Semi-Parametric Estimation and Simulation of Actively Managed Portfolios

Umberto Cherubini * Fabio Gobbi † Sabrina Mulinacci ‡

September 7, 2010

Abstract

In this paper we propose a copula-based technique to recover the distribution of actively managed funds. The copula is meant to represent the dependence structure between the market return (or in general the benchmark) and the investment strategy of the asset manager. The analysis is carried out in a rational investor economy with managed funds, such as that in Merton (1981) and Berk and Green (2004). The distribution of returns on any managed fund turns out to be represented by: i) a marginal distribution representing the asset manager stock picking ability; ii) a copula function representing market timing activity.

Keywords: copula, inference for margins, market timing, stock-picking

1 Introduction

Consider a managed fund Z promising to yield an excess return with respect to a benchmark index X. Assume we know the dynamics and the distribution of the return of X. The distribution of the return on the managed fund Z will depend on the investment policy implemented by the fund manager that will add up to the return on the benchmark. Call Y this component of the return. In principle, Y contains all information needed to describe the management style of the fund. The mean of Y, conditional on X is what is universally known as α of the fund, and it is a measure of the stock-picking ability of the manager. The market timing activity instead impacts on the dependence structure between the

*University of Bologna, Department of Mathematical Economics, Viale Filopanti 5, 40126 Bologna, Italy. Phone: +(39) 0512094370; Fax: +(39) 0512094357. e-mail: umberto.cherubini@unibo.it

†University of Bologna, Department of Mathematical Economics, Viale Filopanti 5, 40126 Bologna, Italy. Phone: +(39) 0512094368; Fax: +(39) 0512094357. e-mail: fabio.gobbi@unibo.it

‡University of Bologna, Department of Mathematical Economics, Viale Filopanti 5, 40126 Bologna, Italy. Phone: +(39) 0512094368; Fax: +(39) 0512094357. e-mail: sabrina.mulinacci@unibo.it

benchmark and the fund. The most straightforward way to see this is to think of active market timing as the introduction of a portfolio of options written on the benchmark, allowing to increase and reduce leverage in periods of higher or lower return on the benchmark. This option based approach was first proposed by Merton (1981) and Henriksson and Merton (1981).

A practical issue that is of utmost relevance is how the distribution of the investment strategy and its dependence with the benchmark impact on the distribution of the returns on the managed fund. In this paper we propose a semi-parametric approach based on copula functions to disentangle the effect of market timing policies on the probability distribution of the returns. The methodology can be applied in two opposite instances: i) if you know the marginal distribution of the investment strategy and its dependence structure with the market (or the benchmark) you can recover both the marginal distribution of the return on the fund and its dependence on the market; ii) if you know the distribution of the return on the fund and its dependence with the market, you can extract the marginal distribution of the return from the investment strategy and its dependence with the market (that is its market timing component).

The analysis is developed within a model of a rational economy with managed funds, such as that assumed in the background of the Henriksson and Merton approach and that formalized in Berk and Green (2004). In this economy, excess returns of managed fund share the same statistical properties of publicly traded assets, that is they cannot be predicted using available information. In this framework of efficient market in semi-strong form we can exploit the statistical model proposed in Cherubini, Mulinacci and Romagnoli (2009).

The structure of the paper is as follows. Section 2 introduces the market model and the motivation of the paper. In Section 3 the copula concept is applied to our problem of disentangling the part of return due to market movements (passive return) from that due to the management strategy implemented by the asset manager. In Section 4 we recover the probability distribution of the return of a managed fund on a given investment horizon using two pieces of information: stock picking skill and market timing skill. Moreover the HM copula is obtained. In Section 5 we address the opposite problem of that of the Section 4: we assume we are given the time series of returns on a fund and that of the market and we back out the record of the asset manager that is running that fund. In Section 6 we show how to estimate the return of a managed fund. More precisely, we estimate via IFM technique the parameters of the copula and of the distribution of Y. In Section 7 we present an empirical application. Section 8 concludes and Section 9 contains the derivation of the HM copula.

2 Market model and motivation

Whenever we evaluate the performance of a managed fund we do so on the background of a model of the capital market. In this paper we use a standard model with rational investors and efficient markets. The rational market model with mutual funds that we have in mind was presented in Berk and Green (2004). In that model, competitive equilibrium ensures that capital flows to efficient asset managers up to the point where expected risk adjusted excess returns equal zero across all possible investments, no matter whether passive and managed.

The starting point is a linear decomposition of the return Z of a managed portfolio in a passive component X and another one due to management decisions Y

$$Z = X + Y \tag{1}$$

The return component Y contains all information needed to characterize and measure the effects of the asset management strategy. In particular, we may single out three hypotheses

- Passive management - Y is independent of X and has zero mean
- Stock picking - Y is independent of X and has non zero mean (called α)
- Market timing - Y is dependent on X

In the standard literature on performance analysis, the asset management strategy is evaluated carrying out a regression. The most famous cases were proposed by Treynor and Matsuy (1966) and Henriksson and Merton (1981).
A problem with this kind of approach is that the regression representation is based on quite restrictive assumptions concerning both the kind of strategy followed by the manager and the institutional environment in which the asset manager operates, such as the kind of compensation scheme. So, for example, the Henriksson-Merton model above is based on the assumption that the asset manager switches the asset allocation between the risk-free investment and the risk-free rate or vice versa whenever she expects a rate of return on the asset higher than the risk-free rate (or vice versa). This is equivalent to a portfolio insurance strategy which uses put options. Of course, even though this representation is sufficient to induce non linearity in the relationship between the rate of return on the managed fund and the market, it remains too simple from the point of view of the management process, that can be path-dependent and may be inspired by different strategies in different periods. For these reasons, non-parametric tests (Jiang, 2003) and graphical (chartist) analysis (Leigh, Paz and Purvis, 2002) have been proposed to assess the timing activity. But beyond tests, one remains with the need to estimate and simulate the joint returns of a managed fund and its reference benchmark, and fully non parametric techniques are not very well suited to accomplish this.
A semi-parametric approach may strike a balance between the need to represent in a general way positive association between the returns on the market and the returns accrued to that by the management activity and a non-parametric representation of the marginal distribution of the managed fund. This semi-parametric approach (in the spirit of Chen and Fan, 2006) applies copula functions to model dependence among the variables and non parametric analysis

to model their marginal distributions. As for copula functions, while we refer the interest reader to Nelsen (2006) for details, we only remind that the copula technique allows to write every joint distribution as a function of marginal distributions. In our case, we can represent the joint distribution of X and Y, say $\Pr(X \leq a, Y \leq b)$, with $a, b \in \Re$ as a function of $F_X(a) \equiv \Pr(X \leq a)$ and $F_Y(a) \equiv \Pr(Y \leq b)$. More formally, there exists a function $C_{X,Y}(u, v)$ such that

$$Pr(X \leq a, Y \leq b) = C_{X,Y}(F_X(a), F_Y(b)) \tag{2}$$

Conversely, given two distribution functions F_X and F_Y and a suitable bivariate function $C_{X,Y}$ we may build joint distribution for the returns. This one to one relationship between joint distributions and copula functions is known as Sklar theorem.
Then, using copula functions we can separately specify the distribution of market returns and those due to the management strategy, and then apply a copula that represents their dependence. From the hypotheses above, the market timing activity generates dependence between the returns due to market movements and those due to the management strategy. So, passive strategies and strategies based on stock picking only are consistent with the product copula $C_{X,Y}(u, v) = uv$ corresponding to independence, and we have

$$Pr(X \leq a, Y \leq b) = F_X(a)F_Y(b) \tag{3}$$

In the general case in which both stock picking and market timing are present, the copula function approach allows to separate the two activities by modelling the returns from stock picking with the distribution F_Y and those from market timing with the copula function $C_{X,Y}$.

3 Passive and active fund management

We are now going to apply the copula concept to our problem of disentangling the part of return due to market movements (passive return) from that due to the management strategy implemented by the asset manager. Before doing that, let us notice that what makes the problem more involved is that the problem is actually tri-variate. We are in fact investigating the relationship among: i) the return on the market; ii) the return on the investment strategy; iii) the return on the fund. Furthermore, we know that the latter is actually the sum of the first two; unfortunately, linearity is by no means a feature that simplifies the analysis, and actually is what makes the issue more involved. The reason is that the relationship between the two variables in the sum is non-linear by definition (at least if we account for market timing).

So, from a statistical point of view, we are in face of a tri-variate compatibility problem. In statistics, the compatibility problem refers to the possibility to build a joint distribution of dimension n with marginals of dimension $n - k$, $k = 1, ..., n - 1$ assigned. Joe (1997) reports some cases of compatibility in dimension three. In our case, we have a compatibility problem in the same dimension, namely among variables Z, X and Y. So, for example, given an assigned relationship between the return on the market and that on the strategy, we must recover a compatible dependence between the market and the fund. The example can be also reversed and assuming that we know the dependence structure between the fund and the market we want to recover that between the market and the investment strategy. In one word, we have a compatibility problem between. This compatibility problem was solved in Cherubini, Mulinacci and Romagnoli (2009). We report here the proposition, referring to the paper for full proof.

Proposition 3.1. *Let X e Y be two real-valued random variables on the same probability space $(\Omega, \Im, \mathbb{P})$ with corresponding copula $C_{X,Y}$ and continuous marginals F_X and F_Y. With $D_1C_{X,Y}(u,v)$ we denote $\frac{\partial C_{X,Y}(u,v)}{\partial u}$. Then,*

$$F_{X+Y}(z) = \int_0^1 D_1 C_{X,Y}\left(w, F_Y(z - F_X^{-1}(w))\right) dw \tag{4}$$

and

$$C_{X,X+Y}(u,v) = \int_0^u D_1 C_{X,Y}\left(w, F_Y(F_{X+Y}^{-1}(v) - F_X^{-1}(w))\right) dw. \tag{5}$$

So, if we are able to model the return on the management strategy conditional on the movements of the market, we can compute both the distribution of the managed fund and its dependence with the market. More precisely, the return on the fund is what Cherubini, Mulinacci and Romagnoli (2009) call *C-convolution*, that is the convolution of two variables with dependence structure represented by a copula function C.

4 The distribution of managed funds returns

Using the compatibility result above, we address the following question. Assume a new fund management program is launched on the market. We would like to estimate the probability distribution of its return on a given investment horizon. Given the analysis above, in order to accomplish this we need two pieces of information:

1. **Stock picking skill**. This is represented by the distribution of the return on the investment strategy.
2. **Market timing skill.** This is represented by a copula function linking the return from the investment strategy

Of course a problem is that if the product has to be launched, this information is not directly available. It is part of the fund analyst to make realistic hypothesis about this information or to use reasonable proxies for such information. Typically, an idea of the investment strategy can be extracted from

- Statements concerning the aims and the investment process issued by the corporate entity issuing the fund in official reports and in road shows to the investors. If for example the management states they would hire a team sector analysts rather than economists to design the investment strategy of the fund, one could in principle design the investment strategy as pure stock picking activity.
- The record of the asset manager hired to run the fund in prior assignments. By record we mean actually the probability distribution of her investment strategy and its co-movement with the market.
- Automatic trading strategies. One could conceive very simplified strategies, such as the portfolio insurance strategy used in the Henriksson and Merton model, or more involved strategies based on simulation.

In the end, the analyst should come up with a probability distribution representing stock picking and a copula function representing market timing.

4.1 Pure stock-picking investment strategies

If one can reasonably conclude, either from the statements of the issuer of the fund or the past investment strategy record of its manager that she would most probably concentrate on a stock picking strategy (or *microeconomic* strategy quoting the term used by Merton), one could immediately provide a more specific representation of the distribution of the return on the fund and its comovement with the market.

In fact, the probability distribution of returns on managed funds that do not engage in market timing is given by

$$F_{X+Y}(z) = \int_0^1 F_Y(z - F_X^{-1}(w))dw. \tag{6}$$

that is the standard convolution of the returns due to market movements and those due to the market strategy. Furthermore, the dependence structure between the fund and the market is

$$C_{X,X+Y}(u,v) = \int_0^u F_Y(F_{X+Y}^{-1}(v) - F_X^{-1}(w))dw. \tag{7}$$

So, the problem would be completely solved once some assumption would have been made concerning the marginal distribution of Y.

4.2 HM-copula

As an example, assume the manager would engage in a market timing strategy based on portfolio insurance as specified in Henriksson and Merton (1981). As it is well known, this model leads to a specification of the stratefy Y as

$$Y = \alpha + \gamma \max(-X, 0) + \epsilon$$

where ϵ is a zero mean disturbance. As for the γ parameter, this is proportional to the concordance between the forecast of the manager and actual movements of the market.

Even though we extend a friendly advice to evaluate the distribution of the return on the fund by simulation, one could also derive the specific copula $C_{X,Y}$ representing the market timing activity in this model. We report here the copula, that we called *HM copula* for obvious reasons, referring the reader to the appendix for the derails of the derivation

Now, the copula between X and Y is given by

$$C_{X,Y}(u,v) = G_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v))\mathbf{1}_{\{u>F_X(0)\}} + H_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v))\mathbf{1}_{\{u\leq F_X(0)\}}$$

where

$$G_{X,Y}(x,y) = F_X(x)F_\epsilon(y-\alpha) - \int_{-\infty}^{y-\alpha} F_X(\frac{e-y+\alpha}{\gamma})F_\epsilon(de)$$

and

$$H_{X,Y}(x,y) = F_X(x)F_\epsilon(\gamma x+y-\alpha) - \int_{-\infty}^{\gamma x+y-\alpha} F_X(\frac{e-y+\alpha}{\gamma})F_\epsilon(de).$$

The distribution of the fund is finally recovered as

$$F_Z(z) = \int_0^{F_X(0)} F_\epsilon(z+(\gamma-1)F_X^{-1}(w)-\alpha)dw + \int_{F_X(0)}^1 F_\epsilon(z-F_X^{-1}(w)-\alpha)dw$$

However the explicit derivation of the copula function enables to compute explicitly the shape of the density function and the dependence between the market and the investment strategy. In fig. 1. we report the shape of the probability density function for different levels of the γ parameter. As expected, the higher the parameter, the more effective the portfolio insurance strategy is, and the lower is the left tail of the distribution.

As for the dependence structure between investment strategy and the market, in fig. 2 we report the value of the corresponding Kendall τ statistic, as a function of different values of the γ parameter. Notice that the γ parameter in the Merton model is a measure of concordance between forecast of the asset manager and market movements. So, the figure shows how this concordance measure translates into a concordance measure (the Kendall τ) between the return on the investment strategy and the market. Since the market timing strategy implies a position in *protective* put options, it is not surprising that association is negative. The more effectively an asset manager is able to forecast future market movements the more effective the protective put position is.

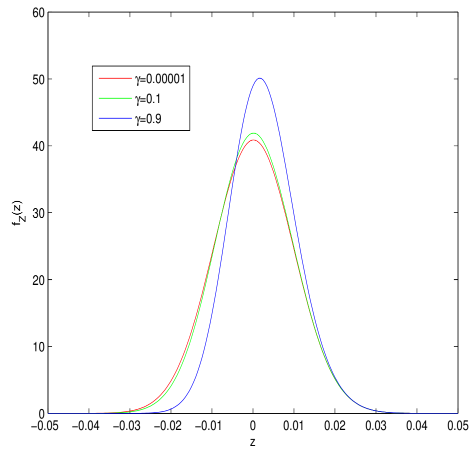

Figure 1: Probability density function of $Z = X + Y$ for different levels of the γ parameter.

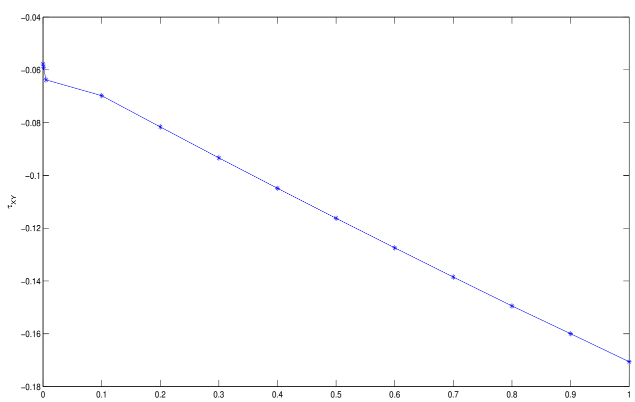

Figure 2: Kendall τ statistic as a function of different values of the γ parameter.

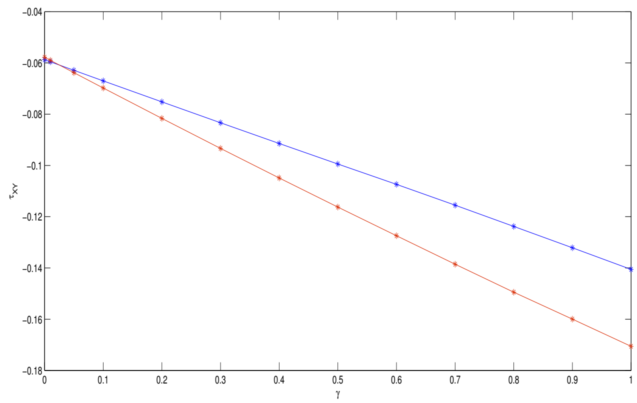

Figure 3: Kendall τ statistic as a function of different values of the γ parameter when σ_ϵ increases by 40 percent (blue line)

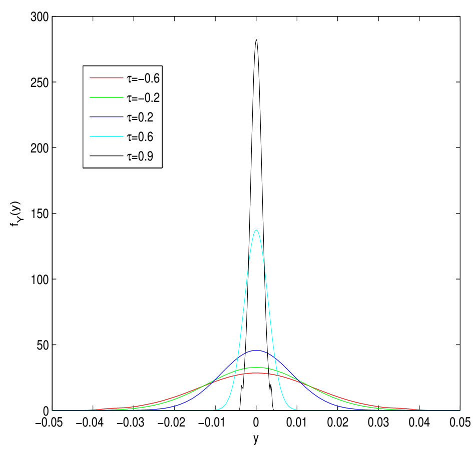

Figure 4: Probability density function of the management strategy Y consistent with selected dependence structures between Z and X: C_{ZX} is a Gaussian copula.

5 The distribution of the return on the investment strategy

In this section we address the opposite problem of that of the previous section. We assume we are given the time series of returns on a fund and that of the market, and we want to back out the record of the asset manager that is running that fund. As in the section before, the record is made up by two components: i) a probability distribution representing stock picking skills; ii) a copula function representing market timing skills.

The solution to this problem can be again found in the result in section 2. Trivially, the problem is to find the probability distribution of $Y = Z - X$ given the dependence structure between Z and X. We then have:

$$F_Y(t) = \int_0^1 D_1 C_{Z,-X}\left(w, F_{-X}(t - F_Z^{-1}(w))\right) dw. \tag{8}$$

It is well known that $C_{Z,-X}$ can be recovered directly from $C_{Z,X}$ using the invariance relationship $C_{Z,-X}(u,v) = u - C_{Z,X}(u, 1-v)$. Therefore, the stock picking skill of the asset manager can be recovered as

$$F_Y(t) = 1 - \int_0^1 D_1 C_{Z,X}\left(w, F_X(F_Z^{-1}(w) - t)\right) dw.$$

Furthermore, the market timing activity of the asset manager will be fully described by the copula function $C_{X,Y}$, that is

$$C_{X,Y}(u,v) = u - \int_0^u D_1 C_{Z,X}\left(w, F_X(F_Z^{-1}(w) - F_Y^{-1}(v))\right) dw \tag{9}$$

In figure 3 we report the density of the management strategy Y consistent with selected dependence structures between Z and X: we namely use Gaussian copulas with increasing dependence (Kendall $\tau = -0.6, ..., 0.9$). In figure 4 we compare the density of Y for the Gaussian copula, the Clayton copula and the Frank copula, with parameters consistent with the same figure of Kendall τ equal to 0.6.

Finally, figure 5 reports the market timing activity measured in terms of Kendall τ. In other terms, we report the measure of association between the return on the investment strategy and the market consistent with the observed association between the return on the fund and the market. For the sake of comparison, we report the Gaussian copula, the Clayton copula and the Frank copula.

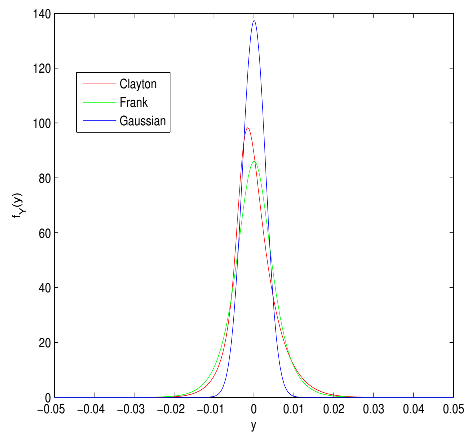

Figure 5: Probability density function of Y for the Gaussian copula, the Clayton copula and the Frank copula, with parameters consistent with the Kendall τ equal to 0.6.

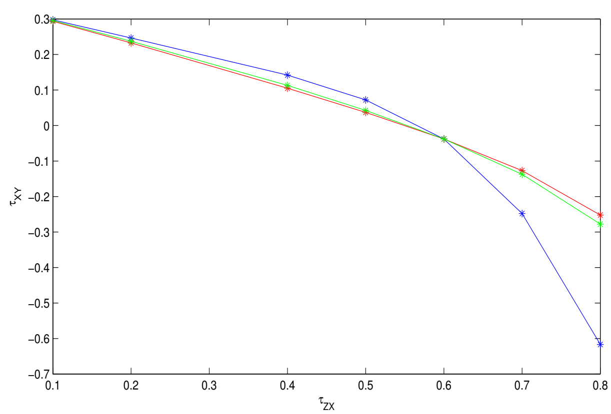

Figure 6: Kendall τ between the return on the investment strategy and the market consistent with the observed Kendall τ between the return on the fund and the market in the case of Clayton copula (red line), Frank copula (green line) and Gaussian copula (blue line).

5.1 Market neutral investment(hedge funds?)

A very interesting specific case of the analysis above applies to so called market neutral funds. It is well known that hedge funds should by definition be endowed with this feature. Market neutrality has been recently questioned on empirical grounds, and for this reason we use a question mark in the title. Anyway, if we were facing a genuine market neutral fund management policy, we could use the technique above to measure the skill of the market neutral asset manager. As a matter of fact, one would simply have to set $C_{Z,X} = uv$ and we would simply obtain

$$F_Y(t) = 1 - \int_0^1 F_X(F_Z^{-1}(w) - t)dw.$$

for the stock picking strategy and

$$C_{X,Y}(u,v) = u - \int_0^u F_X(F_Z^{-1}(w) - t)dw.$$

for the market timing activity.
It is also quite straightforward to describe the distribution of the so called *core-satellite* strategies, in which a small investment in market neutral funds is added to a position in mutual funds to add some α while decreasing correlation in the portfolio. Namely, if a hedge fund H is added to the fund Z, the distribution of the return would be given by

$$F_{\overline{Z}}(t) = \int_0^1 F_H(t - F_Z^{-1}(w)))dw. \tag{10}$$

where $\overline{Z}$ denotes the *core-satellite* portfolio. The gain in diversification can be gauged measuring.

$$C_{Z,H}(u,v) = \int_0^u F_H(F_{\overline{Z}}^{-1}(v) - F_Z^{-1}(w)))dw. \tag{11}$$

6 Estimation

In this section we show how to estimate the return of a managed fund. In other words, we derive the statistical tools to extract the record of an asset manager from market data, and to use this record to simulate the future return on the fund. Here we assume absolutely continuous copula functions and marginal distributions and we will denote with small case letters the corresponding densities.

6.1 Inference function for margin (IFM) estimation

The inference function for margin (IFM) method (Joe and Xu (1996)) is based on the fact that the log-likelihood function of the joint density may be decomposed into three positive terms two of which involving the margins and its parameter only.

For a start, we write the likelihood function of the problem, in order to estimate the parameters of the model. We remind that they are three: i) α, which represents stock-picking ability; ii) the θ parameter, or a set of parameters of the copula function that models dependence between Y and X; iii) volatility of Y, denoted by the standard error σ_Y, which denotes the *tracking error* of the management policy. We remind the assumptions that these variables are endowed with parametric continuous distribution functions $F_X(x;\boldsymbol{\alpha}_X)$, $F_Y(y;\boldsymbol{\alpha}_Y)$ and density functions $f_X(x;\boldsymbol{\alpha}_X)$ and $f_Y(y;\boldsymbol{\alpha}_Y)$. The vectors $\boldsymbol{\alpha}_X$ and $\boldsymbol{\alpha}_Y$ contain the parameters of the margins. We assume that the copula function $C_{X,Y}$ be indexed by a set θ of parameters. In most of the copula functions that are commonly used, this set contains a single parameter, in which case the notation θ refers to that parameter. We rewrite equations (4) and (5) in order to emphasize the dependence on the parameters

$$F_Z(z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) = \int_0^1 D_1C_{X,Y}\left(w, F_Y(z - F_X^{-1}(w;\boldsymbol{\alpha}_X);\boldsymbol{\alpha}_Y)\right)dw$$

and

$$C_{X,Z}(u,v;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) \\ = \int_0^u D_1C_{X,Y}\left(w, F_Y(F_Z^{-1}(v;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) - F_X^{-1}(w;\boldsymbol{\alpha}_X))\right)dw. \quad (12)$$

We can compute the density function of $F_Z(z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta)$, since

$$f_Z(z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) = \frac{d}{dz}F_Z(z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) =$$

$$= \int_0^1 c_{X,Y}(w, F_Y(z - F_X^{-1}(w;\boldsymbol{\alpha}_X);\boldsymbol{\alpha}_Y);\theta) f_Y(z - F_X^{-1}(w;\boldsymbol{\alpha}_X);\boldsymbol{\alpha}_Y)dw \quad (13)$$

with To simplify the notation we denote $F_Z(z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta)$ by $F_Z(z)$ and $f_Z(z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta)$ by $f_Z(z)$.

To construct the likelihood function of $C_{X,Z}$ we need its density function $c_{X,Z}$ which is

$$c_{X,Z}(u,v;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) = \frac{\partial}{\partial u \partial v}C_{X,Z}(u,v;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) =$$

$$\frac{\partial}{\partial v}\Big[D_1C_{X,Y}(u, F_Y(F_Z^{-1}(v) - F_X^{-1}(u;\boldsymbol{\alpha}_X);\boldsymbol{\alpha}_Y);\theta)\Big]$$

$$= \frac{c_{X,Y}(u, F_Y(F_Z^{-1}(v) - F_X^{-1}(u;\boldsymbol{\alpha}_X);\boldsymbol{\alpha}_Y);\theta) f_Y(F_Z^{-1}(v) - F_X^{-1}(u;\boldsymbol{\alpha}_X);\boldsymbol{\alpha}_Y)}{f_Z(F_Z^{-1}(v))}.$$

The joint distribution of (X,Z) is

$$f_{X,Z}(x,z;\boldsymbol{\alpha}_X,\boldsymbol{\alpha}_Y,\theta) = c_{X,Z}(F_X(x;\boldsymbol{\alpha}_X), F_Z(z);\theta) f_X(x;\boldsymbol{\alpha}_X) f_Z(z)$$

$$= c_{X,Y}(F_X(x;\boldsymbol{\alpha}_X), F_Y(z-x;\boldsymbol{\alpha}_Y);\theta) f_Y(z-x;\boldsymbol{\alpha}_Y) f_X(x;\boldsymbol{\alpha}_X).$$

If $(x_1, ..., x_n)$ and $(z_1, ..., z_n)$ are two sample from X and Z respectively, the logarithm of the likelihood function is

$$\ell(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) = \sum_{i=1}^{n} \log[f_{X,Z}(x_i, z_i; \boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta)] =$$

$$= \sum_{i=1}^{n} \log[c_{X,Z}(F_X(x_i; \boldsymbol{\alpha}_X), F_Z(z_i); \theta)] + \sum_{i=1}^{n} \log[f_X(x_i; \boldsymbol{\alpha}_X)] + \sum_{i=1}^{n} \log[f_Z(z_i)]$$

$$= \sum_{i=1}^{n} \log[c_{X,Y}(F_X(x_i; \boldsymbol{\alpha}_X), F_Y(z_i - x_i; \boldsymbol{\alpha}_Y); \theta)] +$$

$$+ \sum_{i=1}^{n} \log[f_Y(z_i - x_i; \boldsymbol{\alpha}_Y)] + \sum_{i=1}^{n} \log[f_X(x_i; \boldsymbol{\alpha}_X)]$$

The key observation is that the log-likelihood is the sum of three terms $\ell(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) = \ell_C(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) + \ell_Y(\boldsymbol{\alpha}_Y) + \ell_X(\boldsymbol{\alpha}_X)$ where

$$\ell_C(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) = \sum_{i=1}^{n} \log[c_{X,Y}(F_X(x_i; \boldsymbol{\alpha}_X), F_Y(z_i - x_i; \boldsymbol{\alpha}_Y); \theta)],$$

$$\ell_Y(\boldsymbol{\alpha}_Y) = \sum_{i=1}^{n} \log[f_Y(z_i - x_i; \boldsymbol{\alpha}_Y)],$$

and

$$\ell_X(\boldsymbol{\alpha}_X) = \sum_{i=1}^{n} \log[f_X(x_i; \boldsymbol{\alpha}_X)].$$

The two terms $\ell_X(\boldsymbol{\alpha}_X)$ and $\ell_Y(\boldsymbol{\alpha}_Y)$ are the log-likelihood of the univariate margins and may be separately maximized to get estimates

$$\tilde{\boldsymbol{\alpha}}_X = \arg\max_{\boldsymbol{\alpha}_X} \ell_X(\boldsymbol{\alpha}_X),$$

$$\tilde{\boldsymbol{\alpha}}_Y = \arg\max_{\boldsymbol{\alpha}_Y} \ell_Y(\boldsymbol{\alpha}_Y).$$

The function $\ell_C(\tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \theta)$ can now maximized over θ to get the estimate

$$\tilde{\theta} = \arg\max_{\theta} \ell_C(\tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \theta).$$

The vector $\tilde{\boldsymbol{\zeta}} = (\tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta})$ is the IFM estimate of the vector of parameters $\boldsymbol{\zeta} = (\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta)$. Using the marginal parameter estimators $\tilde{\boldsymbol{\alpha}}_X$ and $\tilde{\boldsymbol{\alpha}}_Y$ we obtain the marginal distribution estimators $\tilde{F}_X(\cdot; \tilde{\boldsymbol{\alpha}}_X)$ and $\tilde{F}_Y(\cdot; \tilde{\boldsymbol{\alpha}}_Y)$.

Under regularity conditions, the IFM estimator $\tilde{\boldsymbol{\zeta}} = (\tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta})$ is asymptotically multivariate Gaussian (Godambe, 1960, 1976); in particular

$$\sqrt{n}(\tilde{\boldsymbol{\zeta}} - \boldsymbol{\zeta}) \longrightarrow \mathcal{N}(0, \boldsymbol{\mathcal{V}}).$$

The asymptotic covariance matrix $n^{-1}\boldsymbol{\mathcal{V}}$ may be estimate by the jackknife method. Let $\tilde{\boldsymbol{\zeta}}^{(i)}$ be the estimator of $\boldsymbol{\zeta}$ with the ith observation $(x_i, z_i - x_i)$ deleted, $i = 1, ..., n$. The jackknife estimator of $n^{-1}\boldsymbol{\mathcal{V}}$ is

$$\sum_{i=1}^{n}(\tilde{\boldsymbol{\zeta}}^{(i)} - \boldsymbol{\zeta})^T(\tilde{\boldsymbol{\zeta}}^{(i)} - \boldsymbol{\zeta}).$$

In this paper we work with three different copulas between X and Y: Frank, Gaussian and Student's t.

- $C_{X,Y}$ is a **Frank Copula**: $C_{X,Y}(u, v; \theta) = -\frac{1}{\alpha}\log\left(1 + \frac{(e^{-\alpha u}-1)(e^{-\alpha v}-1)}{e^{-\alpha}-1}\right)$. We have

$$c_{X,Y}(u, v; \theta) = \frac{\theta(1 - e^{-\theta})e^{-\theta u}e^{-\theta v}}{[(e^{-\theta} - 1) + (e^{-\theta u} - 1)(e^{-\theta v} - 1)]^2},$$

therefore

$$\ell_C(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) = \sum_{i=1}^{n}\Big\{\log(\theta(1-e^{-\theta})) - \theta F_X(x_i; \boldsymbol{\alpha}_X) - \theta F_Y(z_i - x_i; \boldsymbol{\alpha}_Y) +$$

$$+ \log[(e^{-\theta} - 1) + (e^{-\theta F_X(x_i;\boldsymbol{\alpha}_X)} - 1)(e^{-\theta F_Y(z_i-x_i;\boldsymbol{\alpha}_Y)} - 1)]^2\Big\}.$$

- $C_{X,Y}$ is a **Gaussian Copula**:

$$C_{X,Y}(u, v; \theta) = \int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\frac{1}{2\pi\sqrt{1-\rho^2}}e^{\frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)}}dxdy.$$

We have

$$C_{X,Y}(u, v; \theta) = \frac{f_{\theta\Phi^{-1}(u),1-\theta^2}(\Phi^{-1}(v))}{\phi(\Phi^{-1}(v))},$$

where Φ is the Standard Gaussian distribution, ϕ is the Standard Gaussian density and $f_{\theta\Phi^{-1}(u),1-\theta^2}$ is the density function of a Gaussian r.v. with mean $\theta\Phi^{-1}(u)$ and variance $1 - \theta^2$. Since the denominator is independent of the copula parameter θ the log-likelihood is

$$\ell_C(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) = \sum_{i=1}^{n}\log\left[f_{\theta\Phi^{-1}(F_X(x_i;\boldsymbol{\alpha}_X)),1-\theta^2}(\Phi^{-1}(F_Y(z_i - x_i; \boldsymbol{\alpha}_Y)))\right]$$

- $C_{X,Y}$ is a **Student's t Copula**:

$$C_{X,Y}(u, v; \theta) = \int_{-\infty}^{t_\lambda^{-1}(u)}\int_{-\infty}^{t_\lambda^{-1}(v)}\frac{\Gamma(\frac{\lambda+2}{2})}{\lambda\pi\Gamma(\frac{\lambda}{2})\sqrt{1-\rho^2}}\left(1+\frac{s^2 + t^2 - 2\rho st}{\lambda(1-\rho^2)}\right)^{-\frac{\lambda+2}{2}}dsdt$$

where

$$t_\lambda(x) = \int_{-\infty}^{x}\frac{\Gamma(\frac{\lambda+1}{2})}{\sqrt{\pi\lambda}\,\Gamma(\frac{\lambda}{2})}\left(1 + \frac{z^2}{\lambda}\right)^{-\frac{\lambda+1}{2}}dz.$$

We can see that this elliptical copula is characterized by two parameters (λ, ρ), so that in this case $\theta = (\lambda, \rho)$. We have

$$c_{X,Y}(u, v; \theta) = (1-\rho^2)^{-1/2} \frac{\Gamma(\frac{\lambda+2}{2})\Gamma(\frac{\lambda}{2})}{\Gamma^2(\frac{\lambda+1}{2})} \frac{\left(1 + \frac{\psi_u^2+\psi_v^2-2\rho\psi_u\psi_v}{\lambda(1-\rho^2)}\right)^{-(\lambda+2)/2}}{\left[\left(1+\frac{\psi_u^2}{\lambda}\right)\left(1+\frac{\psi_v^2}{\lambda}\right)\right]^{-(\lambda+1)/2}}$$

where $\psi_u = t_\lambda^{-1}(u)$ and $\psi_v = t_\lambda^{-1}(v)$. Therefore he log-likelihood is

$$\ell_C(\boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta) =$$

$$\sum_{i=1}^{n} \Big\{ \log\Gamma(\frac{\lambda+2}{2}) + \log\Gamma(\frac{\lambda}{2}) - 2\log\Gamma(\frac{\lambda+1}{2}) - \frac{1}{2}\log(1-\rho^2) +$$

$$-\frac{\lambda+2}{2}\log\left(1+\frac{\psi^2_{F_X(x_i;\boldsymbol{\alpha}_X)} + \psi^2_{F_Y(z_i-x_i;\boldsymbol{\alpha}_Y)} - 2\rho\psi_{F_X(x_i;\boldsymbol{\alpha}_X)}\psi_{F_Y(z_i-x_i;\boldsymbol{\alpha}_Y)}}{\lambda(1-\rho^2)}\right)+$$

$$+\frac{\lambda+1}{2}\log\left[\left(1+\frac{\psi^2_{F_X(x_i;\boldsymbol{\alpha}_X)}}{\lambda}\right)\left(1+\frac{\psi^2_{F_Y(z_i-x_i;\boldsymbol{\alpha}_Y)}}{\lambda}\right)\right]\Big\}.$$

We obtain estimators of $F_Z(z; \boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta)$, $C_{X,Z}(z; \boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta)$ and $f_Z(z; \boldsymbol{\alpha}_X, \boldsymbol{\alpha}_Y, \theta)$ using the IFM estimator $(\tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta})$ and the estimated marginal distributions $\tilde{F}_X(\cdot; \tilde{\boldsymbol{\alpha}}_X)$ and $\tilde{F}_Y(\cdot; \tilde{\boldsymbol{\alpha}}_Y)$.

We get

$$\tilde{F}_Z(z; \tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta}) = \int_0^1 D_1 C_{X,Y}(w, \tilde{F}_Y(z - \tilde{F}_X^{-1}(w; \tilde{\boldsymbol{\alpha}}_X); \tilde{\boldsymbol{\alpha}}_Y); \tilde{\theta}) dw$$

$$\tilde{C}_{X,Z}(\tilde{F}_X(x; \tilde{\boldsymbol{\alpha}}_X), \tilde{F}_Z(z; \tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta}); \tilde{\theta}) =$$

$$= \int_0^{\tilde{F}_X(x;\tilde{\boldsymbol{\alpha}}_X)} D_1 C_{X,Y}(w, \tilde{F}_Y(z - \tilde{F}_X^{-1}(w; \tilde{\boldsymbol{\alpha}}_X); \tilde{\boldsymbol{\alpha}}_Y); \tilde{\theta}) dw.$$

and

$$\tilde{f}_Z(z; \tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta}) =$$

$$\int_0^1 c_{X,Y}(w, \tilde{F}_Y(z - \tilde{F}_X^{-1}(w; \tilde{\boldsymbol{\alpha}}_X); \tilde{\boldsymbol{\alpha}}_Y); \tilde{\theta}) \tilde{f}_Y(z - \tilde{F}_X^{-1}(w; \tilde{\boldsymbol{\alpha}}_X); \tilde{\boldsymbol{\alpha}}_Y) dw$$

With the change of variable $w = \tilde{F}_X(x; \boldsymbol{\alpha}_X)$, the piece-wise approximations of the above integrals are

$$\tilde{F}_Z(z; \tilde{\boldsymbol{\alpha}}_X, \tilde{\boldsymbol{\alpha}}_Y, \tilde{\theta}) =$$

$$\sum_{i=1}^{n-1} \Big[D_1 C_{X,Y}(\tilde{F}_X(x_i; \tilde{\boldsymbol{\alpha}}_X), \tilde{F}_Y(z - x_i; \tilde{\boldsymbol{\alpha}}_Y); \tilde{\theta})(\tilde{F}_X(x_{i+1}; \tilde{\boldsymbol{\alpha}}_X) - \tilde{F}_X(x_i; \tilde{\boldsymbol{\alpha}}_X)) \Big]$$

$$\tilde{C}_{X,Z}(\tilde{F}_X(x;\tilde{\boldsymbol{\alpha}}_X),\tilde{F}_Z(z;\tilde{\boldsymbol{\alpha}}_X,\tilde{\boldsymbol{\alpha}}_Y,\tilde{\theta});\tilde{\theta}) =$$

$$\sum_{i=1}^{n-1}\Big[D_1C_{X,Y}(\tilde{F}_X(x_i;\tilde{\boldsymbol{\alpha}}_X),\tilde{F}_Y(z-x_i;\tilde{\boldsymbol{\alpha}}_Y);\tilde{\theta})(\tilde{F}_X(x_{i+1};\tilde{\boldsymbol{\alpha}}_X)-\tilde{F}_X(x_i;\tilde{\boldsymbol{\alpha}}_X))\tilde{F}_X(x_i;\boldsymbol{\alpha}_X)\Big]$$

$$\tilde{f}_Z(z;\tilde{\boldsymbol{\alpha}}_X,\tilde{\boldsymbol{\alpha}}_Y,\tilde{\theta}) =$$

$$\sum_{i=1}^{n-1}\Big[c_{X,Y}(\tilde{F}_X(x_i;\tilde{\boldsymbol{\alpha}}_X),\tilde{F}_Y(z-x_i;\tilde{\boldsymbol{\alpha}}_Y);\tilde{\theta})\tilde{f}_Y(z-x_i;\tilde{\boldsymbol{\alpha}}_Y)(\tilde{F}_X(x_{i+1};\tilde{\boldsymbol{\alpha}}_X)-\tilde{F}_X(x_i;\tilde{\boldsymbol{\alpha}}_X))\Big].$$

6.2 Estimation results

In this section we present an empirical application of our model based on data on Italian mutual funds provided by Prometeia s.p.a, an Italian consulting company. We use four different time series representing the main categories of mutual funds recognnized under the Italian regulation. They are in turn:

- "Bilanciato" (BIL), referring to investment equally split between equity and bonds;
- "Azionario Monodivisa" (AzMon), devoted to the equity of the euro area;
- "Azionario Multidivisa" (AzMul), representing investment in multi-currency equity;
- "Corporate IG" (CIG), representing investment in corporate bonds.

For each fund we were provided with valuation data and the corresponding benchmark. We computed daily returns from December 2000 to September 2009: overall, we are endowed with a sample of 2275 observations. Just for the sake of illustration, we assume that the benchmark return be normally distributed. We focus instead on the shape of the distribution of the return from the investment strategy and its dependence on the return on the benchmark. Our purpose is to estimate the density function of $Z = X + Y$ using the three different copulas presented above: Frank, Gaussian and Student's t. As for the marginal behavior of Y, we consider two different cases:

- Student's t: $\boldsymbol{\alpha}_Y = \nu_Y$;
- Non-central t: $\boldsymbol{\alpha}_Y = (\kappa_Y, \delta_Y)$.

We standardized our returns so that they have zero mean and unit standard deviation. The estimation of the copula parameter was carried out following the IFM technique presented in section 6.1. Tables 1-2 give the copula parameter estimates for the three copulas used in order to model the dependence between X and Y. We note that all the dependence parameter are significant. We notice that only the corporate mutual fund CIG is essentially gaussian, since the degrees of freedom of both the marginal distribution and the Student's t copula are quite high. In the other cases, the data show both leptokurtosis (due to low degrees of freedom of the marginal) and tail dependence (due to low degrees of freedom of the copula function). Extending the analysis to allow for non-central t distribution of the returns from the investment strategy does not alter the results in any way, signalling that the stock picking activity is not significant for the funds under consideration.

As for the sign of the dependence, only in the case of mutual funds equally invested in both equity and bonds we find a positive, albeit weak, dependence. In the other cases, we find evidence of negative dependence, consistent with some portfolio insurance behavior like in Henriksson and Merton (1981). This negative relationship is particularly strong for the domestic equity fund.

In table 3 we finally report the standardized quantiles of the distribution of the returns on the funds and those on the asset management strategy. The table reports in more detail the difference in the densities depicted in figures 7-10. Only in the CIG case we have that the quantiles are close to each other, reflecting that the degrees of freedom of both the marginal distribution and the copula function were quite high. As for the balanced equity and bond case, the tail dependence in the copula function turns out in much higher leptokurtosis of the managed fund returns with respect to that of the investment strategy return: this is consistent with both the positive dependence and high tail dependence found in the data. Contrary to that, in the domestic equity fund the high tail dependence effect seems to be mitigated by the strong negative sign of the dependence structure, so that kurtosis of the managed fund turns out to be substantially lower than that of the strategy. The multi-currency case is finally in the middle with a weak increase of leptokurtosis of the managed fund.

7 Conclusion

In this paper we propose a flexible model to specify and estimate the contribution of an investment strategy to the return of a managed fund. More specifically, the stock picking ability of the asset manager is represented by a probability distribution, while her market timing ability is portrayed by the dependence structure between this distribution and the market, represented by a copula function. We show how to estimate: i) the probability distribution of a managed fund, given the stock picking ability (a marginal distribution) and market timing ability (a copula function) and ii) how to extract the stock picking and market timing ability given the returns on the managed fund and those on the market.

For the sake of illustration we provide application of the method to four cases of Italian mutual funds. We use marginal Student t and non-central t distributions to represent returns from stock picking and elliptical copulas to represent market timing. In all cases we find evidence of market timing activity. While in one case we find that the joint distribution is gaussian, in the other cases we also find evidence of tail dependence of the market timing activity. In some cases, this turns out in higher leptokurtosis of the managed fund returns with respect

to returns on the strategy. In one other case, we find that the leptokurtosis of the managed fund is actually decreased: interestingly, this result corresponds to a case in which there is strong evidence of negative dependence between the market and the strategy corresponding to a portfolio insurance activity of the kind portrayed in the classical Henriksson and Merton (1981) framework.
Future research will include extension of the model to cases in which liquidity issues make more involved the dynamic structure of dependence between managed returns and the market and applications to several classes of investment, included multi-factor models.

References

[1] Berk J.B., Green R.C. (2004): Mutual Funds and Performance in Rational Markets, *Journal of Political Economy*, 112(6), 1269-1295

[2] Chen, X., Fan, Y., (2006): Estimation of copula-based semiparametric time series models, *Journal of Econometrics*, 130, 307-335.

[3] Chen, X., Wu S.,B, and Y. Yi (2009): Efficient estimation of copula-based semiparametric Markov models. *Cowles Foundation Discussion Paper n. 1691*

[4] Cherubini, U., Mulinacci S., Romagnoli S. (2009): A copula-based model of Speculative Price Dynamics in Discrete Time. *working paper*

[5] Godambe, V., P. (1960): An optimal property of regular maximum likelihood estimation, *Ann. Math. Statist.*, 31, 1208-1211

[6] Godambe, V., P. (1976): Conditional likelihood and unconditional optimum estimating equations, *Biometrika*, 63, 277-284

[7] Henriksson R.D., Merton R.C. (1981): On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business*, 54(4), 513-533

[8] Jiang W. (2003): A Non Parametric Test of Market Timing, *Journal of Empirical Finance*, 10, 399-425

[9] Joe H. (1997): Multivariate Models and Dependence Concepts, Chapman & Hall, London

[10] Joe H., Xu J.J. (1996): The Estimation Method of Inference Functions for Margins for Multivariate Models, Dept. of Statistics, University of British Coloumbia, Tech. Rep., 166

[11] Leigh W., Paz N., Purvis R. (2002): Market Timing: A Test of a Chartist Heuristics, *Economics Letters*, 77, 55-63

[12] Merton R.C.(1981): On Market Timing and Investment Performance. II. An Equilibrium Theory of Value For Market Forecests, *Journal of Business*, 54(3), 363-406

[13] Nelsen R.(2006): An Introduction to Copulas, Springer

[14] Treynor J., Mazuy F. (1966): Can Mutual Funds Outguess the Market? *Harvard Business Review*, 44, 131-136

8 Appendix: Derivation of the HM-copula

In this appendix we derive the copula that links X and Y in the Henriksson-Merton model presented in Section 4. We study the case where $\gamma > 0$. Suppose that marginal distributions F_X and F_ϵ are assigned. To construct this copula function we need the joint distribution between X and Y, say $F_{X,Y}$. We have

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x, \alpha + \gamma \max\{K, -X\} + \epsilon \leq y) =$$

$$\int_{-\infty}^{+\infty} \mathbb{P}(X \leq x, \alpha + \gamma \max\{K, -X\} + e \leq y | \epsilon = e) F_\epsilon(de) =$$

$$\int_{-\infty}^{+\infty} \mathbb{P}(X \leq x, \max\{K, -X\} \leq \frac{y - e - \alpha}{\gamma}) F_\epsilon(de) =$$

$$\int_{-\infty}^{+\infty} \Big[\mathbb{P}(X \leq x, X \geq \frac{e - y + \alpha}{\gamma}, X < -K) + \mathbb{P}(X \leq x, K \geq \frac{y - e - \alpha}{\gamma}, X \geq -K)\Big] F_\epsilon(de).$$

We have to distinguish two cases: $x > -K$ and $x \leq -K$. In the first case

$$\mathbb{P}(X \leq x, X \geq \frac{e - y + \alpha}{\gamma}, X < -K) + \mathbb{P}(X \leq x, K \geq \frac{y - e - \alpha}{\gamma}, X \geq -K) =$$

$$\mathbb{P}(X \leq -K, X \geq \frac{e - y + \alpha}{\gamma}) + \mathbb{P}(-K \leq X \leq x, e \geq y - \alpha - \gamma K),$$

therefore, if we indicate by $G_{X,Y}$ the joint distribution in this case

$$G_{X,Y}(x,y) =$$

$$\int_{-\infty}^{y-\alpha-\gamma K} \Big[\mathbb{P}(X \leq -K, X \geq \frac{e - y + \alpha}{\gamma}) + \mathbb{P}(-K \leq X \leq x, e \geq y - \alpha - \gamma K)\Big] F_\epsilon(de) =$$

$$F_X(x) F_\epsilon(y - \alpha - \gamma K) - \int_{-\infty}^{y-\alpha-\gamma K} F_X(\frac{e - y + \alpha}{\gamma}) F_\epsilon(de).$$

in the case where $x \leq -K$ we have

$$\mathbb{P}(X \leq x, X \geq \frac{e - y + \alpha}{\gamma}, X < -K) + \mathbb{P}(X \leq x, K \geq \frac{y - e - \alpha}{\gamma}, X \geq -K) =$$

$$\mathbb{P}(X \leq x, X \geq \frac{e - y + \alpha}{\gamma}),$$

then, the only relevant case is when $\frac{e-y+\alpha}{\gamma} \leq x$ and therefore if we indicate by $H_{X,Y}$ the joint distribution in this case

$$H_{X,Y}(x,y) = \int_{-\infty}^{\gamma x + y - \alpha} \mathbb{P}(X \leq x, X \geq \frac{e - y + \alpha}{\gamma}) F_\epsilon(de) =$$

$$F_X(x) F_\epsilon(\gamma x + y - \alpha) - \int_{-\infty}^{\gamma x + y - \alpha} F_X(\frac{e - y + \alpha}{\gamma}) F_\epsilon(de).$$

Briefly, we can conclude that

$$F_{X,Y}(x,y) = G_{X,Y}(x,y)\mathbf{1}_{\{x>-K\}} + H_{X,Y}(x,y)\mathbf{1}_{\{x\leq -K\}}$$

Now, the copula between X and Y is given by

$$C_{X,Y}(u,v) = F_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v)) =$$

$$G_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v))\mathbf{1}_{\{u>F_X(-K)\}} + H_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v))\mathbf{1}_{\{u\leq F_X(-K)\}},$$

or more explicitly

$$C_{X,Y}(u,v) =$$

$$\Big[uF_\epsilon(F_Y^{-1}(v)-\alpha-\gamma K) - \int_{-\infty}^{F_Y^{-1}(v)-\alpha-\gamma K} F_X(\frac{e-F_Y^{-1}(v)+\alpha}{\gamma})F_\epsilon(de)\Big]\mathbf{1}_{\{u>F_X(-K)\}} +$$

$$+\Big[uF_\epsilon(\gamma F_X^{-1}(u)+F_Y^{-1}(v)-\alpha) - \int_{-\infty}^{\gamma F_X^{-1}(u)+F_Y^{-1}(v)-\alpha} F_X(\frac{e-F_Y^{-1}(v)+\alpha}{\gamma})F_\epsilon(de)\Big]\mathbf{1}_{\{u\leq F_X(-K)\}}.$$

We call this copula function the Henriksson-Merton copula. It remains to determine the marginal distribution F_Y. We have

$$F_Y(y) = \lim_{x\to+\infty} F_{X,Y}(x,y) = \lim_{x\to+\infty} G_{X,Y}(x,y) =$$

$$F_\epsilon(y-\alpha-\gamma K) - \int_{-\infty}^{y-\alpha-\gamma K} F_X(\frac{e-y+\alpha}{\gamma})F_\epsilon(de).$$

Moreover, we determine its density function by derivation

$$f_Y(y) = f_\epsilon(y-\alpha-\gamma K)(1-F_X(-K)) + \int_{-\infty}^{y-\alpha-\gamma K} f_X(\frac{e-y+\alpha}{\gamma})F_\epsilon(de).$$

Finally, we can find the distribution of $Z = \beta X + Y$ by using a generalization of (4). Since we assume that $\beta > 0$

$$F_Z(z) = \int_0^1 D_1 C_{X,Y}\left(w, F_Y(z-\beta F_X^{-1}(w)\right),$$

where

$$D_1 C_{X,Y}(u,v) =$$

$$F_\epsilon(F_Y^{-1}(v)-\alpha-\gamma K)\mathbf{1}_{\{u>F_X(-K)\}} + F_\epsilon(\gamma F_X^{-1}(u)+F_Y^{-1}(v)-\alpha)\mathbf{1}_{\{u\leq F_X(-K)\}}.$$

So, more explicitly

$$F_Z(z) =$$

$$\int_0^{F_X(-K)} F_\epsilon(z+(\gamma-\beta)F_X^{-1}(w)-\alpha)dw + \int_{F_X(-K)}^1 F_\epsilon(z-\beta F_X^{-1}(w)-\alpha-\gamma K)dw.$$

As a consequence we can compute the density of Z, which is

$$f_Z(z) =$$

$$\int_0^{F_X(-K)} f_\epsilon(z+(\gamma-\beta)F_X^{-1}(w)-\alpha)dw + \int_{F_X(-K)}^1 f_\epsilon(z-\beta F_X^{-1}(w)-\alpha-\gamma K)dw$$

where f_X and f_ϵ are the probability density functions of X and ϵ respectively.

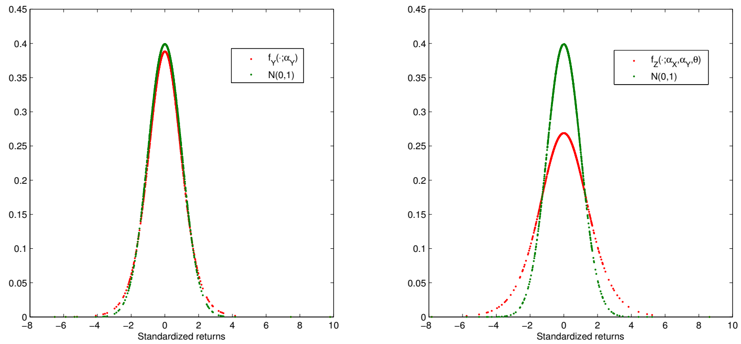

Figure 7: Comparison of the estimated distribution of Y, $\tilde{f}_Y(y;\tilde{\boldsymbol{\alpha}}_Y)$, the estimated distribution of the managed fund with Student's t copula, $\tilde{f}_Z(z;\tilde{\boldsymbol{\alpha}}_X,\tilde{\boldsymbol{\alpha}}_Y,\tilde{\theta})$, and the Standard Gaussian distribution. Data source: BIL.

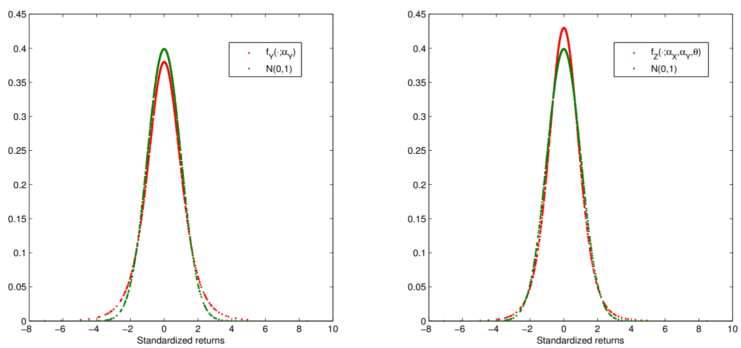

Figure 8: Comparison of the estimated distribution of Y, $\tilde{f}_Y(y;\tilde{\boldsymbol{\alpha}}_Y)$, the estimated distribution of the managed fund with Student's t copula, $\tilde{f}_Z(z;\tilde{\boldsymbol{\alpha}}_X,\tilde{\boldsymbol{\alpha}}_Y,\tilde{\theta})$, and the Standard Gaussian distribution. Data source: Az-Mon.

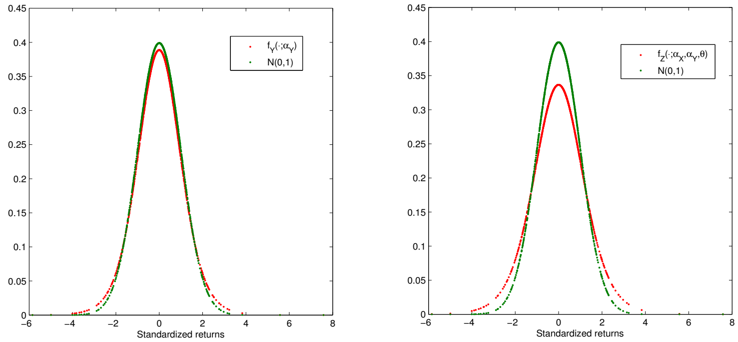

Figure 9: Comparison of the estimated distribution of Y, $\tilde{f}_Y(y;\tilde{\boldsymbol{\alpha}}_Y)$, the estimated distribution of the managed fund with Student's t copula, $\tilde{f}_Z(z;\tilde{\boldsymbol{\alpha}}_X,\tilde{\boldsymbol{\alpha}}_Y,\tilde{\theta})$, and the Standard Gaussian distribution. Data source: Az-Mul.

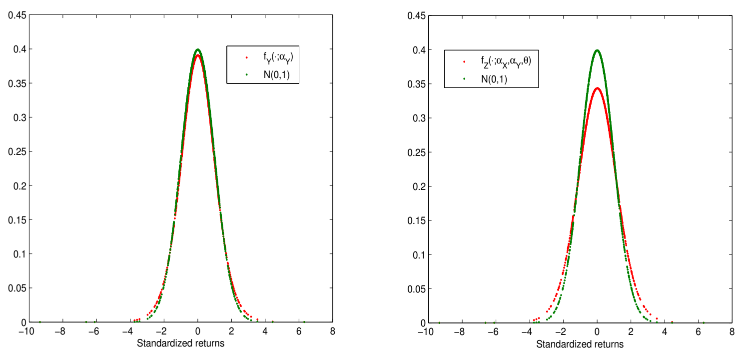

Figure 10: Comparison of the estimated distribution of Y, $\tilde{f}_Y(y;\tilde{\boldsymbol{\alpha}}_Y)$, the estimated distribution of the managed fund with Student's t copula, $\tilde{f}_Z(z;\tilde{\boldsymbol{\alpha}}_X,\tilde{\boldsymbol{\alpha}}_Y,\tilde{\theta})$, and the Standard Gaussian distribution. Data source: CIG.

		Frank	Gaussian	Student's t
	$\tilde{\nu}_Y$	$\tilde{\theta}$	$\tilde{\theta}$	$\tilde{\boldsymbol{\theta}}$
BIL	9.1701	1.1688	0.1887	(0.1558, 5.8211)
	(1.6385)	(0.2859)	(0.0346)	((0.0436), (0.9859))
AzMon	5.0615	-5.6097	-0.4013	(-0.5888, 7.5786)
	(0.4533)	(0.2227)	(0.0228)	((0.0213), (0.6416))
AzMul	9.6126	-1.8349	-0.2276	(-0.2490, 6.4321)
	(1.6241)	(0.2412)	(0.0305)	((0.0332), (1.0418))
CIG	11.8637	-2.3563	-0.2898	(-0.3312, 17.0162)
	(2.8239)	(0.1655)	(0.0212)	((0.0212), (7.2404))

Table 1: IFM estimates of the marginal distribution and copula parameters in the case $Y \sim t_{\nu_Y}$.

			Frank	Gaussian	Student's t
	$\tilde{\kappa}_Y$	$\tilde{\delta}_Y$	$\tilde{\theta}$	$\tilde{\theta}$	$\tilde{\boldsymbol{\theta}}$
BIL	9.1698	0.0013	1.1684	0.1887	(0.1557, 5.8191)
	(1.6391)	(0.0181)	(0.2851)	(0.0361)	((0.0436), (0.9844))
AzMon	5.0610	0.0052	-5.6136	-0.4013	(-0.5890, 7.5638)
	(0.4534)	(0.0146)	(0.2220)	(0.0228)	((0.0212), (0.6388))
AzMul	9.6120	-0.0021	-1.8143	-0.2276	(-0.2489, 6.4302)
	(1.6241)	(0.0181)	(0.2415)	(0.0305)	((0.0296), (0.9891))
CIG	11.8578	0.0070	-2.3572	-0.2898	(-0.3310, 17.0944)
	(2.8013)	(0.0187)	(0.1655)	(0.0212)	((0.0198), (7.0555))

Table 2: IFM estimates of the marginal distribution and copula parameters in the case $Y \sim NCt(\kappa_Y, \delta_Y)$.

	BIL		AzMon		AzMul		CIG	
%	$\tilde{F}_Y$	$\tilde{F}_Z$	$\tilde{F}_Y$	$\tilde{F}_Z$	$\tilde{F}_Y$	$\tilde{F}_Z$	$\tilde{F}_Y$	$\tilde{F}_Z$
1	-2.8088	-4.0441	-3.3383	-2.7253	-2.7874	-3.2654	-2.6768	-2.9210
2.5	-2.2541	-3.2520	-2.5530	-2.1372	-2.2431	-2.6189	-2.1733	-2.3994
5	-1.8287	-2.6390	-2.0024	-1.6938	-1.8224	-2.1215	-1.7760	-1.9774
50	0.0013	0.0019	0.0055	0.0059	-0.0022	-0.0022	0.0071	0.0107
95	1.8308	2.6421	2.0170	1.7076	1.8174	2.1154	1.7921	1.9976
97.5	2.2575	3.2549	2.5696	2.1522	2.2377	2.6159	2.2015	2.4203
99	2.8124	4.0484	3.3578	2.7424	2.7816	3.2665	2.6949	2.9440

Table 3: Percentiles of the estimated distributions $\tilde{F}_Y$ and $\tilde{F}_Z$ (standardized values).