Testing non-linear dependence and market timing ability. An application to the Hedge Fund industry*

Javier Mencía  
CEMFI and UPNA  
&lt;mencia@cemfi.es&gt;  

January 2005  
Revised: April 2006

Abstract  
We show that managers who use market timing signals can generate returns with strong non-linear exposure to the market. When those signals are biased, this effect can be highly harmful for investors. We propose the dependence structures implied by the Student $t$ and the Generalised Hyperbolic distributions to model these features. Specification tests for linear and Student $t$ models are derived and used to analyse the dependence of hedge fund returns on market conditions. We find a significant presence of non-linear dependence and market mis-timing effects, which tends to worsen the performance of hedge funds during episodes of distressed markets.

Keywords: Non-linear dependence, Market timing, Hedge Funds, Correlation, Asymmetry.

JEL: C12,G11,C32,C22

*I am specially indebted to Enrique Sentana, for his advice, comments and suggestions. I am also grateful to Manuel Arellano, Tim Bollerslev, Antonio Díez de los Ríos, Eva Ferreira, Andrew Patton and participants at seminars at CEMFI, CORE, European Central Bank, Banco de España, Duke University, Universidad Carlos III, Universidad de Navarra, Universitat Pompeu Fabra, the XIII Finance Forum (Madrid) and the XXX Symposium on Economic Analysis (Murcia), for helpful comments and valuable insights. Of course, all remaining errors are entirely mine. Financial support from Fundación Ramón Areces is gratefully acknowledged. Address for correspondence: Casado del Alisal 5, E-28014 Madrid, Spain, tel: +34 91 429 05 51, fax: +34 91 429 1056.


1 Introduction

Understanding the risk exposures of financial assets is crucial for investors. They need to know the risks that affect asset returns in order to decide on their portfolio allocations. To measure these risks, the traditional approach has been to estimate a linear factor model with respect to a set of market variables. This has been partly due to the simplicity and intuitiveness of this methodology, but also to the fact that, under joint normality, linear models give a complete characterisation of the dependence between variables. In some cases, such models have been shown to work extremely well. For instance, a high proportion of mutual fund risks can be explained using a linear model with just a few risk factors (see Sharpe, 1992). However, this methodology can be far from fully informative if the dependence between variables goes beyond the simple linear relationship that Gaussianity imposes.

Hedge funds are probably the assets in which the presence of non-linearities can be more dangerous. Hedge fund managers try to exploit investment opportunities with highly leveraged dynamic strategies in order to provide high returns regardless of market conditions. In this sense, Fung and Hsieh (1997) have found that linear models of market factors are only able to explain a minor proportion of returns in this industry. As a result, traditional performance and risk measures, such as the Sharpe ratio, Jensen’s alpha, or the market’s beta, might induce a naive investor to invest unreasonably high proportions of her wealth in hedge funds (see e.g. Lhabitant, 2004). However, this picture can be quite misleading in the presence of non-linear dependence. In this sense, Fung and Hsieh (2001) have found that lookback straddles provide a useful characterisation of the returns of trend following funds. Similarly, Agarwal and Naik (2004) find that a large number of equity oriented hedge funds exhibit payoffs that resemble short positions on put options on the market index. In addition, Lo (2001) has also noted the “phase-locking” behaviour of hedge funds. He uses this term to denote those situations in which otherwise uncorrelated assets suddenly become almost perfectly correlated. One of the most famous examples of this phenomenon is the general debacle that followed the Russian debt restructuring in August 1998, with LTCM as the most prominent victim.

Hence, hedge funds market exposure seems to be higher during episodes of distressed markets, which are precisely the situations most feared by investors. In consequence, this phenomenon can have a profound impact on risk management decisions. Linear
dependence (i.e. a pure correlation effect), can be easily hedged with a static strategy.\footnote{In the terminology of the CAPM, we only need to know the beta of the portfolio in this case} However, a dependence with the market beyond what is captured by linear correlation can make optimal allocation a much more difficult task (see Patton, 2004).

In this paper, we show that market timing managers can generate returns with a strong non-linear exposure to market factors. This effect can be highly beneficial if managers have valuable private information about the market. Unfortunately, if they are not able to correctly time the market, this non-linear exposure can be strongly harmful for investors. Our approach is closely related to those of Glosten and Jagannathan (1994), Mitchell and Pulvino (2001) and Díez de los Ríos and García (2003). These authors test the empirical relevance of the market timing model of Merton (1981). Nevertheless, our starting point is the timing model of Admati, Bhattacharya, Pfleiderer, and Ross (1986), for which we provide a multivariate generalisation. One of the main advantages of our approach is that it can be related to the most popular statistical distributions in the literature. In particular, we derive specification tests for the linear factor model (Gaussian dependence) and the non-linear dependence structure imposed by the Student $t$ distribution. Our tests use the dependence structure introduced by the Generalised Hyperbolic (GH) as our alternative hypothesis. This is a flexible family of multivariate asymmetric and leptokurtic distributions that nests the multivariate Gaussian and Student $t$ distributions as particular cases. We will show that this distribution can model market timing effects with high flexibility. We will exploit these features to derive one-sided versions of our tests that consider market timing alternatives.

We are mainly interested in dependence in this paper. Hence, we will rely on the Gaussian pseudo-Maximum likelihood methodology developed by Bollerslev and Wooldridge (1992) among others, to ensure that our tests do not assume any particular distribution about the marginals. Furthermore, although the tests are derived with a GH alternative hypothesis in mind, they can be easily interpreted as tests of intuitive moment conditions. Hence, they will also yield non-trivial power against other alternatives that introduce non-linear dependence. We assess this issue, as well as the finite sample properties of the tests with a Monte Carlo exercise.

Our tests can be a useful tool in order to decide on the model of dependence that is best suited to correctly account for the risk-return characteristics of a specific set
of assets. In particular, we will analyse the dependence structure between hedge fund returns and the market, defined as the S&P 500 index and the FT 100. Using monthly returns on around 1,500 hedge funds, we find that Gaussian dependence is rejected for a significant number of funds. Although some of these rejections can be attributed to market timing ability, inability to time the market seems to be more common. We also consider the dependence in this industry as a whole by testing the dependence between the 10 CSBF-Tremont main style indices and the market, and we can also easily reject the null of linear dependence. In contrast, Student $t$ dependence provides a much better empirical fit to the data. Finally, we investigate the impact of non-linear dependence on the diversification offered by hedge funds. Our results show that it improves their performance in normal times. Unfortunately, when strong negative shocks to the market occur, their diversification benefits seem to deteriorate due to non-linear dependence.

The rest of the paper is organised as follows. The next section discusses the effects of market timing ability on hedge fund returns. In section 3 we give a brief introduction to the GH distribution. Section 4 focuses on testing Gaussian dependence, while in section 5 we derive a test of Student $t$ dependence. We analyse dependence in the hedge fund industry in section 6. Finally, concluding remarks and directions for future research are suggested in section 7. Proofs and auxiliary results can be found in the appendix.

2 Market timing and non-linear dependence

We will start with an economy that satisfies a linear factor model. Then, we will show that hedge fund managers that use market timing signals generate returns that cannot be explained by a linear factor model. Consider a world with a safe and $A$ risky assets whose net returns are $r$ and the $A$–dimensional vector $y_a$, respectively. We will assume that none of these assets is redundant. There are also $k \leq A$ market factors in this economy, denoted by $y_m$. By assumption, these factors are either directly traded or can be spanned by the set of traded assets. In addition, we will assume that

\[
\begin{pmatrix}
y_m \\
y_a
\end{pmatrix} \sim N\left[\begin{pmatrix}
\mu_m \\
\mu_a
\end{pmatrix}, \begin{pmatrix}
\Sigma_{mm} & \Sigma_{am} \\
\Sigma_{am}^\prime & \Sigma_{aa}
\end{pmatrix}\right]
\]

Hence, we can express $y_a$ in terms of the following linear factor model:

\[
y_a = \mu_a + \Sigma_{am} \Sigma_{mm}^{-1} (y_m - \mu_m) + \varepsilon_{a|m}
\]
where $\varepsilon_{a|m} \sim N(0, \Sigma_{a|m})$ is independent of $y_m$, and $\Sigma_{a|m} = \Sigma_{aa} - \Sigma_{am}\Sigma_{mm}^{-1}\Sigma_{am}'$. The implicit assumption for $y_m$ to be considered a meaningful set of factors is that the residuals from this model, $\varepsilon_{a|m}$, should only reflect idiosyncratic or diversifiable risk. However, no explicit restriction will be imposed.

There are two types of hedge fund managers in this economy: those who use market timing signals to allocate the capital of their funds and those who do not.\(^2\) Both types are CARA utility maximisers:

$$U(w) = -\exp[-\alpha [(1 - w'\iota_A) r + w'y_a]]$$

where $\alpha$ is the coefficient of absolute risk aversion, initial capital is normalised to unity, $w$ is the vector of allocations to the risky assets and $\iota_A$ is a vector of $A$ ones. Market timing managers observe the signal

$$s = l(y_m - \mu_m) + \varepsilon_{s|m}$$

(1)

where $\varepsilon_{s|m} \sim N(0, \lambda \Sigma_{mm})$ is independent of $y_a$, and $\lambda > 0$. This type of signal is a multivariate generalisation of the signal considered by Admati, Bhattacharya, Pfleiderer, and Ross (1986). Jointly Gaussian signals are also considered by Sentana (2005). However, we additionally consider the possibility that the beliefs of the managers about the signals are wrong. In particular, the managers believe that $l = 1$ (unbiased), but in reality it might be $-1$ (biased) or even 0 (noise).

First, let us consider the optimal hedge fund return obtained by managers that do not observe (1). In this case, the optimal portfolio can be written as $y^0_h = (1 - w^0'y_A) r + w^0'y_a$ where $w^0 = \alpha^{-1}\Sigma_{aa}^{-1}(\mu_a - r\iota_A)$. Thus, $y^0_h$ still satisfies a linear factor dependence structure with respect to $y_m$. Indeed, the mean of $y^0_h$ conditional on the market variables is linear, while the conditional variance is constant:

$$E[y^0_h | y_m] = r + \alpha^{-1}(\mu_a - r\iota_A)'\Sigma_{aa}^{-1}(\mu_a - r\iota_A)$$

$$+ \alpha^{-1}(\mu_a - r\iota_A)'\Sigma_{aa}^{-1}\Sigma_{am}\Sigma_{mm}^{-1}(y_m - \mu_m)$$

$$V[y^0_h | y_m] = \alpha^{-2}(\mu_a - r\iota_A)'\Sigma_{aa}^{-1}\Sigma_{a|m}\Sigma_{aa}^{-1}(\mu_a - r\iota_A)$$

In consequence, in this simple setting the linear factor model is able to capture the real exposure of hedge funds to market factors. For instance, this kind of dependence should

\(^2\)The model has an infinite number of investors, so individual agents, and in particular managers, cannot influence assets prices.
be found in market neutral funds, since their managers are not supposed to time the market.

Now, let us consider the returns generated by the more sophisticated market timing managers. Their asset allocations will be based on the observed signals, under the assumption that they are unbiased:

\[ w^t = \alpha^{-1} V^{-1} (y_a | s, l = 1) \left[ E (y_a | s, l = 1) - r_{\lambda A} \right] \]

The returns from these funds will satisfy the following properties:

**Proposition 1**

\[
E \left[ y_h^t | y_m, l \right] = c_0 + c_1'(l) (y_m - \mu_m) + c_2(l) \varsigma_{mt} \\
V \left[ y_h^t | y_m, l \right] = d_0 + d_1'(y_m - \mu_m) + d_2 \varsigma_{mt}
\]

where

\[
\varsigma_{mt} = (y_m - \mu_m)' \Sigma_{mm}^{-1} (y_m - \mu_m)
\]

\[
c_0 = r + \alpha^{-1} (R\mu_a - r_{\lambda A-k})' (R \Sigma_{aa} R')^{-1} (R\mu_a - r_{\lambda A-k}) \\
+ \alpha^{-1} \lambda^{-1} (1 + \lambda) (\mu_m - r_{\lambda k})' \Sigma_{mm}^{-1} (\mu_m - r_{\lambda k})
\]

\[
c_1(l) = \alpha^{-1} \lambda^{-1} (1 + \lambda + l) \Sigma_{mm}^{-1} (\mu_m - r_{\lambda k})
\]

\[
c_2(l) = \alpha^{-1} \lambda^{-1} l
\]

\[
d_0 = \alpha^{-2} (R\mu_a - r_{\lambda A-k})' (R \Sigma_{aa} R')^{-1} (R\mu_a - r_{\lambda A-k}) \\
+ \alpha^{-2} \lambda^{-2} (\mu_m - r_{\lambda k})' \Sigma_{mm}^{-1} (\mu_m - r_{\lambda k})
\]

\[
d_1 = 2 \alpha^{-2} \lambda^{-1} \Sigma_{mm}^{-1} (\mu_m - r_{\lambda k})
\]

\[
d_2 = \alpha^{-2} \lambda^{-1}
\]

where \( R \) is a \((A - k) \times A\) matrix of rank \( A - k \), such that \( R \Sigma_{am} = 0 \).

In consequence, now the conditional mean and variance are quadratic functions of the market. The vector \( \mu_m - r_{\lambda k} \) should be strictly positive, so that the market is demanded in equilibrium. Hence, the coefficients \( d_1 \) and \( d_2 \) should be positive. However, the sign of the quadratic term in the conditional mean will depend on whether the signal is unbiased in reality or not. In particular, if \( l = 1 \), \( c_2(l) \) will be positive and non-linear exposure will have a beneficial effect on the expected return of the fund. Unfortunately, if the signal is biased \( l = -1 \), its effect might be very harmful, since \( c_2(l) \) will be negative. Finally, when the signal is pure noise \( l = 0 \), the conditional mean will be linear, but the conditional variance will still be a quadratic function of the market. This feature might also deteriorate the performance of this fund, by increasing its variance.
when there are strong shocks to the market. Hence, if we want to analyse the exposure of hedge fund returns to the market, a linear factor model is not able to capture the true dependence structure when market timing information is used. Even though investors might implement strategies to reduce the linear component in the conditional mean, the remaining terms cannot be so easily dealt with.

Now, assume that we have data on the returns of \( N \) hedge funds, denoted by \( y_h \). We are interested in assessing the exposure of these returns to the \( k \) market indices \( y_m \). We will assume that the unconditional distribution of our assets is characterised by the following mean and variance\(^3\)

\[
E \begin{bmatrix} \mathbf{y}_m \\ \mathbf{y}_h \end{bmatrix} = \begin{bmatrix} \mu_m \\ \mu_h \end{bmatrix}; \quad V \begin{bmatrix} \mathbf{y}_m \\ \mathbf{y}_h \end{bmatrix} = \begin{bmatrix} \Sigma_{mm} & \Sigma'_{hm} \\ \Sigma_{hm} & \Sigma_{hh} \end{bmatrix}.
\]

We will tackle this problem by modelling \( E (y_h | y_m) \) and \( V (y_h | y_m) \). However, as we have learnt in this section, the shape of these moments might be highly non-linear, depending on the investment strategy followed by the hedge fund managers. Hence, we will need to consider several degrees of dependence. We will start with Gaussian dependence, under which we can completely characterise the relationship between the alternative investments and the market by means of the linear regression:

\[
y_h = \mu_h + \Sigma_{hm} \Sigma_{mm}^{-1} (y_m - \mu_m) + \varepsilon_{h|m} \quad (2)
\]

where the residual \( \varepsilon_{h|m} \) is independent of \( y_m \). Thus, in this case we can write:

\[
E (y_h | y_m) = \mu_h + \Sigma_{hm} \Sigma_{mm}^{-1} (y_m - \mu_m) \quad (3)
\]

\[
V (y_h | y_m) = \Sigma_{h|m} = \Sigma_{hh} - \Sigma_{hm} \Sigma_{mm}^{-1} \Sigma'_{hm}
\]

This type of dependence, consistent with the absence of market timing signals, is completely characterised by a linear combination of the market factors. Therefore, we will also denote it as “linear dependence”.

Let us now turn to the more general case of Student \( t \) dependence. Now the conditional mean is still given by (39), but the conditional variance is a quadratic function of the market:

\[
V (y_h | y_m) = \frac{1 - 2\eta + \eta \Sigma_{mm}}{1 + \eta (k - 2)} \Sigma_{h|m} \quad (4)
\]

\(^3\)For expositional simplicity, these moments will be assumed to be constant for the moment. More general parametrisations will be considered in the following sections.
where $s_{mt} = (y_m - \mu_m)' \Sigma_{mm}^{-1} (y_m - \mu_m)$ and $\eta > 0$. Notice that this type of market exposure can be generated by market timing managers who observe uninformative signals. Unfortunately, (4) implies that the risk in these investments will be higher under conditions of stressed markets. Those situations are characterised by extreme moves of $y_m$ from its expected return, which imply high values of $s_{mt}$.

Furthermore, we may also want to account for non-linear effects of the market in the conditional mean, as well as for asymmetric features in the conditional variance. In our case, we will model these effects with the dependence structure of the Generalised Hyperbolic distribution. Interestingly, we will see in the paper that this distribution can model the market timing and mis-timing effects in the conditional mean and variance that we have studied at the beginning of this section.

### 3 The Generalised Hyperbolic distribution

The Generalised Hyperbolic distribution is a family of multivariate asymmetric and leptokurtic distributions, with the Gaussian and the Student $t$ distributions as particular cases. In general, this family induces tail dependence between variables, which can be flexibly modelled by appropriately choosing the parameters of the distribution. The source of this dependence can be understood by considering the location-scale mixture of normals interpretation of the GH distribution. Let $u$ be a standardised $N$-dimensional GH random vector, then $u$ can always be written as

$$u = \alpha(\beta, \eta, \psi) + \Upsilon(\beta, \eta, \psi)\beta\xi^{-1} + \xi^{-1/2} \Upsilon^1(\beta, \eta, \psi) r$$

(5)

where $\Upsilon(\beta, \eta, \psi) = \psi^{-1} + \sqrt{1 + 4[D_1 - 5\eta^{-1} - 1]β'β}/2[D_1 - 5\eta^{-1} - 1]β'β$, $D_\nu(x) = K_{\nu+1}(x)K_\nu(x)/K^2_{\nu+1}(x)$, $R_\nu(x) = K_{\nu+1}(x)/K_\nu(x)$, $\beta \in \mathbb{R}^N$, $\eta \in \mathbb{R}$, $\psi \in [0,1]$ and $K_\nu(\cdot)$ is the modified Bessel function of the third kind (see Abramowitz and Stegun, 1965). In addition, $r \sim N(0, I_N)$, and the positive mixing variable $\xi$ is an independent Generalised Inverse Gaussian (GIG) variate with parameters $.5\eta^{-1}$, $\psi^{-1} - 1$ and 1, or $\xi \sim GIG (.5\eta^{-1}, \psi^{-1} - 1, 1)$ for short (see Jørgensen,
Tail dependence is obtained because very small values of \( \xi \) produce extreme realisations in all the components of \( \mathbf{u} \). Additionally, through the vector \( \mathbf{\beta} \), we can make the transmission of this effect stronger in certain directions, thus introducing asymmetric tail dependence.

For the purposes of this paper, it is also important to note that, if we consider two subgroups of variables in \( \mathbf{u} \), say \( \mathbf{u}_1 \) and \( \mathbf{u}_2 \), then the conditional distribution of \( \mathbf{u}_1 \) given \( \mathbf{u}_2 \) is also GH. This result, first proved by Blæsild (1981), will be fundamental to develop the tests of this paper. We will specifically use the mean and variance of this conditional distribution to describe \( E[\mathbf{y}_h | \mathbf{y}_m] \) and \( V[\mathbf{y}_h | \mathbf{y}_m] \) (see appendix B). In this sense, we will rely on the dependence structure of the GH distribution to model the conditional mean and variance, and it can be shown that when either \( \eta \) or \( \psi \) are zero, then we obtain a Gaussian dependence structure, regardless of the value of \( \mathbf{\beta} \). Additionally, for \( \eta > 0 \), \( \psi = 1 \) and \( \mathbf{\beta} = \mathbf{0} \), the relationship between \( \mathbf{y}_h \) and \( \mathbf{y}_m \) is of Student \( t \) type, with \( 1/\eta \) being the degrees of freedom of this distribution (see 4). Finally, when \( \mathbf{\beta} \neq \mathbf{0} \), this distribution induces non-linear asymmetric features in both the conditional mean and variance.

As an illustration, we have plotted in figures 1a and 1b the shape of the conditional mean and variance when \( N = k = 1 \) and \( \mathbf{y}_m \) and \( \mathbf{y}_h \) are uncorrelated with zero unconditional mean and unit variance. Due to the absence of correlation, the Gaussian dependence structure is characterised by constant conditional mean and variance. The Student \( t \) dependence example is characterised by a quadratic conditional variance, which is consistent with a pure noise market timing signal. We have also plotted two examples in which the vector \( \mathbf{\beta} \) is not identically zero to illustrate the flexibility of the GH distribution to model non-linear effects. In the first case, the conditional mean decreases as the market goes either up or down. What is worse, the conditional variance shows that the uncertainty is reduced as the market moves down. Therefore, this case is consistent with inability to time the market. Finally, the last example, where non-linear dependence has always a positive effect on the expected return of \( y_h \), illustrates the case of market timing ability. Of course, these are just two examples of the flexibility of this distribution and with other values of \( \mathbf{\beta} \), \( \psi \) and \( \eta \), we could model other interesting patterns.

\(^4\)\( \alpha(\mathbf{\beta}, \eta, \psi) \) and \( \mathbf{\Upsilon}(\mathbf{\beta}, \eta, \psi) \) are, respectively, a vector and a matrix of location and scale parameters which are parametrised in terms of \( \mathbf{\beta} \), \( \eta \) and \( \psi \) so that \( E(\mathbf{u}) = \mathbf{0} \) and \( V(\mathbf{u}) = \mathbf{I}_N \).
4 Testing linear dependence

In section 2, we illustrated the implications of non-linear dependence on financial returns. However, finding a good model for a specific set of assets can be costly and time consuming. Hence, it seems important to have analytical tools that can serve us as a guide in order to select the preferred one. In this section, we derive a test of the null of linear conditional mean and constant variance, as well as several interesting variants.

4.1 Testing approach

We are interested in models with possibly time varying conditional means and variances. Thus, to account for these features, we will now adopt a more general notation, that includes a time subindex in the variables names. We will assume that, conditional on past information and possibly \( p_z \) contemporaneous conditioning variables that we will denote as \( z_t \), the first two moments of the variables under study are given by:

\[
E_{t-1} \left[ \begin{array}{c} y_{mt} \\ y_{ht} \end{array} \right | \theta, z_t] = \left[ \begin{array}{c} \mu_{mt}(\theta) \\ \mu_{ht}(\theta) \end{array} \right],
\]

(6)

\[
V_{t-1} \left[ \begin{array}{c} y_{mt} \\ y_{ht} \end{array} \right | \theta, z_t] = \left[ \begin{array}{cc} \Sigma_{mmmt}(\theta) & \Sigma_{mmt}(\theta) \\ \Sigma_{hmt}(\theta) & \Sigma_{ht}(\theta) \end{array} \right],
\]

(7)

where the time subindex at the mean and variance operators indicates that we are also conditioning on the available information set at \( t-1 \). To remain as general as possible, we use a rather generic notation for the conditional mean and variance, which are functions known up to the \( p_{\theta} \times 1 \) vector of true parameter values \( \theta_0 \).

The main goal of this section is to test the null of a linear dependence between \( y_{ht} \) and \( y_{mt} \). Under the null, we can write the relationship between these variables as

\[
\varepsilon_{ht}(\theta) = \Sigma_{hmt}(\theta)\Sigma_{mmmt}^{-1}(\theta)\varepsilon_{mt}(\theta) + \varepsilon_{h|mmt}(\theta)
\]

(8)

where \( \varepsilon_{jt}(\theta) = y_{jt} - \mu_{jt}(\theta) \), for \( j = h, m \), and \( \varepsilon_{h|mmt}(\theta) \) will be assumed to be independent of \( \varepsilon_{mt}(\theta) \), given past information up to \( t-1 \) and \( z_t \). Hence, the null implies that linear correlation is a sufficient statistic of the dependence between \( y_{ht} \) and \( y_{mt} \). Or, in other words, there is no relationship between \( y_{ht} \) and \( y_{mt} \) beyond linear correlation. In particular, we will test linearity by considering the implications of the null on the first
two conditional moments, i.e.

\[
E_{t-1} [\varepsilon_{ht}(\theta) | \varepsilon_{mt}(\theta), \mathbf{z}_t; \theta] = \Sigma_{hmt}(\theta) \Sigma_{mmt}^{-1}(\theta) \varepsilon_{mt}(\theta)
\]

\[
V_{t-1} [\varepsilon_{ht}(\theta) | \varepsilon_{mt}(\theta), \mathbf{z}_t; \theta] = \Sigma_{hjmt}(\theta)
\]

where \( \Sigma_{hjmt}(\theta) = \Sigma_{hht}(\theta) - \Sigma_{hmt}(\theta) \Sigma_{mmt}^{-1}(\theta) \Sigma_{mht}(\theta) \).

In this sense, for the particular case of \( k = N = 1 \), Patton (2005) considers the following moment conditions for the mean and variance under the alternative hypothesis

\[
E(y_{ht} | y_{mt}) = \delta_0 + \delta_1 y_{mt} + \delta_2 y_{mt}^2
\]

\[
V(y_{ht} | y_{mt}) = \omega_0 + \omega_1 y_{mt} + \omega_2 y_{mt}^2
\]

To test mean neutrality of \( y_{ht} \) with respect to \( y_{mt} \), he considers the null of \( \delta_1 = \delta_2 = 0 \). Analogously, he also considers testing variance neutrality as a test of \( \omega_1 = \omega_2 = 0 \). Our linear dependence test can be seen under Patton’s framework as a test of \( \delta_2 = \omega_1 = \omega_2 = 0 \).

Nevertheless, our alternative hypothesis will be the conditional mean and variance implied by the GH distribution of \( y_{ht} \) given \( y_{mt} \). In particular, we consider a GH model with shape parameters \( \beta_t = \Sigma_t^{1/2} (\theta) \mathbf{b}, \eta \) and \( \psi \) for the orthogonalised variables \( \varepsilon_{ht}^t(\theta) = [\varepsilon_{mt}^t(\theta), \varepsilon_{hmt}^t(\theta)]' \).

From this model, we obtain the formulas for the conditional mean and variance of \( \varepsilon_{hmt}(\theta) \) given \( \varepsilon_{mt}(\theta) \) under the GH distribution (see appendix B). These are our counterparts of Patton’s moment conditions (11) and (12). In this sense, Patton’s methodology can be interpreted as a truncated Taylor expansion of the moment conditions that we will consider. Indeed, we will obtain the same result in the bivariate case \( (k = N = 1) \). However, our GH approach can be easily implemented in multivariate settings, where a wide variety of dependence structures can be obtained with just \( N + k + 2 \) parameters. Furthermore, it can easily deal with specific parametric features in the mean and variance (e.g. AR, GARCH effects).

### 4.2 Implementation

We will follow two main steps to derive the test. First, we will estimate the mean and variance parameters \( \theta \) by Gaussian Pseudo-Maximum Likelihood (PML). The PML estimator, which we will denote as \( \hat{\theta}_T \), can be obtained from

\[
\hat{\theta}_T = \arg \max_{\theta} \frac{1}{T} \sum_{t=1}^T \ell_t (y_t; \theta)
\]

\[\]following Mencía and Sentana (2005), we model \( \beta \) as a time varying function of the conditional variance and \( \mathbf{b} \), in order to make our results invariant to the choice of the square root of \( \Sigma_t^{1/2} (\theta) \).
where $\ell_t (y_t; \theta)$ is the log-likelihood of the multivariate Gaussian distribution:

$$
\ell_t (y_t; \theta) = -\frac{N + k}{2} \log (2\pi) - \frac{1}{2} \log |\Sigma_t(\theta)| \\
- \frac{1}{2} \left[ y_t - \mu_t(\theta) \right]' \Sigma_t^{-1}(\theta) \left[ y_t - \mu_t(\theta) \right]
$$

(14)

Importantly, $\hat{\theta}_T$ will be consistent even under absence of normality. As Bollerslev and Wooldridge (1992) have shown, the consistency of $\hat{\theta}_T$ only requires correct specification of the conditional mean and variance (equations 6 and 7). Now, consider the Gaussian pseudo-likelihood of $\epsilon_{ht}(\theta)$ conditional on $\epsilon_{mt}(\theta)$ under the alternative of GH dependence

$$
\ell_{bmt}(\theta, \eta, \psi, b) = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log |V[\epsilon_{ht}(\theta)| \epsilon_{mt}(\theta); \theta, \eta, \psi, b]| \\
- \frac{1}{2} \left[ \epsilon_{ht}(\theta) - E[\epsilon_{ht}(\theta)| \epsilon_{mt}(\theta); \theta, \eta, \psi, b] \right]' \\
\times V^{-1}[\epsilon_{ht}(\theta)| \epsilon_{mt}(\theta); \theta, \eta, \psi, b] \\
\times \{ \epsilon_{ht}(\theta) - E[\epsilon_{ht}(\theta)| \epsilon_{mt}(\theta); \theta, \eta, \psi, b] \},
$$

(15)

where $E[\epsilon_{ht}(\theta)| \epsilon_{mt}(\theta); \theta, \eta, \psi, b]$ and $V[\epsilon_{ht}(\theta)| \epsilon_{mt}(\theta); \theta, \eta, \psi, b]$ denote the conditional mean and variance under the GH distribution (see appendix B). We will use the first-step parameter estimates $\hat{\theta}_T$ in (15), and then focus on testing if the remaining parameters $\eta$, $\psi$ and $b$ satisfy the null by means of a Lagrange multiplier test. Note again that, since we are also adopting a PML approach for the specification of the conditional log-likelihood, our test remains valid even if the conditional distribution is not Gaussian.$^7$

As Mencía and Sentana (2005) show, when $\eta$ or $\psi$ tend to zero, normality is achieved, regardless of the value of $b$. Similarly, in this case the null of linear dependence is achieved under the same circumstances. This implies that $b$ and $\psi$ are not identified when $\eta$ tends to zero, nor $b$ and $\eta$ when $\psi$ tends to the null. Following Mencía and Sentana (2005) again, we solve this problem by computing a Lagrange multiplier test for fixed values of the unidentified parameters. Then, we can compute as an overall test statistic the supremum of the test with respect to the unidentified parameters (see Andrews, 1994, for a formal justification).

Through tedious but otherwise straightforward algebra, it can be shown that the

---

$^7$See Fiorentini and Sentana (2005) for a discussion of the efficiency of sequential approaches of this type.
pseudo-score with respect to \( \eta \) and \( \psi \) for a given \( b \) can be written under the null as
\[
s_{\eta t} (y_{ht} | y_{mt}; \theta) = \lim_{\eta \to 0^+} \frac{\partial \ell_{h|m}(\theta, \eta, \psi, b)}{\partial \eta} = b^+ v_{\eta t} (y_{ht} | y_{mt}; \theta)
\]
\[
= - \lim_{\eta \to 0^-} \frac{\partial \ell_{h|m}(\theta, \eta, \psi, b)}{\partial \eta} = 2 \lim_{v \to 0} \frac{\partial \ell_{h|m}(\theta, \eta, \psi, b)}{\partial \psi}
\]
where \( b^+ = (1, b')' \),
\[
v_{\eta t} (y_{ht} | y_{mt}; \theta) = \begin{bmatrix} v_{\eta t} (y_{ht} | y_{mt}; \theta) \\ \v_{\eta t} (y_{ht} | y_{mt}; \theta) \end{bmatrix}
\]
and
\[
v_{\eta t} (y_{ht} | y_{mt}; \theta) = \left\{ \frac{1}{2} \left[ s_{m|m}(\theta) - N \right] \left[ s_{m|m}(\theta) - k \right] \right\}
\]
\[
v_{\eta t} (y_{ht} | y_{mt}; \theta) = \left[ s_{m|m}(\theta) - k \right] \v_{h|m}(\theta)
\]
with \( s_{m|m}(\theta) = e'_{m|m}(\theta) \Sigma_{m|m}^{-1}(\theta) e_{m|m}(\theta) \) and \( e_{h|m}(\theta) = e_{h|m}(\theta) \Sigma_{h|m}^{-1}(\theta) e_{h|m}(\theta) \). Further, we can show that the asymptotic distribution of the pseudo-score can be expressed as follows:

**Proposition 2** Let \( \hat{\theta}_T \) be the quasi-maximum likelihood estimator of \( \theta \), defined in (13). Then,
\[
\sqrt{T} \sum_t v_{\eta t} (y_{ht} | y_{mt}; \hat{\theta}_T) \overset{d}{\to} N [0, F(\theta_0)]
\]
where
\[
F(\theta_0) = B_{y\eta} (\theta_0) + A_{\theta\eta} (\theta_0) A_{\theta\eta}^{-1} (\theta_0) B_{y\theta} (\theta_0) A_{\theta\theta}^{-1} (\theta_0) A_{\theta\eta} (\theta_0)
\]
\[
B(\theta_0) = \begin{bmatrix} B_{y\theta} (\theta_0) & B_{y\theta} (\theta_0) \\ B_{\eta\theta} (\theta_0) & B_{\eta\eta} (\theta_0) \end{bmatrix}
\]
\[
= \lim_{T \to \infty} V \left\{ \frac{\sqrt{T}}{T} \sum_t \begin{bmatrix} s_{\theta t} (y_t; \theta) \\ v_{\eta t} (y_{ht} | y_{mt}; \theta) \end{bmatrix} \right\},
\]
\[
A_{\theta\eta} (\theta_0) = p \lim_{T \to \infty} \frac{1}{T} \sum_t A_{\theta\eta} (\theta_0),
\]
\[
A_{\theta\theta} (\theta_0) = p \lim_{T \to \infty} \frac{1}{T} \sum_t A_{\theta\theta} (\theta_0),
\]
\[
A_{\theta\theta} (\theta) = \frac{\partial \mu (\theta)}{\partial \sigma (\theta)} \Sigma^{-1}(\theta) \frac{\partial \mu (\theta)}{\partial \theta}
\]
\[
+ \frac{1}{2} \frac{\partial \text{vec}^T(\Sigma (\theta))}{\partial \theta} \left[ \Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta) \right] \frac{\partial \text{vec}(\sigma (\theta))}{\partial \theta},
\]
and
\[
A_{\theta\eta} (\theta) = \begin{bmatrix} \frac{\partial \text{vec}^T(\Sigma (\theta))}{\partial \theta} \left[ \Sigma^{-1}(\theta) \otimes I_N \right] \\ -\frac{\partial \text{vec}^T(\Sigma (\theta))}{\partial \theta} \left[ \Sigma^{-1}(\theta) \otimes \Sigma^{-1}(\theta) \Sigma (\theta) \right] \end{bmatrix}
\]
\[
\times s_{m|m}(\theta) \left[ e_{m|m}(\theta) \otimes I_N \right] L_h
\]
\( L_h = [0_{N \times k+1}, I_N] \) and \( s_{\theta t} (y_t; \theta) \) is the score of the Gaussian log-likelihood (14).
Hence, our proposed test statistic for given \( b \) will be
\[
\tau_d(\hat{\theta}_T, b) = \frac{b^+ \sqrt{T} \sum_{t} v_{qt}(y_{ht}|y_{mt}; \hat{\theta}_T)}{b^+ \mathcal{F}(\theta_0) b^+}
\]
which is asymptotically distributed as a chi-square with one degree of freedom. By maximising \( \tau_d(\hat{\theta}_T, b) \) with respect to \( b \), we find that the supremum test

**Proposition 3**
\[
\sup_b \tau_d(\hat{\theta}_T, b) = \left[ \frac{\sqrt{T}}{T} \sum_t v_{qt}(y_{ht}|y_{mt}; \hat{\theta}_T) \right] \times \mathcal{F}^{-1}(\theta_0) \left[ \frac{\sqrt{T}}{T} \sum_t v_{qt}(y_{ht}|y_{mt}; \hat{\theta}_T) \right]
\]
converges in distribution to a \( \chi^2_{N+k+1} \).

We can see that, both in (21) and (22), the vector \( v_{qt}(y_{ht}|y_{mt}; \hat{\theta}_T) \) plays a fundamental role. In particular, the supremum test can be interpreted as a moment test of the conditions given in this vector,
\[
E \left[ v_{qt}(y_{ht}|y_{mt}; \theta_0) \right] = 0
\]
These moment conditions consist of three covariances between \( \epsilon_{mt}(\theta) \) and the residual of the regression (8), which under the null must all be zero. We can better understand the test by considering the bivariate case. For \( k = N = 1 \), it is straightforward to show that the moment conditions in (23) can be written as
\[
E \left[ E_{t-1} \begin{bmatrix} H_2(\epsilon^*_{h|mt})H_2(\epsilon^*_{mt}) \\ \sigma_{mt}\sqrt{2}H_1(\epsilon^*_{mt})H_2(\epsilon^*_{h|mt}) \\ \sigma_{h|mt}\sqrt{2}H_1(\epsilon^*_{h|mt})H_2(\epsilon^*_{mt}) \end{bmatrix} \right] = E \begin{bmatrix} 0.5\text{cov}_{t-1}(\epsilon^*_{h|mt}, \epsilon^*_{mt}) \\ \sigma_{mt}\text{cov}_{t-1}(\epsilon^*_{mt}, \epsilon^*_{h|mt}) \\ \sigma_{h|mt}\text{cov}_{t-1}(\epsilon^*_{h|mt}, \epsilon^*_{mt}) \end{bmatrix}
\]
where \( H_k(\cdot) \) denotes the standardised Hermite polynomial of order \( k \) (see Stuart and Ord, 1977), and \( \epsilon^*_{h|mt} = \epsilon_{h|mt}/\sigma_{h|mt} \), while \( \epsilon^*_{mt} = \epsilon_{mt}/\sigma_{mt} \). Hence, these moment conditions are the co-kurtosis and co-skewness between \( \epsilon^*_{h|mt} \) and \( \epsilon^*_{mt} \). Intuitively, we are testing if, once we have removed linear dependence, non-linear exposure still remains in \( \epsilon^*_{h|mt} \).

Now, consider the moment conditions in the normality test of Mencía and Sentana (2005). In the bivariate case, these conditions can be expressed as
\[
\sqrt{3/2} \left[ H_4(\epsilon^*_{h|mt}) + H_4(\epsilon^*_{mt}) \right] + H_2(\epsilon^*_{h|mt}) H_2(\epsilon^*_{mt})
\]
\[
\sigma_{mt} \left[ \sqrt{6}H_3(\epsilon^*_{mt}) + \sqrt{2}H_1(\epsilon^*_{mt}) H_2(\epsilon^*_{h|mt}) \right]
\]
\[
\sigma_{h|mt} \left[ \sqrt{6}H_3(\epsilon^*_{h|mt}) + \sqrt{2}H_1(\epsilon^*_{h|mt}) H_2(\epsilon^*_{mt}) \right]
\]

\(^8\)For ease of notation, we will not specify the dependence on \( \theta \) in this discussion.
Equation (25) is the kurtosis component of the normality test, whereas (26) and (27) are the skewness components. If we compare (24) with these equations, we can see that the linearity test retains from the normality test only those moments that are related to co-skewness and co-kurtosis, but not those that deal with the skewness and kurtosis of each variable. In other words, the test is not using the conditions that imply making a particular assumption about the marginal distributions.

In general, for \( k \) and/or \( N \) greater than one, we obtain a generalised multivariate measure of co-kurtosis and co-skewness. It may be noticed that, although we have assumed a GH dependence structure under the alternative, we have obtained easily interpretable moment conditions. In fact, it can be shown that if we consider as our alternative hypothesis

\[
E(\mathbf{y}_{ht}|\mathbf{y}_{mt}) = \mathbf{\mu}_h + \Sigma_{hm}\Sigma_{mm}^{-1}(\mathbf{y}_{mt} - \mathbf{\mu}_m) + \delta_2(s_{mt} - k),
\]

\[
V(\mathbf{y}_{ht}|\mathbf{y}_{mt}) = \Sigma_{hm}[1 + \omega_1'(\mathbf{y}_{mt} - \mathbf{\mu}_m) + \omega_2(s_{mt} - k)],
\]

we will again obtain (22) if we test the null of \( \delta_2 = \omega_1 = 0 \) and \( \omega_2 = 0 \) by means of a Lagrange multiplier test. Thus, (28) and (29) give a natural multivariate extension of the polynomial approach of Patton (2005). In addition, these equations show that our alternative hypothesis can model the types of non-linear exposures generated by the market timing signals of section 2. Specifically, the coefficient \( \delta_2 \) is related to the moment condition (18), while \( \omega_1 \) and \( \omega_2 \) are related to (17). Hence, by considering one-sided versions, we can test linear dependence versus unbiased, biased and pure noise signals. In particular, \( \omega_1 \) and \( \omega_2 \) are strictly positive under these three alternatives (see proposition 1). Therefore, only the strictly positive elements of the moment conditions related to these coefficients will be taken into account. However, the sign of \( \delta_2 \) will depend on the quality of the signal.

**Linear dependence versus market timing ability** When \( l = 1 \) in (1), \( \delta_2 \) is positive under the alternative hypothesis. In consequence, only the strictly positive elements of sample mean of (18) will be considered in this test. Before formally writing the explicit formula of the test, we first need to introduce some notation.

Consider a vector \( \mathbf{x} = (x_1, x_2, \ldots, x_R)' \), where only the \( q \leq R \) elements \( \mathbf{x}_+ = (x_{p_1}, x_{p_2}, \ldots, x_{p_q})' \) are strictly positive. Then, we can define the \( q \times R \) matrix \( \mathbf{L}_+(\mathbf{x}) = [\mathbf{e}_{p_1}, \mathbf{e}_{p_2}, \ldots, \mathbf{e}_{p_q}] \), where \( \mathbf{e}_i \) is the \( i \)--th column of the identity matrix of order \( R \). Hence
\( L'_+(x)x = x_+ \), that is, it yields the strictly positive elements of \( x \).

Using this matrix, we can express our test as:

\[
\tau_{mt1}(\hat{\theta}_T, b) = \left[ \frac{\sqrt{T}}{T} \sum_t v'_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right] L'_1 \\
\times [L_1 F(\theta_0) L'_1]^{-1} L_1 \left[ \frac{\sqrt{T}}{T} \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right]
\]

(30)

where \( L_{+1} = L_+ \left[ (\sqrt{T}/T) \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right] \).

**Linear dependence versus market mis-timing ability**  Now \( l = -1 \), so \( \delta_2 \) will be negative under the alternative hypothesis. Hence, only the strictly negative elements of (18) will be taken into account. Hence, in this case we can define a new matrix \( L_-(x) \) such that \( L_-(x)x \) yields the negative elements of \( x \). Hence, we can express this test as:

\[
\tau_{mt2}(\hat{\theta}_T, b) = \left[ \frac{\sqrt{T}}{T} \sum_t v'_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right] L'_2 \\
\times [L_2 F(\theta_0) L'_2]^{-1} L_2 \left[ \frac{\sqrt{T}}{T} \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right]
\]

(31)

where

\[
L_2 = \begin{bmatrix}
L_+ \left[ (\sqrt{T}/T) \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right] & 0 \\
0 & L_+ \left[ (\sqrt{T}/T) \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right]
\end{bmatrix}
\]

**Linear dependence versus pure noise signals**  Finally, if \( l = 0 \), we will maintain the hypothesis that \( \delta_2 = 0 \). Hence, this test can be can expressed as:

\[
\tau_{mt3}(\hat{\theta}_T, b) = \left[ \frac{\sqrt{T}}{T} \sum_t v'_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right] L'_3 \\
\times [L_3 F_{vv}(\theta_0) L'_3]^{-1} L_3 \left[ \frac{\sqrt{T}}{T} \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right]
\]

(32)

where

\[
L_3 = L_+ \left[ \frac{\sqrt{T}}{T} \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \right]
\]

and \( F_{vv}(\theta_0) \) is the sub-matrix of (19) that yields the variance of \( \sqrt{T}/T \sum_t v_{nl}(y_{ht}|y_{mt}; \hat{\theta}_T) \).

The three tests can be interpreted as Kuhn-Tucker multiplier tests (see Gouriéroux, Holly, and Monfort, 1982). Hence, their asymptotic distributions are mixtures of chi-squares. Unfortunately, finding the probability weights of these mixtures requires a rather
involved procedure, specially for more than two inequality constraints. In this paper, we will obtain those p-values by simulation, by generating independent Gaussian vectors with zero mean and covariance matrix $F(\theta_0)$ and computing the tests (30), (31) and (32) for each simulation. The p-values will be obtained from the empirical distribution of these simulated tests.

4.3 Extension to a joint test of linear dependence and absence of correlation

Both (21) and (22) allow for the presence of linear correlation between $y_{ht}$ and $y_{mt}$ under the null. As an extension, in this subsection we consider a test that also imposes zero correlation between these variables under the null. In particular, suppose we can partition $\theta$ as $(\delta', \rho')'$, where $\rho$ is a vector of $p$ parameters that only appears in the covariance term $\Sigma_{mht}(\delta, \rho)$, so that $\Sigma_{mht}(\delta, 0) = 0$. In this notation, we consider a joint test of $\rho = 0$ and $\psi_\eta = 0$ versus $\rho \neq 0, \psi_\eta \neq 0$. We will essentially follow the same approach as before, that is, we will obtain a consistent estimator of $\delta$ by PML, and then we will use this estimator to compute the test. However, this test has an additional moment condition based on the pseudo-score with respect to $\rho$:

$$s_{\rho t}(\delta, 0) = \lim_{\rho \to 0} \frac{\partial \ell_t(\theta)}{\partial \rho}$$

$$= Y_{hmt}(\delta) \left[ \Sigma_{mmt}^{-1}(\delta) \otimes \Sigma_{hht}^{-1}(\delta) \right] \text{vec} \left[ \varepsilon_{ht}(\delta) \varepsilon_{mt}'(\delta) \right],$$

where

$$Y_{hmt}(\delta) = \lim_{\rho \to 0} \frac{\partial \text{vec}' \Sigma_{mht}(\delta, \rho)}{\partial \rho}$$

Let $\hat{\delta}_T$ be the PML estimator of $\delta_0$. The asymptotic distribution of the moment conditions is as follows:

**Proposition 4**

$$\sqrt{T} \sum_t \left\{ v_{nt} \left[ s_{\rho t}(\hat{\delta}_T, 0) \right] \right\} \xrightarrow{d} N \left[ 0; D(\delta_0) \right]$$

where

$$D(\delta_0) = \left[ \begin{array}{cc} D_{pp}(\delta_0) & D_{p\eta}(\delta_0) \\ D_{\eta p}(\delta_0) & D_{\eta\eta}(\delta_0) \end{array} \right] = p \lim_{T \to \infty} \sum_t D_t(\delta_0)$$

$$D_t(\delta_0) = V_{t-1} \left\{ v_{nt} \left[ s_{\rho t}(\delta_0, 0) \right] \right\}$$
Instead of the original moment conditions, to make the derivation of the tests more simple, we will work with the orthogonalised moment conditions, $s_{\rho t}(\hat{\delta}_T,0)$, and

$$v_{\rho t}^+ \left[ y_{ht|y_{mt}}; (\hat{\delta}_T,0) \right] = v_{\rho t} \left[ y_{ht|y_{mt}}; (\hat{\delta}_T,0) \right] - D_{\eta\rho}(\delta_0) D_{\rho\rho}^{-1}(\delta_0) s_{\rho t} \left[ y_t; (\hat{\delta}_T,0) \right].$$

Hence, the test for given $b$ can be expressed as

$$\tau_l = \left[ \frac{\sqrt{T}}{T} \sum_t s_{\rho t} \left[ y_t; (\hat{\delta}_T,0) \right] \right]^T D_{\rho\rho}^{-1}(\delta_0) \left[ \frac{\sqrt{T}}{T} \sum_t s_{\rho t} \left[ y_t; (\delta_0,0) \right] \right]$$

$$+ \frac{b^+ \left[ \frac{\sqrt{T}}{T} \sum_t v_{\rho t}^+ \left[ y_{ht|y_{mt}}; (\hat{\delta}_T,0) \right] \right] \left[ \frac{\sqrt{T}}{T} \sum_t v_{\rho t} \left[ y_{ht|y_{mt}}; (\delta_0,0) \right] \right]^T b^+}{b^+ \left[ D_{\rho\eta}(\delta_0) - D_{\eta\rho}(\delta_0) D_{\rho\rho}^{-1}(\delta_0) D_{\rho\eta}(\delta_0) \right] b^+},$$

which has a chi-square distribution with $p + 1$ degrees of freedom. In consequence, the supremum test can be obtained by maximizing (33) with respect to $b$, which yields

**Proposition 5**

$$\sup_b \tau_l(\hat{\delta}_T, b) = \left[ \frac{\sqrt{T}}{T} \sum_t s_{\rho t}(\hat{\delta}_T,0) \right]^T D_{\rho\rho}^{-1}(\delta_0) \left[ \frac{\sqrt{T}}{T} \sum_t s_{\rho t}(\delta_0,0) \right]$$

$$\times D_{\rho\rho}^{-1}(\delta_0) \left[ \frac{\sqrt{T}}{T} \sum_t v_{\rho t}^+ \left[ y_{ht|y_{mt}}; (\delta_0,0) \right] \right] \left[ \frac{\sqrt{T}}{T} \sum_t v_{\rho t} \left[ y_{ht|y_{mt}}; (\delta_0,0) \right] \right]^T b^+$$

(34)

This test converges in distribution to a $\chi^2_{N+p+k+1}$. Again, it is straightforward to derive one sided versions of this test by proceeding analogously to the previous subsection.

### 4.4 Power of the test

It is interesting to study the power properties of the multivariate linear dependence tests derived in the previous sections. For simplicity, we will first assume that $\theta$ is known. As our D.G.P. for this exercise we will assume that $y_{mt}$ and $y_{ht}$ are jointly $iid$ GH with zero mean, unit variances and $\text{cor}(y_{mti}, y_{htj}) = \rho$, for all $1 \leq i \leq k$ and $1 \leq j \leq N$. We will compute power for a sample size of $T = 100$, based on the asymptotic distribution of the test under the alternative, which is known in explicit form. Specifically, in this study we will set the parameter $\psi$ equal to 1 and $\eta > 0$, so that our alternative hypothesis will be the member of the GH family known as the multivariate asymmetric $t$ distribution. In this distribution, we can use $\eta$ as a measure of dependence, since $\eta = 0$ yields the null hypothesis, and as $\eta$ gets larger, tail dependence increases. Importantly, when $b = 0$, $\eta$ can be interpreted as the reciprocal of the degrees of freedom of the Student $t$ distribution.
Results, at the usual 5% level, are displayed in figure 2 for the linear dependence test (22), whereas in figure 3 we plot the results for the joint test of linear dependence and absence of correlation (34). In these figures, symmetric alternatives must be interpreted as those that only introduce non-linear dependence in the conditional variance, while in asymmetric alternatives non-linear dependence also affects the conditional mean. We can obtain several important conclusions. First, the power of the linearity test against symmetric alternatives is noticeably lower than under asymmetric alternatives. Second, power can increase substantially if we consider dependence tests between vectors of higher dimension. Finally, we can see in figure 3 that non-zero linear correlation can yield large increases in power when it is included in the null of the test.

4.5 Finite sample properties

In this subsection, we assess the finite sample properties of the testing procedures discussed above, by means of several Monte Carlo exercises. We will consider a bivariate model with the following VAR(1) structure for the conditional mean:

\[
\mu_{mt}(\theta) = \alpha_m + \phi_m y_{mt-1}, \tag{35}
\]

\[
\mu_{ht}(\theta) = \alpha_h + \phi_{mh} y_{mt-1} + \phi_h y_{ht-1}, \tag{36}
\]

The reasons for this choice will be explained in the empirical application. The conditional variance will be assumed to be constant in most of the simulations. Nevertheless, we will also consider a GARCH(1,1) model with constant conditional correlation (CCC),

\[
\sigma_{mmt}(\theta) = \omega_m + \gamma_{m1} \sigma_{mmt-1}(\theta) + \gamma_{m2} [y_{mt-1} - \mu_{mt-1}(\theta)]^2, \tag{37}
\]

\[
\sigma_{hht}(\theta) = \omega_h + \gamma_{h1} \sigma_{hht-1}(\theta) + \gamma_{h2} [y_{ht-1} - \mu_{ht-1}(\theta)]^2, \tag{38}
\]

to analyse the effect of time varying variances on the properties of the tests. As for parameter values, we have chosen \(\alpha_i = 0.2, \phi_i = 0, \omega_i = 0.1, \gamma_{i1} = 0.1, \gamma_{i2} = 0.85\) for \(i = m, h\), and \(\phi_{mh} = 0\). In the unconditional case, we use the unconditional variances implied by the CCC-GARCH(1,1) model. The results, based on 30,000 Monte Carlo replications of samples of size \(T = 100\), are summarised by means of Davidson and MacKinnon’s (1998) p-value discrepancy plots. These plots show the difference between actual and nominal test sizes for every possible nominal size. Figures 4a-4f report the main results concerning the size of the tests. In figures 4a-4b we assume that \(y_{mt}\) and
are jointly Gaussian. The variances are kept constant through time in this case. As we can observe, the linear dependence test is subject to slightly higher size distortions than the zero correlation test. For example, for a 10% nominal size, we find that the actual size is around 9%. In figure 4b we check that the size of the linear dependence test remains correct when linear correlation is different from zero. In figures 4c-4d, we repeat the same exercise, assuming now that the shocks to \( y_{mt} \) and \( y_{ht} \) given \( y_{mt} \) are two independent Student t variates with 10 degrees of freedom. Thus, we keep a linear dependence structure but with non-Gaussian marginal distributions. As expected, the properties of the tests do not seem to be greatly affected by the presence of fat tails. Furthermore, size distortions are also of a similar magnitude when we parametrise the variance with the CCC-GARCH(1,1) model of (37) and (38) (see figures 4e-4f).  

In figures 5a-5f we explore the effect of several types of non-linear dependence on the power of the tests of absence of correlation and linear dependence. In all cases, the marginal distributions are Gaussian. In figure 5a-5b, we assume Student t dependence, which only introduces a quadratic effect on the conditional variance of \( y_{ht} \) given \( y_{mt} \). As a result, the co-kurtosis component of the test is the only moment condition that is able to capture this effect. In consequence, we obtain relatively low power. In contrast, in figures 5c-5d we introduce asymmetric dependence that affects both the conditional mean and variance. Hence, we now obtain substantial increases in power, since all the moment conditions are useful to detect non-linear dependence in this case. Notice that despite the fact that linear correlation is zero in both cases, we obtain some size distortions in the zero correlation test, due to the short sample sizes that we are considering.  

Finally, in figures 5e-5f we study the ability of our tests to detect the non-linear dependence introduced by the Clayton copula. This copula is modelled with only one parameter, that we will denote \( \kappa \). When \( \kappa = 0 \), \( y_{mt} \) and \( y_{ht} \) will be independent. As \( \kappa \) increases, this copula introduces not only linear correlation between the two variables, but also non-linear dependence. Eventually, as \( \kappa \) tends to infinity, both variables tend to be comonotonic. Since in this exercise these variables are drawn from the same marginal distribution, this implies that they tend to be perfectly correlated as \( \kappa \to \infty \) (see McNeil, Frey, and Embrechts, 2005, for further details). Hence, the only parameter of this copula induces both linear and non-linear dependence, except for the extreme cases of zero cor-

---

9In this case, the conditional mean is modelled with a diagonal VAR(1) model.
relation and comonotonicity. For the intermediate cases, however, we can see in figure 5f that the linear dependence test is able to capture the presence of non-linear dependence, although linear correlation seems to be the strongest effect introduced by this copula, as figure 5e shows.

5 Testing Student \( t \) dependence

5.1 Testing approach

The Student \( t \) distribution is nested in the GH family when \( \psi = 1 \) and \( b = 0 \). We can use this fact to derive a test of the dependence structure imposed by this distribution. In this case, the conditional mean and variance under the null can be written as

\[
E_{t-1}(y_{ht}|y_{mt}, z_t; \theta) = \mu_{ht}(\theta) + \sum_{hmt}(\theta) \sum_{mmt}^{-1}(\theta) [y_{mt} - \mu_{mt}(\theta)]
\]

\[
V_{t-1}(y_{ht}|y_{mt}, z_t; \theta) = \frac{1 - 2\eta + \eta\varsigma_{mt}(\theta)}{1 - \eta(2 - k)} \sum_{h|m,t}(\theta)
\]

where \( \eta \geq 0 \).

In consequence, in this case the conditional mean that we want to test is again linear in the conditioning variable, whereas the conditional variance is a function of \( s_{mt}(\theta) = \varepsilon'_{mt}(\theta) \sum_{mmt}^{-1}(\theta) \varepsilon_{mt}(\theta) \). Hence, we are allowing for quadratic effects in the variance. As in the previous section, there will be two steps. First, we will estimate \( \theta \) by Gaussian PML, which will yield \( \hat{\theta}_T \). Additionally, we also have to estimate \( \eta \). Fiorentini and Sentana (2005) also use a sequential approach to estimate the degrees of freedom of the multivariate Student \( t \) distribution. In particular, they exploit the relationship between the coefficient of multivariate kurtosis and the degrees of freedom in order to estimate \( \eta \). However, since we are interested in a test that remains valid even if the marginal distribution is not Student \( t \), we will base our estimate of \( \eta \) on the conditional pseudo-likelihood of \( y_{ht} \) given \( y_{mt} \). In consequence, we can obtain a consistent estimator, \( \hat{\eta}_T \), from

\[
\hat{\eta}_T = \arg \max_{\eta \geq 0} \left\{ \frac{1}{T} \sum_{t=1}^{T} \ell_{h|mt}(\hat{\theta}_T, \eta, \psi, b) \right\}
\]

subject to the restrictions of the null: \( \psi = 1 \) and \( b = 0 \). In this case, the conditional pseudo-log-likelihood under the null can be written as

\[
\ell_{h|mt}(\hat{\theta}_T, \eta, 1, 0) = -\frac{N}{2} \log (2\pi) - \frac{N}{2} \log [1 - 2\eta + \eta s_{mt}(\theta)] + \frac{N}{2} \log [1 - \eta(2 - k)]
\]

\[
-\frac{1}{2} \log \left| \sum_{h|m,t}(\theta) \right| - \frac{1}{2} \frac{1 - \eta(2 - k)}{1 - 2\eta + \eta s_{mt}(\theta)} \varsigma_{h|mt}(\theta)
\]

---

Footnote: \( \eta \) would the reciprocal of the degrees of freedom if the joint distribution of \( y_{mt} \) and \( y_{ht} \) were Student \( t \).
which, once differentiated with respect to \( \eta \), yields the pseudo-score

\[
s_{\eta t}(y_{ht}|y_{mt}; \pi) = \frac{s_{mt}(\theta) - k}{2[1 - 2\eta + \eta s_{mt}(\theta)][1 + \eta(k - 2)]} \times \left[ \frac{[1 + \eta(k - 2)]s_{htmt}(\theta)}{1 - 2\eta + \eta s_{mt}(\theta)} - N \right]
\]

(41)

where \( \pi = (\theta, \eta) \). As regards the asymptotic distribution of \( \hat{\eta}_T \), we can show that, when the true value of \( \eta \) is \( \eta_0 > 0 \):

**Proposition 6** \( \sqrt{T}[\hat{\eta}_T - \eta_0] \) is asymptotically normal with zero mean and variance

\[
T_{\eta}^{-2}(\pi_0) H(\pi_0) p \lim_{T \to \infty} V \left\{ \frac{\sqrt{T}}{T} \sum_t s_{\theta t}(y_i; \theta_0) s_{\eta t}(y_{ht}|y_{mt}; \pi_0) \right\} H'(\pi_0)
\]

where

\[
H(\pi_0) = \left[ -T_0'(\pi_0) A_{\theta \theta}^{-1}(\theta_0) \ 1 \right],
\]

(42)

and

\[
T_{\eta}(\pi_0) = -p \lim_{T \to \infty} \frac{1}{T} \sum_t \frac{\partial s_{\eta t}(y_{ht}|y_{mt}; \pi_0)}{\partial \eta}
\]

(43)

\[
T_{\theta}(\pi_0) = -p \lim_{T \to \infty} \frac{1}{T} \sum_t \frac{\partial s_{\eta t}(y_{ht}|y_{mt}; \pi_0)}{\partial \theta}
\]

(44)

### 5.2 Moment conditions of the test

In principle, we would like to base the test on the pseudo-scores with respect to \( b \) and \( \psi \). In this sense, we can show that the score with respect to \( b \) is

\[
s_{bt}(y_{ht}|y_{mt}; \pi) = \eta \left[ \frac{[1 + \eta(k - 2)]s_{htmt}(\theta)}{1 - 2\eta + \eta s_{mt}(\theta)} - N \right] \frac{\varepsilon_{mt}(\theta)}{1 - 2\eta + \eta s_{mt}(\theta)} .
\]

(45)

However, the pseudo score with respect to \( \psi \) is identically zero under the null. To deal with this unusual type of testing problem, also present in the multivariate Student t test of Mencía and Sentana (2005), Lee and Chesher (1986) propose to replace this moment condition by what they call an “extremum test” (see also Bera, Ra, and Sarkar, 1998).

Given that the first order conditions are identically zero, their suggestion is to study the restrictions that the null imposes on higher order conditions. In our case, we will use the second order derivative of the pseudo-log-likelihood with respect to \( \psi \), i.e.

\[
s_{\psi \psi t}(y_{ht}|y_{mt}; \pi) = \eta^2 \left[ \frac{[1 + \eta(k - 2)]s_{htmt}(\theta)}{1 - 2\eta + \eta s_{mt}(\theta)} - N \right] \times \left[ \frac{(1 - 4\eta)^{-1}}{1 - 2\eta + \eta s_{mt}(\theta)} - (1 - 2\eta)^{-1} \frac{[1 - 2\eta + \eta s_{mt}(\theta)]}{[1 + \eta(k - 2)][1 + \eta(k - 4)]} \right]
\]

(46)
It is possible to show that (46) has zero expected mean under the null.

We will first consider a two-sided moment test. The following proposition gives the asymptotic variance of these conditions

**Proposition 7**

\[
\frac{\sqrt{T}}{T} \sum_t \left[ \frac{s_{bt}(y_{ht}|y_{mt}; \hat{\pi}_T)}{s_{\psi\psi t}(y_{ht}|y_{mt}; \hat{\pi}_T)} \right] \xrightarrow{d} N[0, V(\pi_0)]
\]

where

\[
V(\pi_0) = J(\pi_0) p \lim_{T \to \infty} V \left\{ \frac{\sqrt{T}}{T} \sum_t \left[ \begin{array}{c} s_{\theta t}(y_{ht}|y_{mt}; \theta_0) \\ s_{\eta t}(y_{ht}|y_{mt}; \pi_0) \\ s_{bt}(y_{ht}|y_{mt}; \pi_0) \\ s_{\psi\psi t}(y_{ht}|y_{mt}; \pi_0) \end{array} \right] \right\} J'(\pi_0)
\]

and \( J(\pi_0) = \begin{bmatrix} J_{\pi}(\pi_0) & I_{N+k+1} \end{bmatrix} \), with

\[
J_{\pi}(\pi_0) = p \lim_{T \to \infty} \left\{ \frac{-1}{T} \sum_t \frac{\partial}{\partial \pi} \left[ \begin{array}{c} s_{bt}(y_{ht}|y_{mt}; \pi_T) \\ s_{\psi\psi t}(y_{ht}|y_{mt}; \pi_T) \end{array} \right] \right\} \times \left[ \begin{array}{c} A_{\theta \theta}^{-1}(\theta_0) \\ -T_\eta^{-1}(\pi_0) T_\theta(\pi_0) A_{\theta \theta}(\theta_0) \end{array} \right]
\]

Hence, our proposed test is

\[
\tau_{2s}(\hat{\pi}_T) = \left\{ \frac{\sqrt{T}}{T} \sum_t \left[ \begin{array}{c} s_{bt}(y_{ht}|y_{mt}; \hat{\pi}_T) \\ s_{\psi\psi t}(y_{ht}|y_{mt}; \hat{\pi}_T) \end{array} \right] \right\} V^{-1}(\hat{\pi}_T)
\]

which converges asymptotically to a chi-square with \( N+k+1 \) degrees of freedom. However, higher power against GH alternatives can be achieved with a one-sided test. In this respect, following Lee and Chesher (1986), it is possible to show that the expected value of \( s_{\psi\psi t}(y_{ht}|y_{mt}; \pi_T) \) should be positive for \( \psi < 1 \). Based on this result, we can construct a partially one-sided version of the previous result. We first need to orthogonalise the moment conditions, since \( V(\pi_0) \) will not be generally block-diagonal. Specifically, instead of the pseudo-score with respect to \( b \), we consider

\[
s_{bt}^{\perp}(y_{ht}|y_{mt}; \hat{\pi}_T) = s_{bt}(y_{ht}|y_{mt}; \hat{\pi}_T) - V_{b\psi}(\pi_0) V_{\psi\psi}^{-1}(\pi_0) s_{\psi\psi t}(y_{ht}|y_{mt}; \hat{\pi}_T)
\]

where again we have partitioned the asymptotic covariance matrix as

\[
V(\pi_0) = \begin{bmatrix} V_{bb}(\pi_0) & V_{b\psi}(\pi_0) \\ V_{\psi b}^T(\pi_0) & V_{\psi\psi}(\pi_0) \end{bmatrix}
\]
It is then straightforward to see that the asymptotic distribution of
\[\tau_1s(\hat{\pi}_T) = \sqrt{\frac{T}{T}} \sum_t s_{bt}(y_{ht}|y_{mt}; \hat{\pi}_T)\]
\[\times \left[ \nu_{bb}(\hat{\pi}_T) - \frac{\nu_{bb}(\pi_0) \nu_{bb}'(\pi_0)}{\nu_{\psi\psi}(\pi_0)} \right]^{-1} \sqrt{\frac{T}{T}} \sum_t s_{bt}(y_{ht}|y_{mt}; \hat{\pi}_T)\]
\[+ \sqrt{\frac{T}{T}} \sum_t s_{\psi\psi}(y_{ht}|y_{mt}; \hat{\pi}_T) \nu_{\psi\psi}(\pi_0) \left( \sum_t s_{\psi\psi}(y_{ht}|y_{mt}; \hat{\pi}_T) > 0 \right)\]
is a 50:50 mixture of chi-squares with \(N + k\) and \(N + k + 1\) degrees of freedom under the null.

### 5.3 Interpretation

Similarly to what we did with the linear dependence test, it can be shown that (28) and
\[V(y_{ht}|y_{mt}) = \sum_{h|m} \left[ \frac{1 - 2\eta + \eta smt(\theta)}{1 + \eta(k - 2)} + \omega_1(y_{mt} - \mu_m) \right.\]
\[+ \omega_2 \left. \left( \frac{1}{1 - 4\eta} - (1 - 2\eta)^{-1} (1 - 2\eta + \eta smt(\theta))^2 \right) \right] \]
give a locally equivalent alternative. Under this interpretation, the coefficients \(\delta_2\) and \(\omega_1\) yield the moment conditions (45), whereas \(\omega_2\) is responsible for (46). As we can see in (47), this latter condition introduces a fourth order effect in conditional variance under the alternative hypothesis. We can also exploit this equivalence to obtain one-sided versions of the test against market (mis-)timing ability. In this case we would test \(\delta_2 = \omega_1 = 0\) versus \(\delta_2 > 0, \omega_1 > 0\) (market timing) or \(\delta_2 < 0, \omega_1 > 0\) (market mis-timing).

### 6 Dependence in the hedge Fund industry

In this section, we will use the tests derived in the previous sections to analyse the dependence between the hedge fund industry and the market. We will use a combined database form HFR and TASS. Our dataset consists of monthly returns for individual hedge funds from December 1992 to August 2003. It is organised in five groups of styles: event driven, market neutral, equity hedge, equity non hedge and fund of funds.\(^\text{11}\) The database contains both live and dead funds, although it is not reported whether funds with an incomplete history are dead or they have just stopped reporting their returns.

We will test dependence between each fund in the database and the market. Patton

\(^{11}\)See Lhabitant (2002), for a detailed description of hedge fund strategies
(2005) has studied “market neutrality” for this same dataset. However, in this paper we will explicitly focus on non-linear dependence, as well as on market (mis-)timing ability. Additionally, we will also study dependence between the hedge fund industry as a whole and the market, using data on the 10 CSBF-Tremont main style indices from April 1994 to July 2005. Since both datasets are based on returns of live and dead funds, it is not clear what effect survivorship bias will have on the results. Selection and backfill biases may also be present. It is very difficult to quantify those biases, since we would need to know the characteristics of the funds that are not in the database. Nevertheless, we are more interested in exposure to the market than in expected returns themselves. Hence, for our purposes, to the extent that those biases have remained stable, or at least they have not been non-linearly influenced by the market during the sample period of our study, they will not affect the validity of our tests.

6.1 Individual tests

We will start by considering the dependence between each hedge fund in our database and the market, proxied by the S&P 500 index. In consequence, we will work on a pair by pair basis (i.e. \( k = N = 1 \)). Under the methodology of this paper, we can easily account for the presence of autocorrelation and volatility clustering in the conditional mean and variance. However, it is important to take into account that the largest sample sizes are only of 129 months, while the smallest ones can be as low as one month. Because of this, we will only compute the tests for those funds with \( T \geq 30 \). In consequence, the selection of a parsimonious model will be crucial in order to ensure a reasonable power to reject the null.

We will start by considering a VAR(1) model with constant conditional variance. As Asness, Krail, and Liew (2001) have emphasised, many hedge funds hold illiquid securities, which sometimes are difficult to price. For some of them, there are no publicly available traded prices at the end of each month, the result being a considerable amount of flexibility in how hedge funds can mark their positions for their monthly reports. One important consequence of this lack of liquidity is the presence of time series autocorrelation (see Lo, 2001). In addition, under the presence of these stale prices, hedge fund returns may respond to market’s moves with some lag. Consequently, following Scholes and Williams (1977), it may also be important to consider a feedback effect of lagged market returns.
on current hedge funds returns. Our methodology can easily account for both features. In particular, we consider the triangular VAR(1) model given by (35) and (36).

The proportions of tests rejections are reported in table 1a. If we want to interpret these results as joint tests of dependence in a specific category, we need to compute the 95% percentile of the distribution of the proportion of funds failing the tests under the null.\textsuperscript{12} Intuitively, there is a 5% (nominal) probability of rejecting the null for a specific fund when it is satisfied. Thus, a non-zero proportion of test rejections does not necessarily imply a rejection of the null for a hedge fund style group. In consequence, we need to find a critical value for that proportion in order to determine when those percentages are significantly different from zero. Assuming that the funds are independent, this critical value can be obtained from a binomial distribution. However, the independence assumption does not appear to be very realistic given that we are considering homogeneous groups of funds that presumably follow similar strategies. To deal with this type of issue, Patton (2005) uses a bootstrap method. In our case, we will obtain asymptotically consistent critical values by computing the dependence tests between the original hedge fund returns and 1,000 replications of the market index, generated from an independent Gaussian distribution with the same parameters as those estimated for the true market index (see appendix C for more details).

Table 2a reports critical values for our triangular bivariate VAR(1) model with constant conditional variance, together with the binomial critical values. The critical values obtained with the simulation method are larger than the binomial ones. This shows the importance of taking into account the presence of dependence between funds. In we compare the values of table 1a with the corresponding 95% percentiles of table 2a, we can see that absence of correlation is significant for the five styles. Nevertheless, this proportion is noticeably lower in the market neutral group, with a value of around 36%. In consequence, as far as linear correlation is concerned, this group seems to be more “market neutral” than the rest, although theoretically we should not find any correlation at all. On the other hand, around 87% of Equity non Hedge funds have a significant correlation with the market. This result is consistent with the characteristics of these funds, which, also known as “stock pickers”, usually rely on directional bets on some assets, without necessarily hedging their positions. However, quite a different pattern emerges when we

\textsuperscript{12}Of course the mean of this distribution is 5% when we consider 95% confidence levels for the individual tests, but if we take this number as our benchmark, we will incur in size distortions.
consider the linear dependence tests. Non-linear dependence is significant for all groups except for Equity Hedge and Equity non Hedge funds. The last columns show that in the Funds of funds and Event Driven groups there are more rejections of linear dependence due to inability to time the market than to ability. In addition, pure noise market signals also seem to play an important role, and Equity Hedge is the only style where non-linear dependence remains insignificant even when we consider market timing alternatives.

It can be argued that the S&P 500 only gives an imperfect estimate of the true world market portfolio. To explore this issue, we now consider dependence with respect to a broader measure of the market, composed by the S&P 500 and the FT 100.13 In this case, we have also considered a model with constant covariance matrix and a triangular VAR(1) that allows for a feedback effect of lagged values of both market variables on the expected funds returns. Due to the higher degree of parametrisation of this trivariate model, now we will only consider funds with at least 50 months of data. As we can see in table 1b, the inclusion of the second factor reveals a presence of correlation with the market similar to what we found with just one factor. Those proportions are significant at the 95% confidence level, as table 2b shows. In sharp contrast, non-linear dependence is no longer significant in the market neutral group. A higher survivorship bias may be a more plausible explanation of this result, since we are using funds with longer histories than in table 1a. However, when we consider market-timing effects, the inability to time the market remains significant for the Funds of Funds, Event Driven and Market Neutral styles. In contrast, evidence in favour of market timing ability and uninformative signals is only significant for the first two groups.

Finally, we study whether the results that we have obtained are affected by the presence of volatility clustering effects. We model the market, proxied by the S&P 500, and each fund’s returns with a diagonal VAR(1) mean and a GARCH(1,1) parametrisation for the variances. To keep the number of parameters at a minimum, we do not consider the feedback effect of the market in this case. The covariance between each hedge fund and the market is assumed to satisfy the constant conditional correlation (CCC) model of Bollerslev (1990). Results, shown in table 1c, are barely different to those of table 1a. Hence, our previous conclusions do not seem to be caused by time varying volatilities.

\footnote{13Converted into dollars}
6.2 Industry tests

The above results suggest a considerable presence of non-linear dependence in the hedge fund industry, specially in funds of funds and the event driven style, and in a somewhat lower degree in the market neutral group. However, these are only tests of individual market exposure, and not joint tests for the industry as a whole. Furthermore, they only represent four groups of strategies, a relatively small proportion of the much larger universe of available styles. In this subsection we test the joint dependence between the 10 CSFB-Tremont hedge fund style indices and the previous market factors.\textsuperscript{14} CSFB-Tremont calculate value-weighted indices, with sensible criteria in order to decide on the inclusion of a fund in the index. Consequently, these style indices broadly represent the state of the whole industry. Table 3 shows the results for a triangular VAR(1) model with feedback of the lagged market return, in the spirit of (35) and (36), where the variance is modelled with a CCC-GARCH(1,1) parametrisation.\textsuperscript{15} The second column reports the exposure to the S&P 500, while the third column also considers the FT 100 as a factor. Linear dependence is easily rejected for both models, thus confirming our conclusions in the previous subsection. However, we cannot reject Student $t$ dependence, although we obtain a p-value close to 10\% when we use the two market factors and our alternative hypothesis is market mis-timing. Hence, at the aggregate level, the empirical evidence is consistent with the use of uninformative market signals in this industry, but not with systematically good or bad information.

Finally, we will study the implications of the observed non-linear dependence on hedge funds performance. Again, we consider investments in the 10 CSFB-Tremont style indices. We will model the mean and variance with the same VAR(1)-CCC GARCH(1,1) parametrisation of table 3, focusing on the analysis of investment opportunities in August 1998 and 2005. August 1998 is the month of the Russian debt restructuring, while the other date has been selected as a robustness check. Specifically, we will compute the mean-variance frontier for the orthogonalised asset returns $y_{h_{t+1}} = y_{h_{t+1}} - \Sigma_{hm\,t+1}(\hat{\theta}_T)\sigma_{mm\,t+1}(\hat{\theta}_T)y_{m\,t+1}$ in August 1998 (2005), conditional on a prospect for the

\textsuperscript{14}These styles are: Convertible Arbitrage, Dedicated Short Bias, Emerging Markets, Equity Market Neutral, Event Driven, Fixed Income Arbitrage, Global Macro, Long/Short Equity, Managed Futures and Multi-Strategy.

\textsuperscript{15}We have assessed the adequacy of the CCC assumption by computing the test proposed by Engle and Sheppard (2001) with lags between 1 and 6, but we have not been able to reject this null at the 10\% level for any of the two models that we analyse.
S&P 500 return for this period and on available information up to the end of July 1998 (2005). To analyse the effect of non-linear dependence, we will compare the frontiers obtained from the Student $t$ and GH models with the Gaussian dependence one. Since we are using the orthogonalised returns, linear correlation with the market is zero in all models. Hence, differences will only be due to non-linear dependence. The interest of the mean-variance frontier is that it allows us to carry out this comparison graphically. Figures 7 and 8 depict these frontiers for four possible scenarios. In each scenario, we assume that the standardised shocks to the market in the month of interest belong to one of the quartiles of the historical distribution of these shocks, represented in figure 6. In other words, we consider four possible deviations of the market return from its expected mean, each of which implies a different degree of optimism/pessimism.

To obtain the mean and variance of $y_{ht+1}$ conditional on past information and $y_{mt+1}$ being on a specific quartile, we exploit the law of iterated expectations with the following formulas:

$$E_t [y_{ht+1} | y_{mt+1} \in (q_{ai}, q_{bi}), \theta] = E_t [E_t [y_{ht+1} | y_{mt+1}, \theta] | y_{mt+1} \in (q_{ai}, q_{bi}), \theta],$$

$$V_t [y_{ht+1} | y_{mt+1} \in (q_{ai}, q_{bi}), \theta] = E_t [V_t [y_{ht+1} | y_{mt+1}, \theta] | y_{mt+1} \in (q_{ai}, q_{bi}), \theta] + V_t [E_t [y_{ht+1} | y_{mt+1}, \theta] | y_{mt+1} \in (q_{ai}, q_{bi}), \theta],$$

where $q_{ai}$ and $q_{bi}$ are, respectively, the lower and upper bounds of the $i^{th}$ quartile of the distribution of $y_{mt+1}$ conditional on information up to time $t$. The inner expectations are known in close form for the three dependence structures that we consider, while the outer ones are obtained from the empirical distribution of the S&P 500. 16

As we can observe in the figures, in the central quartiles, which can be interpreted as normal times, we see that the Gaussian dependence model clearly underestimates the investment opportunities, compared to the other two models. In sharp contrast, in the top quartile the three models give a very similar frontier, whereas in the bottom quartile it seems that the Gaussian model overestimates the investment opportunities, specially in 1998. The conclusion that emerges from these figures is that non-linear dependence in the hedge fund industry seems to generate a very good risk-return tradeoff in normal times, but these gains basically disappear in very good times. What is more important,

16Specifically, we use the artificial market variables $\tilde{y}_{mt+1,s} = E_t(y_{mt+1}; \theta) + V_t^{1/2}(y_{mt+1}; \theta)\varepsilon_s^{(i)}$, where $\varepsilon_s^{(i)} = [y_{ms} - E_{s-1}(y_{ms}; \theta)]/V_{s-1}^{1/2}(y_{ms}; \theta)$ are the historical market standardised returns that belong to the $i^{th}$ quartile of its empirical distribution.
due to non-linear dependence, hedge funds appear to be worse investments when extreme downside moves of the market occur. Interestingly, this asymmetric effect is also observed for Student $t$ dependence, even though this model imposes symmetric dependence. This is due to the negative skewness of the empirical distribution of the market (see figure 6). Hence, the larger magnitude in absolute terms of shocks in the very bad scenario compared to those of the very good scenario is responsible for this asymmetry. In addition, this result seems to be robust to selected month of analysis, although its magnitude is of course different.

7 Conclusions

This paper analyses dependence between asset returns and market conditions. We first study this issue from a theoretical perspective, and show that non-linear dependence on market factors can arise when market timing investment strategies are followed. Good market timers generate a non-linear dependence that can be beneficial. Unfortunately, a potentially harmful exposure is obtained when they are unable to correctly time the market. We also show the relationship of these effects with three types of dependence structures: Gaussian or linear, Student $t$ and Generalised Hyperbolic. Linear factor models, although extensively used, are not able to account for non-linear dependence on the factors. Interestingly, Student $t$ dependence is consistent with the use of uninformative market signals by managers, whereas the Generalised Hyperbolic distribution can model active market timing as well as market mis-timing effects. We derive tests of Gaussian and Student $t$ dependence between two groups of assets, using the dependence structure of the Generalised Hyperbolic distribution as alternative. Thanks to our pseudo-maximum likelihood approach, these tests are only subject to correct specification of the mean and variance, but do not make any additional assumptions about the marginal distributions. In addition, our methodology can be easily coupled with a wide variety of models for the conditional mean and variance, such as autoregressive processes or GARCH models.

One of the areas in which the results of this paper are more relevant is the hedge fund industry. Hedge funds are alternative investment vehicles with absolute return targets regardless of market conditions. However, in our empirical application to a database of around 1,500 hedge funds, we find an important presence of non-linear dependence with respect to the market. Unfortunately, market timing is less common
than market mis-timing or just uninformative market predictions. Furthermore, we use data on 10 style indices to study dependence in this industry as a whole, and find that linear dependence with respect to the S&P 500 and the FT 100 can be easily rejected, whereas Student $t$ dependence seems to provide a much better fit. This is specially worrying due to the rapid growth of this industry in the previous years. From 2000 to 2005, hedge funds have doubled in size and number (see The Economist, 2005). Many new investors are beginning to move their money to alternative investments. Therefore, it is extremely important that they are fully aware of the advantages and inherent risks in these instruments. In this sense, we find that the presence of non-linear dependence improves the diversification benefits of hedge funds in normal times. However, those benefits essentially disappear in good times, whereas in bad times the gains from investing in hedge funds appear to deteriorate because of non-linear dependence.

A fruitful avenue for future research will be to study whether hedge funds whose managers systematically mis-time the market have shorter lives than those with managers that correctly predict market trends. Otherwise, in equilibrium the first type of managers should be able to provide a higher expected return in exchange for higher risk in order to attract investors.
References


Appendix

A  Proofs of Propositions

A.1 Proposition 1

There are no short sales constraints. Hence, since $y_m$ can be spanned by $y_a$, investing in $y_a$ is equivalent to investing in the synthetic assets $\tilde{y}_a = (y'_m, y^\perp_m)'$, where $y^\perp_m = Ry_a$ and $R$ is an $(A - k) \times A$ matrix of rank $A - k$, such that $R\Sigma_{am} = 0$. In other words, $y^\perp_m$ are the $A - k$ non-redundant risky assets spanned by $y_a$ that are orthogonal to $y_m$. In this context, the optimal allocations to $\tilde{y}_a$ can be written as:

$$w_t = \alpha^{-1} \left[ V^{-1} (y_m | s, l = 1) E (y_m | s, l = 1) - r\iota_k \right] (A1)$$

where, through tedious but straightforward algebra, it can be shown that

$$E (y_m | s, l = 1) = \mu_m + (1 + \lambda)^{-1} s$$

$$V (y_m | s, l = 1) = \lambda (1 + \lambda)^{-1} \Sigma_{mm}$$

We can substitute (1) in (A2) and use this result together with (A3) and (A1) to express $y^\prime_h = (1 - w^\prime \iota_A) r + w^\prime \tilde{y}_a$ as a function of $y_m, y^\perp_m$ and $\varepsilon_{s|m}$. Finally, we can exploit the independence between these three vectors to compute the conditional mean and variance provided in the proposition.

A.2 Proposition 2

Consider the following Taylor expansion for the first order conditions of the pseudo score

$$\frac{\sqrt{T}}{T} \sum_t s_{\theta t} (y_t; \hat{\theta}_T) = 0 = \frac{\sqrt{T}}{T} \sum_t s_{\theta t} (y_t; \theta_0)$$

$$- A_{\theta \theta} (\theta_0) \sqrt{T} \left[ \hat{\theta}_T - \theta_0 \right] + o_p (1)$$

Hence, we can write

$$\sqrt{T} \left[ \hat{\theta}_T - \theta_0 \right] = [A_{\theta \theta} (\theta_0)]^{-1} \frac{\sqrt{T}}{T} \sum_t s_{\theta t} (y_t; \theta_0) + o_p (1)$$

(A4)
Now consider an analogous Taylor expansion of the moment conditions of the non-linear dependence test

\[
\frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \hat{\theta}_T \right) = \frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \theta_0 \right) \\
+ \left[ \frac{1}{T} \sum_t \frac{\partial v_{nt} \left( y_{ht} | y_{mt} ; \theta_0 \right)}{\partial \theta} \right] \sqrt{T} \left[ \hat{\theta}_T - \theta_0 \right] + o_p \left( 1 \right)
\]

\[
= \frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \theta_0 \right) - A_{\theta \eta} \left( \theta_0 \right) \sqrt{T} \left[ \hat{\theta}_T - \theta_0 \right] + o_p \left( 1 \right)
\]

where

\[
A_{\theta \eta} \left( \theta_0 \right) = \lim_{T \to \infty} - \frac{1}{T} \sum_t \frac{\partial v_{nt} \left( y_{ht} | y_{mt} ; \theta_0 \right)}{\partial \theta} \\
= \lim_{T \to \infty} \frac{1}{T} \sum_t A_{\theta \eta} \left( \theta_0 \right)
\]

\[
A_{\theta \eta} \left( \theta_0 \right) = \left[ \frac{\partial \text{vec} \left( \Sigma_{hmt} \left( \theta \right) \right)}{\partial \theta} \left[ \Sigma_{mmt}^{-1} \left( \theta \right) \otimes I_N \right] \\
- \frac{\partial \text{vec} \left( \Sigma_{mmt} \left( \theta \right) \right)}{\partial \theta} \left[ \Sigma_{mmt}^{-1} \left( \theta \right) \otimes \Sigma_{mmt}^{-1} \left( \theta \right) \Sigma_{mht} \left( \theta \right) \right] \varepsilon_{mt} \left( \theta \right) \right] \\
\times \left[ \varepsilon_{mt} \left( \theta \right) \otimes I_N \right] \left[ L_h \right]
\]

and, if we use (A4), we obtain

\[
\frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \hat{\theta}_T \right) = \left[ - A_{\theta \eta} \left( \theta_0 \right) A_{\theta \eta}^{-1} \left( \theta_0 \right) \right] I_{N+k+1} \\
\times \frac{\sqrt{T}}{T} \sum_t \left[ s_{\theta t} \left( y_t ; \theta_0 \right) \right] + o_p \left( 1 \right) \quad (A5)
\]

In consequence, the asymptotic distribution under the null of (A5) is Gaussian with zero mean and covariance matrix \( \mathcal{F} \left( \theta_0 \right) \), using the fact that \( v_{nt} \left( y_{ht} | y_{mt} ; \theta_0 \right) \) and the score with respect to \( \theta \) are orthogonal under the null for known \( \theta_0 \)

### A.3 Proposition 3

\( \tau_d \left( \hat{\theta}_T, b \right) \) can be trivially expressed as

\[
\tau_d \left( \hat{\theta}_T, b \right) = \frac{T b^+ \left[ \frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \hat{\theta}_T \right) \right] \left[ \frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \hat{\theta}_T \right) \right]^\prime b^+}{b^\prime \mathcal{F} \left( \theta_0 \right) b^+} \quad (A6)
\]

Since the maximisation of (A6) is a well-known generalised eigenvalue problem, its solution will be proportional to

\[
\mathcal{F}^{-1} \left( \theta_0 \right) \left[ \frac{\sqrt{T}}{T} \sum_t v_{nt} \left( y_{ht} | y_{mt} ; \hat{\theta}_T \right) \right]
\]

35
We can make sure that the first element of \( b^+ \) is equal to 1 by choosing the constant of proportionality appropriately. Substituting this value in (A6) yields the required result. Finally, the asymptotic distribution of the sup test follows directly from the result of proposition 2.

**A.4 Proposition 4**

\[
\sqrt{T} \sum_t \left\{ \begin{array}{l}
\frac{s_{pt}(\delta_T, 0)}{v_{pt}[y_{ht}|m_{yt};(\delta_T, 0)]} \\
\end{array} \right\} = \sqrt{T} \sum_t \left\{ \begin{array}{l}
\frac{s_{pt}(\delta_0, 0)}{v_{pt}[y_{ht}|m_{yt};(\delta_0, 0)]} \\
\end{array} \right\}
\]

\[
+ p \lim_{T \to \infty} \frac{1}{T} \sum_t \frac{\partial}{\partial \delta} \left\{ \begin{array}{l}
\frac{s_{pt}(\delta_0, 0)}{v_{pt}[y_{ht}|m_{yt};(\delta_0, 0)]} \\
\end{array} \right\} \sqrt{T} \left[ \delta_T - \delta_0 \right] + o_p(1)
\]

where

\[
p \lim_{T \to \infty} \frac{1}{T} \sum_t \frac{\partial}{\partial \delta} \left\{ \begin{array}{l}
\frac{s_{pt}(\delta_0, 0)}{v_{pt}[y_{ht}|m_{yt};(\delta_0, 0)]} \\
\end{array} \right\} = \]

\[
= p \lim_{T \to \infty} \frac{1}{T} \sum_t \left| E_{t-1} \right| \left\{ \begin{array}{l}
\frac{\partial}{\partial \delta} \left\{ \begin{array}{l}
\frac{s_{pt}(\delta_0, 0)}{v_{pt}[y_{ht}|m_{yt};(\delta_0, 0)]} \\
\end{array} \right\} \right| \delta_0 \}
\]

(A7)

and, computing the derivatives in (A7) and using the assumptions of the null, it is fairly easy to show that (A7) equals zero. Hence, uncertainty in the estimation of \( \delta \) does not affect the asymptotic distribution of the moment conditions.

**A.5 Proposition 5**

The proof of this proposition can be straightforwardly obtained by following the same steps of the proof of proposition 3.

**A.6 Propositions 6 and 7**

First, consider a Taylor expansion for the pseudo-score with respect to \( \eta \)

\[
0 = \frac{\sqrt{T}}{T} \sum_t s_{pt}(y_{ht}|m_{yt}; \hat{\eta}_T) = \frac{\sqrt{T}}{T} \sum_t s_{pt}(y_{ht}|m_{yt}; \eta_0)
\]

\[
+ \left[ \frac{1}{T} \sum_t \frac{\partial s_{pt}(y_{ht}|m_{yt}; \eta_0)}{\partial \theta} \right] \sqrt{T} \left( \hat{\eta}_T - \eta_0 \right)
\]

\[
+ \left[ \frac{1}{T} \sum_t \frac{\partial s_{pt}(y_{ht}|m_{yt}; \eta_0)}{\partial \eta} \right] \sqrt{T} \left( \hat{\eta}_T - \eta_0 \right) + o_p(1)
\]
Using (A4), it is straightforward to show that
\[
\sqrt{T}(\hat{\eta}_T - \eta_0) = T_{\eta}^{-1}(\pi_0) \mathcal{H}(\pi_0) \frac{\sqrt{T}}{T} \sum_t \left[ s_{\theta t}(y_t; \theta_0) \right. \\
\left. + s_{nt}(y_{ht}|y_{mt}; \pi_0) \right] + o_p(1) \tag{A8}
\]
where \( \mathcal{H}(\pi_0) \) and \( T_{\eta}(\pi_0) \) are defined by (42), (43) and (44). From (A8), it is straightforward to obtain the result of proposition 6.

Now, consider a Taylor expansion of the moment conditions of the test around the true parameter values, i.e.
\[
\sqrt{T} \sum_t s_{bt}(y_{ht}|y_{mt}; \hat{\pi}_T) = \sqrt{T} \sum_t s_{bt}(y_{ht}|y_{mt}; \pi_0) \\
+ \frac{1}{T} \sum_t \frac{\partial s_{bt}(y_{ht}|y_{mt}; \pi_0)}{\partial \pi'} \sqrt{T}(\hat{\pi}_T - \pi_0) + o_p(1)
\]
and
\[
\sqrt{T} \sum_t s_{\psi t}(y_{ht}|y_{mt}; \hat{\pi}_T) = \sqrt{T} \sum_t s_{\psi t}(y_{ht}|y_{mt}; \pi_0) \\
+ p \lim_{T \to \infty} \left\{ \frac{1}{T} \sum_t \frac{\partial s_{\psi t}(y_{ht}|y_{mt}; \pi_0)}{\partial \pi'} \right\} \sqrt{T}(\hat{\pi}_T - \pi_0) + o_p
\]
Using the fact that, from (A4) and (A8), we can write
\[
\sqrt{T}(\hat{\pi}_T - \pi_0) = \left[ \begin{array}{cc} A_{\theta \theta}^{-1}(\theta_0) & 0 \\
-T_{\eta}^{-1}(\pi_0) T_{\theta}^{-1}(\pi_0) A_{\theta \theta}^{-1}(\theta_0) & T_{\eta}^{-1}(\pi_0) \end{array} \right] \\
\times \frac{\sqrt{T}}{T} \sum_t \left[ s_{\theta t}(y_t; \theta_0) \right. \\
\left. + s_{nt}(y_{ht}|y_{mt}; \pi_0) \right],
\]
we can straightforwardly obtain the result of the proposition.

\section*{B Form of the conditional mean and variance under the GH alternative}

Consider the GH standardised random vector \( u \) described in section 3. Blæsild (1981) shows that, if we partition this vector as \( u = [u_1', u_2']' \), the distribution of \( u_2 \), conditional on \( u_1 \) is also GH, with mean and variance
\[ E(u_2|u_1) = \alpha_2(\beta, \eta, \psi) + \Xi_{21}(\beta, \eta, \psi)\Xi_{11}^{-1}(\beta, \eta, \psi)[u_1 - \alpha_1(\beta, \eta, \psi)] \]

\[ + \frac{\gamma\delta^*(u_1)R_{\nu-\frac{1}{2}}(\delta^*(u_1)\gamma^*)}{\gamma^*R_{\nu}(\gamma)}\Xi_{21}(\beta, \eta, \psi)\beta_2 \]

\[ V(u_2|u_1) = \frac{\gamma\delta^*(u_1)R_{\nu-\frac{1}{2}}[\delta^*(u_1)\gamma^*]}{\gamma^*R_{\nu}(\gamma)}\Xi_{21}(\beta, \eta, \psi) \]

\[ + R_{\nu-\frac{1}{2}+1}[\delta^*(u_1)\gamma^*]R_{\nu-\frac{1}{2}}[\delta^*(u_1)\gamma^*] \left[ \frac{\gamma\delta^*(u_1)}{\gamma^*R_{\nu}(\gamma)} \right]^2 \Xi_{21}(\beta, \eta, \psi)\beta_2\beta_2'\Xi_{21}(\beta, \eta, \psi) \]

\[ - \left[ \frac{\gamma\delta^*(u_1)R_{\nu-\frac{1}{2}}[\delta^*(u_1)\gamma^*]}{\gamma^*R_{\nu}(\gamma)} \right]^2 \Xi_{21}(\beta, \eta, \psi)\beta_2\beta_2'\Xi_{21}(\beta, \eta, \psi) \]

where \( \gamma = \psi^{-1} - 1 \), \( \delta^*(u_1) = \sqrt{\xi_1\gamma^{-1}R_{\nu}(\gamma) + 1} \),

\[ \gamma^* = \sqrt{\beta'} \left[ \frac{\Xi_{11}(\beta, \eta, \psi)}{\Xi_{21}(\beta, \eta, \psi)} \right] \Xi_{11}^{-1}(\beta, \eta, \psi) \left[ \Xi_{11}(\beta, \eta, \psi) - \Xi_{21}(\beta, \eta, \psi) \right] \beta \frac{\gamma}{R_{\nu}(\gamma)} + \gamma^2. \]

and \( \xi_m = (u_1 - \alpha_1(\beta, \eta, \psi))'\Xi_{11}^{-1}(\beta, \eta, \psi)(u_1 - \alpha_1(\beta, \eta, \psi)) \). Note that we have partitioned \( \beta \) as \((\beta_1', \beta_2')', \alpha(\beta, \eta, \psi) \) as \((\alpha_1'(\beta, \eta, \psi), \alpha_2'(\beta, \eta, \psi))' \),

\[ \Xi(\beta, \eta, \psi) = \left[ \begin{array}{cc} \Xi_{11}(\beta, \eta, \psi) & \Xi_{21}'(\beta, \eta, \psi) \\ \Xi_{21}(\beta, \eta, \psi) & \Xi_{22}(\beta, \eta, \psi) \end{array} \right] \]

and we have defined \( \Xi_{21}(\beta, \eta, \psi) = \Xi_{22}(\beta, \eta, \psi) - \Xi_{21}(\beta, \eta, \psi)\Xi_{11}^{-1}(\beta, \eta, \psi)\Xi_{21}'(\beta, \eta, \psi) \).

**C Distribution of the proportion of funds failing the tests**

Our simulation method involves two main steps. First, we estimate the distribution of the market return. Secondly, with those estimates, we generate \( s = 1 \cdots S \) replications of the market index. We have generated these simulations from the Gaussian distribution, but more flexible densities can also be used. For each of those replications, we compute the 5 dependence tests between each hedge fund, \( j = 1 \cdots F \), and the market, obtaining an indicator function \( \iota_{is} = 1(\tau_{is} > \tau_{0.95}) \) that takes a value of one if the test is rejected, and zero otherwise. As the sample size grows, \( \iota_{is} \) will converge in distribution to a Bernoulli variate with a 5% probability of 1. Hence, the distribution of the proportion of test rejections can be obtained from the empirical distribution of these simulated tests.

Importantly, this method does not make any assumption about the distribution of hedge fund returns. Hence, it accounts for the fact that not only correlation but also non-linear
dependence between the funds can increase the proportion of rejections even when the null hypothesis is true.
### Table 1a

Percentage of funds with test rejections at the 95% level.

Bivariate triangular VAR(1) between each fund and the S&P 500.

<table>
<thead>
<tr>
<th>Hedge fund strategies</th>
<th>Correlation</th>
<th>Non-linear</th>
<th>Ability</th>
<th>Inability</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund of Funds</td>
<td>55.9</td>
<td>23.8</td>
<td>20.8</td>
<td>30.7</td>
<td>24.5</td>
</tr>
<tr>
<td>Event driven</td>
<td>79.8</td>
<td>34.5</td>
<td>28.6</td>
<td>38.1</td>
<td>33.3</td>
</tr>
<tr>
<td>Market Neutral</td>
<td>35.9</td>
<td>14.1</td>
<td>19.6</td>
<td>19.6</td>
<td>21.7</td>
</tr>
<tr>
<td>Equity non Hedge</td>
<td>87.3</td>
<td>10.1</td>
<td>16.5</td>
<td>16.5</td>
<td>12.7</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>64.8</td>
<td>14.1</td>
<td>17.7</td>
<td>17.7</td>
<td>18.0</td>
</tr>
</tbody>
</table>

### Table 1b

Percentage of funds with test rejections at the 95% level.

Trivariate triangular VAR(1) between each fund and the S&P 500 and FT 100.

<table>
<thead>
<tr>
<th>Hedge fund strategies</th>
<th>Correlation</th>
<th>Non-linear</th>
<th>Ability</th>
<th>Inability</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund of Funds</td>
<td>62.5</td>
<td>20.4</td>
<td>24.4</td>
<td>28.0</td>
<td>25.8</td>
</tr>
<tr>
<td>Event driven</td>
<td>79.1</td>
<td>29.9</td>
<td>29.9</td>
<td>38.8</td>
<td>32.8</td>
</tr>
<tr>
<td>Market Neutral</td>
<td>27.5</td>
<td>5.8</td>
<td>8.7</td>
<td>13.0</td>
<td>11.6</td>
</tr>
<tr>
<td>Equity non Hedge</td>
<td>86.8</td>
<td>10.3</td>
<td>11.8</td>
<td>16.2</td>
<td>13.2</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>69.5</td>
<td>10.4</td>
<td>12.5</td>
<td>14.0</td>
<td>14.3</td>
</tr>
</tbody>
</table>

### Table 1c

Percentage of funds with test rejections at the 95% level.

Bivariate diagonal AR(1)-CCC-GARCH(1,1) between each fund and the S&P 500.

<table>
<thead>
<tr>
<th>Hedge fund strategies</th>
<th>Correlation</th>
<th>Non-linear</th>
<th>Ability</th>
<th>Inability</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund of Funds</td>
<td>62.7</td>
<td>26.8</td>
<td>20.2</td>
<td>33.1</td>
<td>24.4</td>
</tr>
<tr>
<td>Event driven</td>
<td>79.7</td>
<td>37.5</td>
<td>32.8</td>
<td>43.8</td>
<td>37.5</td>
</tr>
<tr>
<td>Market Neutral</td>
<td>31.4</td>
<td>17.1</td>
<td>24.3</td>
<td>21.4</td>
<td>20.0</td>
</tr>
<tr>
<td>Equity non Hedge</td>
<td>86.9</td>
<td>13.1</td>
<td>13.1</td>
<td>18.0</td>
<td>13.1</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>68.9</td>
<td>14.0</td>
<td>13.7</td>
<td>18.2</td>
<td>16.5</td>
</tr>
</tbody>
</table>

Notes: Triangular VAR(1) models allow for a feedback effect of lagged values of the market variable(s) on the conditional means of Hedge fund returns. Tables 1a and 1c: funds $T \geq 30$. Table 1b: funds with $T \geq 50$. Except for the zero correlation test, the null is linear dependence in all cases. The last three columns report the proportion of rejections from testing the null against market timing (Ability), market mis-timing (Inability) and non-informative signals (Noise).
Table 2a

95th percentile of the distribution of proportion of funds failing the tests under the null.

Bivariate triangular VAR(1) between each fund and one simulated index.

<table>
<thead>
<tr>
<th>Hedge fund strategies</th>
<th>Correlation</th>
<th>Non-linear</th>
<th>Ability</th>
<th>Inability</th>
<th>Noise</th>
<th>Binomial (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund of Funds</td>
<td>13.3</td>
<td>17.3</td>
<td>17.8</td>
<td>17.2</td>
<td>17.8</td>
<td>6.9 (404)</td>
</tr>
<tr>
<td>Event driven</td>
<td>17.8</td>
<td>19.0</td>
<td>20.6</td>
<td>21.9</td>
<td>21.4</td>
<td>9.5 (84)</td>
</tr>
<tr>
<td>Market Neutral</td>
<td>10.8</td>
<td>11.9</td>
<td>12.1</td>
<td>12.3</td>
<td>12.0</td>
<td>8.6 (92)</td>
</tr>
<tr>
<td>Equity non Hedge</td>
<td>15.1</td>
<td>13.9</td>
<td>16.3</td>
<td>16.5</td>
<td>17.5</td>
<td>8.8 (79)</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>12.1</td>
<td>14.3</td>
<td>16.3</td>
<td>14.3</td>
<td>15.4</td>
<td>6.7 (412)</td>
</tr>
</tbody>
</table>

Table 2b

95th percentile of the distribution of proportion of funds failing the tests under the null.

Trivariate triangular VAR(1) between each fund and two simulated indices.

<table>
<thead>
<tr>
<th>Hedge fund strategies</th>
<th>Correlation</th>
<th>Non-linear</th>
<th>Ability</th>
<th>Inability</th>
<th>Noise</th>
<th>Binomial (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund of Funds</td>
<td>16.0</td>
<td>16.0</td>
<td>21.9</td>
<td>21.1</td>
<td>22.0</td>
<td>7.2 (275)</td>
</tr>
<tr>
<td>Event driven</td>
<td>14.9</td>
<td>17.9</td>
<td>20.4</td>
<td>22.7</td>
<td>23.0</td>
<td>8.9 (67)</td>
</tr>
<tr>
<td>Market Neutral</td>
<td>11.5</td>
<td>10.1</td>
<td>11.9</td>
<td>12.4</td>
<td>12.2</td>
<td>10.1 (69)</td>
</tr>
<tr>
<td>Equity non Hedge</td>
<td>14.7</td>
<td>14.7</td>
<td>16.3</td>
<td>17.6</td>
<td>17.4</td>
<td>10.2 (68)</td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>13.6</td>
<td>13.9</td>
<td>15.2</td>
<td>15.3</td>
<td>15.9</td>
<td>7.1 (279)</td>
</tr>
</tbody>
</table>

Notes: Tests computed between hedge fund returns and 1,000 replications of the market return, obtained by simulation from the Gaussian distribution under the null of independence with respect to hedge fund data. The column labelled binomial reports the 95% percentile under independence of hedge fund returns, with the number of funds in each style in parenthesis. Except for the zero correlation test, the null is linear dependence in all cases. The last three columns report critical values of tests of the null against market timing (Ability), market mis-timing (Inability) and non-informative signals (Noise).
Table 3
Dependence tests between 10 CSFB-Tremont Hedge fund style indices and the Market
Triangular VAR(1)-GARCH(1,1) with constant conditional correlation

<table>
<thead>
<tr>
<th>Test</th>
<th>Market: S&amp;P 500</th>
<th>Market: S&amp;P 500 and FT 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian (linear)</td>
<td>49.570 (0.000)</td>
<td>30.874 (0.003)</td>
</tr>
<tr>
<td>Student $t$</td>
<td>6.579 (0.858)</td>
<td>16.128 (0.214)</td>
</tr>
<tr>
<td>Student $t$ vs. Ability</td>
<td>3.601 (0.623)</td>
<td>7.387 (0.275)</td>
</tr>
<tr>
<td>Student $t$ vs. Inability</td>
<td>5.576 (0.376)</td>
<td>10.154 (0.134)</td>
</tr>
</tbody>
</table>

Notes: Triangular VAR(1) models allow for a feedback effect of lagged values of the market variable on
the conditional means of Hedge fund returns. Gaussian and Student $t$ reported dependence tests are
partially one-sided. Sample size $T = 136$. 
Figure 1: Examples of shapes of the conditional mean and variance under Gaussian, Student $t$ and GH dependence structures

(a) $E(y_h | y_m)$

(b) $V(y_h | y_m)$

Type of Dependence: Gaussian Student $t$ GH GH
Market timing signal: None Noise Biased Unbiased

Note: in all cases $E(y_h) = E(y_m) = 0$, $V(y_h) = V(y_m) = 1$, and $\text{cov}(y_h, y_m) = 0$
Figure 2: Power of the test of linearity of mean and constancy of variance

(a) Symmetric alternatives

(b) Asymmetric alternatives $b = (1, 1)'$

$k = 1, N = 1$

$k = 1, N = 5$

$k = 5, N = 1$

$k = 10, N = 1$

Figure 3: Power of the joint test of linearity of mean, constancy of variance and absence of correlation

(a) Symmetric alternatives

(b) Asymmetric alternatives $b = (1, 1)'$

$\rho = 0$

$\rho = 0.1$

$\rho = 0.2$

Notes: Size= 5%, $T=100$, $\psi = 1$. In figure 3, $k = N = 1$ in all cases.
Figure 4: p-value discrepancy plots of the tests of zero correlation and linear dependence under the null of linear dependence

Gaussian VAR(1) with constant variance
(a) $\rho = 0$

(b) $\rho = 0.1$

Non-Gaussian VAR(1) with constant variance
(c) $\rho = 0$

(d) $\rho = 0.1$

Gaussian VAR(1)-CCC-GARCH(1,1)
(e) $\rho = 0$

(f) $\rho = 0.1$

Absence of Correlation Linear dependence Joint test

Notes: p-value discrepancy plots obtained from a Monte Carlo study with 30,000 simulations of samples with $T = 100$ under the null of linear dependence. In the non-Gaussian model the marginal distributions are Student $t$ with 10 degrees of freedom. $\rho$ denotes linear correlation.
Figure 5: p-value discrepancy plots of the tests of zero correlation and linear dependence under non-linear dependence

Symmetric Generalised Hyperbolic dependence: $\beta = (0, 0)'$, $\psi = 1$

(a) Test of absence of correlation

(b) Test of linear dependence

Asymmetric Generalised Hyperbolic dependence: $\eta = 0.1, \psi = 1$

(c) Test of absence of correlation

(d) Test of linear dependence

Clayton copula dependence

(e) Test of absence of correlation

(f) Test of linear dependence

Notes: p-value discrepancy plots obtained from a Monte Carlo study with 30,000 simulations of samples with $T = 100$ under non-linear dependence. Gaussian VAR(1) with constant variance. Linear correlation is zero in figures (a) to (d). $\kappa$ is the parameter of the Clayton copula.
Figure 6: Kernel estimation of the density of the historical standardised shocks to the S&P 500 from 1994 to 2005
Figure 7: Mean-variance frontier of $y^\perp_{ht+1}$ conditional on the market return $y_{m t+1}$ in August 1998

(a) Very bad scenario:
First quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t^5(y_{m t+1})$

(b) Bad scenario:
Second quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t^5(y_{m t+1})$

(c) Good scenario:
Third quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t^5(y_{m t+1})$

(d) Very good scenario:
Fourth quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t^5(y_{m t+1})$

Notes: $y^\perp_{ht+1}$ denotes the 10 monthly returns of the CSFB-Tremont style indices, neutralised with respect to the market in terms of conditional correlation. Monthly returns of the S&P 500 index are denoted with $y_{m t+1}$. Portfolio weights are restricted to be between 0 and 1. Conditional means and variances are modelled with a VAR(1)-GARCH(1,1) parametrisation with constant conditional correlation that allows for a feedback effect of lagged observations of $y_m$ on the conditional means of hedge funds index returns. For each scenario, we consider that the deviation of $y_{m t+1}$ from its expected return will be on each of the quartiles of the empirical distribution of deviations. This analysis is conditional on information up to July 1998, though parameters are estimated with the whole sample.
Figure 8: Mean-variance frontier of $y_{h t+1}$ conditional on the market return $y_{m t+1}$ in August 2005

(a) Very bad scenario:
First quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t(y_{m t+1})$

(b) Bad scenario:
Second quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t(y_{m t+1})$

(c) Good scenario:
Third quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t(y_{m t+1})$

(d) Very good scenario:
Fourth quartile of $[y_{m t+1} - E_t(y_{m t+1})]/V_t(y_{m t+1})$

Notes: $y_{h t+1}$ denotes the 10 monthly returns of the CSFB-Tremont style indices, neutralised with respect to the market in terms of conditional correlation. Monthly returns of the S&P 500 index are denoted with $y_{m t+1}$. Portfolio weights are restricted to be between 0 and 1. Conditional means and variances are modelled with a VAR(1)-GARCH(1,1) parametrisation with constant conditional correlation that allows for a feedback effect of lagged observations of $y_m$ on the conditional means of hedge funds index returns. For each scenario, we consider that the deviation of $y_{m t+1}$ from its expected return will be on each of the quartiles of the empirical distribution of deviations. This analysis is conditional on information up to July 2005.