	2.836227619
EDHEC-Short Selling	-24.29013217	18.8506452
EDHEC-Funds of Funds	2.033494462	8.749446216
Average of the hedging errors over the indices	-4.324502488	7.142377997
HFRI Convertible Arbitrage Index	-4.675913708	2.609102503
HFRI Distressed Securities Index	3.722398591	16.58984332
HFRI Emerging Markets (Total)	7.097564556	12.66959323
HFRI Equity Hedge Index	-1.643622346	11.19037495
HFRI Equity Market Neutral Index	-2.258515275	2.472596466
HFRI Equity Non-Hedge Index	7.453328183	4.603254198
HFRI Event-Driven Index	2.862294451	12.29110626
HFRI Fixed Income (Total)	-2.603139406	2.402853856
HFRI Fixed Income: Arbitrage Index	-4.087640896	4.977681096
HFRI Fixed Income: High Yield Index	2.638073684	2.387582196
HFRI FOF: Conservative Index	-2.598299696	2.863947585
HFRI FOF: Diversified Index	-5.7248332	-2.263005293
HFRI FOF: Market Defensive Index	-7.19063663	3.850690389
HFRI FOF: Strategic Index	-8.584197214	7.510126485
HFRI FOF Composite Index	-4.800243375	3.856993799
HFRI FOF Composite Index (Off.)	-7.540482923	5.850515308
HFRI Fund Weighted Composite Index	2.244013529	15.42135204
HFRI Macro Index	-2.491140954	7.274298256
HFRI Merger Arbitrage Index	-3.635490602	2.457954561
HFRI Regulation D Index	-4.354249204	3.999422953
HFRI Relative Value Arbitrage Index	-1.757126584	4.245628798
HFRI Short Selling Index	-30.41350326	21.98728382
Average of the hedging errors over the indices	-3.106425558	6.989782153

Table 16: Hedging errors (basis per points) of the EDHEC and HFRI indices for each of two reserve assets over the first sub-period (1997–2001).

Fund	Hedging error	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	-5.854414114	7.331726159
EDHEC-CTA Global	-3.261304874	15.48344278
EDHEC-Distressed Securities	-10.63141111	18.21688996
EDHEC-Emerging Markets	-41.58467617	10.40839934
EDHEC-Equity Market Neutral	0.216747837	4.117171088
EDHEC-Event Driven	5.530304616	15.02572238
EDHEC-Fixed Income Arbitrage	-10.48685482	12.72732957
EDHEC-Global Macro	-1.950399253	10.81371999
EDHEC-Long/Short Equity	-5.472302407	8.63029379
EDHEC-Merger Arbitrage	-7.268360093	9.778517204
EDHEC-Relative Value	12.74567524	8.974747668
EDHEC-Short Selling	9.60796941	55.3754198
EDHEC-Funds of Funds	-12.45957574	9.552013774
Average of the hedging errors over the indices	-5.4514308	14.34118411
HFRI Convertible Arbitrage Index	-4.443351952	3.162705469
HFRI Distressed Securities Index	-9.790346341	17.16305189
HFRI Emerging Markets (Total)	-35.23487925	15.15013559
HFRI Equity Hedge Index	-15.50411415	10.73944888
HFRI Equity Market Neutral Index	-3.470329903	4.313327157
HFRI Equity Non-Hedge Index	-23.74524481	12.04989301
HFRI Event-Driven Index	-14.22269812	2.919363113
HFRI Fixed Income (Total)	-8.573296126	5.631050796
HFRI Fixed Income: Arbitrage Index	-4.419486285	7.911636813
HFRI Fixed Income: High Yield Index	-11.13974405	8.52434995
HFRI FOF: Conservative Index	-0.431323396	5.769353502
HFRI FOF: Diversified Index	-35.19300862	10.4684057
HFRI FOF: Market Defensive Index	-10.3549352	11.11226975
HFRI FOF: Strategic Index	-12.24470309	11.52847099
HFRI FOF Composite Index	-7.634859635	9.191341753
HFRI FOF Composite Index (Off.)	-9.434687073	11.45793373
HFRI Fund Weighted Composite Index	-24.92998374	8.097748863
HFRI Macro Index	3.853207823	15.60074834
HFRI Merger Arbitrage Index	-3.887143693	5.073019858
HFRI Regulation D Index	-2.431374814	13.42772207
HFRI Relative Value Arbitrage Index	-12.36688683	-0.033279555
HFRI Short Selling Index	-11.04287744	4.216592611
Average of the hedging errors over the indices	-11.66554849	8.794331377

Table 17: Hedging errors (basis per points) of the EDHEC and HFRI indices for each of two reserve assets over the second sub-period (2002–2006).

Fund	Hedging error	
	Reserve 1	Reserve 2
EDHEC-Convertible Arbitrage	-0.216644211	14.71418971
EDHEC-CTA Global	-6.582901453	36.52509979
EDHEC-Distressed Securities	-1.016305328	13.72616412
EDHEC-Emerging Markets	-10.13043018	31.53349596
EDHEC-Equity Market Neutral	-1.061827125	5.534718588
EDHEC-Event Driven	6.84618159	12.12860342
EDHEC-Fixed Income Arbitrage	-0.469146864	9.667485841
EDHEC-Global Macro	-0.423280278	15.97182143
EDHEC-Long/Short Equity	-4.781157282	11.93987295
EDHEC-Merger Arbitrage	2.432702745	11.46724918
EDHEC-Relative Value	1.392743596	6.816342792
EDHEC-Short Selling	8.422391583	44.59416583
EDHEC-Funds of Funds	-0.055444422	12.87295253
Average of the hedging errors over the indices	-0.434085972	17.49939709
HFRI Convertible Arbitrage Index	0.262326757	12.38729712
HFRI Distressed Securities Index	0.409148738	14.75409472
HFRI Emerging Markets (Total)	-9.733473043	18.61313588
HFRI Equity Hedge Index	-7.44974449	20.13641716
HFRI Equity Market Neutral Index	-0.864838511	7.778622358
HFRI Equity Non-Hedge Index	-12.1917044	22.30909531
HFRI Event-Driven Index	2.282364634	16.12105281
HFRI Fixed Income (Total)	0.028468682	4.514146651
HFRI Fixed Income: Arbitrage Index	-0.000830608	8.164211829
HFRI Fixed Income: High Yield Index	0.197970856	10.94636336
HFRI FOF: Conservative Index	-3.168935239	6.353466396
HFRI FOF: Diversified Index	-0.756117028	12.47152293
HFRI FOF: Market Defensive Index	-4.356542614	12.3839458
HFRI FOF: Strategic Index	-5.764923859	19.62428265
HFRI FOF Composite Index	-0.639004576	8.640850198
HFRI FOF Composite Index (Off.)	-0.945047981	9.668544712
HFRI Fund Weighted Composite Index	-4.808300645	13.62398725
HFRI Macro Index	1.031221433	30.93949304
HFRI Merger Arbitrage Index	2.456653636	12.85418557
HFRI Regulation D Index	0.291293742	28.74394205
HFRI Relative Value Arbitrage Index	0.400473588	7.062009818
HFRI Short Selling Index	7.951857962	38.58089671
Average of the hedging errors over the indices	-1.607621953	15.30325292