

IS A NORMAL COPULA THE RIGHT COPULA?

Dante Amengual and Enrique Sentana

CEMFI Working Paper No. 1504

August 2015

CEMFI
Casado del Alisal 5; 28014 Madrid
Tel. (34) 914 290 551 Fax (34) 914 291 056
Internet: www.cemfi.es

We would like to thank Manuel Arellano, Yanqin Fan, Pedro García Ares, Pascal Lavergne, Javier Mencía, Andrew Patton and Javier Perote for their comments, advice and suggestions. We have also benefited from the feedback provided by seminar participants at CEMFI, Duke, Erasmus, Exeter, Murcia and Yale, as well as audiences at the TSE Financial Econometrics Conference (May 2013), the XXI Finance Forum (Segovia, November 2013), the XXXVIII Symposium on Economic Analysis (Santander, December 2013), the LATAM Econometrics Workshop (São Paulo, December 2013), the IEEM Meeting (Montevideo, December 2013), the 68th ESEM (Toulouse, August 2014) and the 11th World Congress of the Econometric Society (Montreal, August 2015). We thank Difang Huang and Zaici Li for their help in collecting the data and especially Julio Crego for his excellent research assistance. Of course, the usual caveat applies. Financial support from the Spanish Ministries of Science & Innovation and Economy & Competitiveness through grants ECO 2011-26342 and 2014-59262, respectively, is gratefully acknowledged.

IS A NORMAL COPULA THE RIGHT COPULA?

Abstract

We derive computationally simple and intuitive expressions for score tests of Gaussian copulas against Generalised Hyperbolic alternatives, including symmetric and asymmetric Student t , and Hermite polynomial expansions. We decompose our tests into third and fourth moment components, and obtain one-sided Likelihood Ratio analogues, whose asymptotic distribution we provide. We conduct Monte Carlo exercises to assess the finite sample properties of asymptotic and bootstrap versions of our tests. In an empirical application to CRSP stocks, we find that short-term reversals and momentum effects are better captured by non-Gaussian copulas. We estimate their parameters by indirect inference, and devise successful trading strategies.

JEL Codes: C12, C46, C52, G11, G14.

Keywords: Cokurtosis, coskewness, indirect inference, Kuhn-Tucker test, momentum strategies, non-linear dependence, short-term reversals, supremum test, underidentified parameters.

Dante Amengual
CEMFI
amengual@cemfi.es

Enrique Sentana
CEMFI
sentana@cemfi.es

1 Introduction

Although the efficient markets hypothesis implies that current stock prices reflect all the information included in previous prices and financial statements, there is a substantial body of empirical evidence which suggests that individual stock returns are not cross-sectionally independent of some of their observable characteristics. In particular, a broad consensus has emerged on the presence of momentum, reversals and value effects in individual stock returns all over the world (see e.g. Asness, Moskowitz and Pedersen (2013) and the references therein). As is well known, short term reversal strategies regularly invest in those stocks that have underperformed in the past month, while momentum strategies typically invest in stocks that outperformed over the previous months of the past year. In turn, value strategies focus on stocks whose price-to-book or price-to-earnings ratios are low. However, most of the literature has focused on assessing the excess returns obtained by various trading strategies rather than on precisely analyzing the cross-sectional dependence between stock returns this month and those characteristics at the time investment decisions are made. Such an analysis is important not only because it can potentially lead to better decisions, but also because it can shed some light on the sources of the dependence.

There are several different ways of characterizing dependence. The most straightforward one is by means of linear relationships, as it is often done in the extensive growth convergence literature in macroeconomics. Specifically, a researcher could cross-sectionally regress individual stock returns this month on a constant and returns over previous months, and look at the size and significance of Pearson correlation coefficient. Figures 1a (reversals) and 1b (momentum) contain the results of applying such a procedure to stocks in the CRSP database in March 1998, an uneventful month. Similarly, Figures 2a and 2b present the corresponding evidence for August 2007, a period at the beginning of the global financial crises with wild gyrations in the US stock market due to the actions of quantitative equity market-neutral trading strategies (see Lo (2007)). The problem with this linear approach is that the OLS coefficient estimates may be extremely sensitive to a few outliers, as illustrated in Figure 2b, in which the slightly negative slope is largely driven by the southeasternmost stocks. In fact, if we trim the sample of 2,463 observations by simply excluding those five stocks whose cumulative return over the period September 2006 - June 2007 exceeded 300%, we obtain a positive correlation.

A far more robust procedure would involve rank regressions, whereby one regresses the cross-

sectional rank of stocks this month on a constant and their cross-sectional rank over the relevant period in the past. Figures 1c and 2c (reversals) and 1d and 2d (momentum) contain the scatterplots of the corresponding normalized ranks for the same two months, the associated regression lines and Spearman correlation coefficients. This procedure is closely related to the concept of “copula”, which allows us to separate joint distributions from marginal ones by fixing the latter. In the case of rank regressions, in particular, the empirical marginal distributions are discrete uniform by construction. But this is not the only possibility. An equally robust approach is to look at the dependence between the so-called Gaussian ranks, which are simple monotonic transformations of the usual ranks by means of the standard normal quantile function. In fact, one may convincingly argue that scatterplots of Gaussian ranks are easier to interpret than scatterplots of uniform ranks, if only because empirical researchers are more used to analyzing real data with approximately Gaussian marginals than uniform ones (see Joe (2015) for a more formal justification). Figures 1e and 2e (reversals) and 1f and 2f (momentum) show the scatterplots of these alternative ranks, the level curves of bivariate kernel density estimates, the corresponding regression lines and what we shall henceforth call the Gaussian rank correlation coefficients.

Both the Gaussian rank correlation and Spearman correlation coefficients confirm the presence of momentum and short term reversals in individual stock returns. However, feasible trading strategies that exploit such dependence require more specificity. In particular, we can devise sound trading strategies along the following lines:

1. We look at the (Gaussian) rank of the chosen characteristic of an individual stock over the relevant observation period.
2. Conditional on that rank, a copula allows us to make probabilistic predictions about the rank of the return on that stock over the next month.
3. If the predicted probability of the rank being high is large, we buy the stock.
4. If the predicted probability of the rank being low is large, we sell it short.
5. Otherwise, we do not hold any position on it.

The (Gaussian) rank correlation is obviously very important in determining those probabilities, but it will not necessarily be a sufficient statistic in the non-Gaussian copula case. In

general, non-linear (i.e. tail) dependence matters, especially when the rank correlation coefficient is low, as in Figures 1 and 2.

Nowadays copulas are extensively used in many other economic and finance applications, with the Gaussian copula being very popular despite ruling out non-linear dependence, particularly in the lower tail. Nevertheless, the validity of this copula in finance has been the subject of considerable public debate, to the extent that the media declared it “the formula that felled Wall Street” (see the provocative article by Salmon (2009), the more nuanced analysis by MacKenzie and Spears (2012), and the academic response by Donnelly and Embrechts (2010)). To be fair, the statistics and econometric literatures were well aware of the possibility of misspecification of the assumed copula, and several more or less formal diagnostics have already been proposed (see e.g. Malevergne and Sornette (2003), Berg and Quessy (2009) and Genest, Rémillard and Beaudoin (2009)).

The first objective of our paper is to provide computationally simple and intuitive expressions for score-based specification tests of Gaussian copulas against both Generalized Hyperbolic alternatives, which include the symmetric and asymmetric Student t , normal-gamma mixtures, hyperbolic, normal inverse Gaussian and symmetric and asymmetric Laplace distributions, and Hermite polynomial expansions.

We decompose our tests into third and fourth moment analogues, and obtain more powerful one-sided Kuhn-Tucker versions that are equivalent to the Likelihood Ratio test. Importantly, we show that the asymptotic distribution of our proposed tests is standard despite the non-standard features of the problem in the case of the asymmetric Student t (underidentified parameters under the null), Generalized Hyperbolic (three different ways of converging to the null), and Hermite expansions (identically zero scores). In all cases, we derive closed-form expressions for the asymptotic covariance matrices of the influence functions we use for testing, which should improve the finite sample reliability of our tests. As an aside, we also show that the empirical Gaussian rank correlation coefficient is an efficient estimator of its population counterpart when the true copula is Gaussian, and remains consistent even when it is not, thereby inheriting the usual properties of the Gaussian pseudo maximum likelihood estimators of Pearson correlation coefficient.

For pedagogical reasons, we initially assume known margins, but since this rarely happens in practice, we also consider the usual two-step estimation procedure, whereby the marginal

distributions are replaced by their (re-scaled) empirical cdf counterparts. For the distributions that we consider as alternatives, we show that it is possible to capture the variance modification in the scores of the shape parameters by adding linear combinations of third and fourth Hermite polynomials in the Gaussian ranks. In contrast, the adjustment for the correlation coefficients only involves linear combinations of second order polynomials in each of those variables.

We also study the finite sample properties of the asymptotic and parametric bootstrap versions of our proposed tests with an extensive Monte Carlo analysis. We find that the asymptotic p-values are very reasonable in moderately large samples sizes, but not in small ones. In contrast, the parametric bootstrap rejection rates are close to being perfect for all samples sizes. In addition, the finite sample power of the different test statistics agrees with what the asymptotic results would suggest.

Finally, we employ our proposed tests to assess the suitability of the Gaussian copula for capturing the short-term reversals and momentum effects observed in individual stock returns in the CRSP data base. In both cases, we reject the null hypothesis of a Gaussian copula by a long margin, the source of the rejection being not only the “cokurtosis” between the Gaussian ranks, but also their “coskewness”, especially for momentum strategies. For that reason, we estimate the parameters of the non-Gaussian copulas by means of a constrained indirect inference approach that uses the Gaussian rank correlations and our score tests as sample statistics to match, as suggested by Calzolari, Fiorentini and Sentana (2004). Importantly, we find that it is possible to devise successful trading strategies that exploit those features of the data.

The rest of the paper is divided as follows. In section 2, we discuss the relevant theoretical background to the problem, and develop our proposed tests. Next, we report the results from an extensive Monte Carlo exercise in section 3. We then analyze the cross-sectional dependence between monthly returns on individual U.S. stocks in the CRSP database and some of their observable characteristics in section 4, followed by our conclusions. Proofs and auxiliary results are relegated to appendices.

2 Theoretical background

2.1 Econometric model

Let \mathbf{x} denote a vector of K continuous random variables. The traditional way of modelling the dependence between the elements of \mathbf{x} is through the joint distribution function $F_K(\mathbf{x})$ or the

associated density function $f_K(\mathbf{x})$ when it is well defined. These functions are often recursively factorized for a predetermined ordering as the sequence of conditional distributions of x_k given $x_{k-1}, x_{k-2}, \dots, x_1$ ($k = 2, \dots, K$) times the marginal distribution of x_1 .

In contrast, the standard copula approach first instantaneously transforms each of the elements of \mathbf{x} into a uniform random variable by means of the probability integral transform $u_k = F_{1k}(x_k)$, where $F_{1k}(\cdot)$ is the marginal cumulative distribution function of x_k , and then models the dependence of the random vector $\mathbf{u} = (u_1, \dots, u_K)'$ through a joint distribution function $F_{\mathbf{U}}(\mathbf{u})$ with uniform marginals defined over the unit hypercube in \mathbb{R}^K . This distribution function is known as the copula distribution function, and the associated density as the copula density function.

Although there are many well known examples of bivariate copulas, some of them are popular simply because they are mathematically convenient, as opposed to being motivated by empirical observations on real life phenomena. More importantly, they are difficult to generalize to multiple dimensions. On the other hand, the Gaussian copula is a popular choice both in bivariate and multivariate contexts since it is easily scalable. Moreover, as its name suggests, it is the copula function that corresponds to the multivariate Gaussian distribution, which remains dominant in multivariate statistical analysis.

More formally, define $\mathbf{y} = (y_1, \dots, y_K)'$, where $y_k = \Phi_1^{-1}(u_k)$, $\Phi_1(\cdot)$ denotes the univariate standard normal cumulative distribution function and $\Phi_1^{-1}(\cdot)$ the corresponding quantile function. The Gaussian copula with correlation matrix $\mathbf{P}(\boldsymbol{\rho})$, where $\mathbf{P}(\boldsymbol{\rho})$ is a positive definite matrix which contains $K(K-1)/2$ possibly distinct functions of the $p \times 1$ vector of parameters $\boldsymbol{\rho}$, is derived from the cumulative distribution function of a multivariate random vector $\mathbf{y} \sim N[\mathbf{0}, \mathbf{P}(\boldsymbol{\rho})]$. Specifically, its density function will be given by

$$f_{\mathbf{U}}(\mathbf{u}; \boldsymbol{\rho}) = |\mathbf{P}(\boldsymbol{\rho})|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{y}' [\mathbf{P}^{-1}(\boldsymbol{\rho}) - \mathbf{I}_K] \mathbf{y} \right\} = |\mathbf{P}(\boldsymbol{\rho})|^{-1/2} \exp \left\{ -\frac{1}{2} [\varsigma(\boldsymbol{\rho}) - \varsigma(\mathbf{0})] \right\},$$

where $\varsigma(\boldsymbol{\rho}) = \mathbf{y}' \mathbf{P}^{-1}(\boldsymbol{\rho}) \mathbf{y}$ and $\varsigma(\mathbf{0}) = \mathbf{y}' \mathbf{y}$, as we assume throughout that we have parametrized $\mathbf{P}(\boldsymbol{\rho})$ in such a way that $\mathbf{P}(\mathbf{0}) = \mathbf{I}_K$. Figure 3 displays a bivariate Gaussian copula density with Spearman correlation of .115 ($\rho = .12$), both with uniform (Figures 3a-b) and Gaussian margins (Figures 3c-d). Finally, Figure 3e presents the conditional densities generated by the Gaussian copula density with Gaussian margins. As expected, only the mean depends on the value of the conditioning variable.

In principle, we could consider more complex models by conditioning on past values of \mathbf{x} or

present and past values of some exogenous variables \mathbf{z} , but for the sake of clarity we will only explicitly cover multivariate distributions without conditioning variables.¹

2.2 Some parametric models that nest the Gaussian copula

As we mentioned in the introduction, the Gaussian copula rules out any type of non-linear dependence between the elements of \mathbf{y} . For that reason, empirical researchers have considered more flexible copulas that nest the normal copula as a special case, which are generated from a multivariate distribution that in turn nests the multivariate normal distribution. Some important examples are the symmetric Student t , the far more flexible Generalized Hyperbolic distribution, as well as Hermite polynomial expansions.

2.2.1 Student t copulas

The Student t distribution generalizes the multivariate normal distribution through a single additional parameter, which is usually known as the degrees of freedom. The log density of a (standardized) multivariate Student t distribution with mean $\mathbf{0}$, correlation matrix $\mathbf{P}(\boldsymbol{\rho})$ and $\nu > 2$ degrees of freedom is given by the following expression:

$$\begin{aligned} l(\boldsymbol{\phi}; K) &= h(\eta; K) + d(\boldsymbol{\rho}) + g[\varsigma(\boldsymbol{\rho}), \eta; K], \\ h(\eta; K) &= \ln \left[\Gamma \left(\frac{K\eta + 1}{2\eta} \right) \right] - \ln \left[\Gamma \left(\frac{1}{2\eta} \right) \right] - \frac{K}{2} \ln \left(\frac{1 - 2\eta}{\eta} \right) - \frac{K}{2} \ln \pi, \\ d(\boldsymbol{\rho}) &= -\frac{1}{2} \ln |\mathbf{P}(\boldsymbol{\rho})|, \\ g[\varsigma(\boldsymbol{\rho}), \eta; K] &= - \left(\frac{K\eta + 1}{2\eta} \right) \ln \left[1 + \frac{\eta}{1 - 2\eta} \varsigma(\boldsymbol{\rho}) \right] \end{aligned} \tag{1}$$

where $\Gamma(\cdot)$ is Euler's gamma (or generalized factorial) function and $\eta = 1/\nu$ (see Fiorentini, Sentana and Calzolari (2003)).² This distribution is such that its marginal components are also univariate Student t 's with mean 0, unit variance and ν degrees of freedom.

Therefore, the Student t copula will be given by the expression

$$f_{\mathbf{U}}(\mathbf{u}; \boldsymbol{\rho}, \eta) = |\mathbf{P}(\boldsymbol{\rho})|^{-1/2} \exp[h(\eta; K) - Kh(\eta; 1)] \frac{[1 + \eta\varsigma(\boldsymbol{\rho})/(1 - 2\eta)]^{(K\eta+1)/(2\eta)}}{\left\{ \prod_{k=1}^K [1 + \eta y_k^2/(1 - 2\eta)] \right\}^{(\eta+1)/(2\eta)}},$$

where $\mathbf{y} = [F_1^{-1}(u_1; \eta), \dots, F_1^{-1}(u_K; \eta)]'$, with $F_1^{-1}(u_k; \eta)$ denoting the quantile function of a univariate standardized Student t .

¹One common complication in dynamic copula models is that in theory the "marginal" distributions of the components used to transform observations into ranks should be conditional on the past of all the random variables involved.

²This distribution differs from the textbook multivariate Student t distribution in that the kernel is $\ln[1 + \eta(1 - 2\eta)^{-1}\varsigma]$ instead of $\ln[1 + \eta\varsigma]$ in order to guarantee that $\mathbf{P}(\boldsymbol{\rho})$ coincides with the correlation matrix. This difference is inconsequential in the neighbourhood of the null hypothesis $H_0 : \eta = 0$.

As expected, this expression converges to the normal copula as $\eta \rightarrow 0^+$ but otherwise it induces tail dependence even when $\boldsymbol{\rho} = \mathbf{0}$. Figure 4 displays a bivariate Student t copula density with the same Spearman correlation as in Figure 3 ($\rho = .122$), both with uniform (Figures 4a-b) and Gaussian margins (Figures 4c-d). Figure 4e presents the conditional densities that arise from the Student t copula density with Gaussian margins. In this case, not only the conditional mean changes but also the higher order moments. Nevertheless, those conditional distributions will be symmetric when $\rho = 0$.

2.2.2 Generalized hyperbolic copulas

The Generalized Hyperbolic (GH) distribution introduced by Barndorff-Nielsen (1977) and studied in detail by Blæsild (1981) is a rather flexible family of multivariate distributions that nests not only the normal and Student t but also many other examples such as the asymmetric Student t , the hyperbolic and normal inverse Gaussian distributions, as well as symmetric and asymmetric versions of the normal-gamma mixture and Laplace.

We can gain some intuition about the GH distribution by considering its interpretation as a location-scale mixture of normals in which the mixing variable is a Generalized Inverse Gaussian (GIG). If $\boldsymbol{\varepsilon}$ is a GH vector, then it can be expressed as

$$\boldsymbol{\varepsilon} = \boldsymbol{\alpha} + \boldsymbol{\Upsilon}\boldsymbol{\beta}\xi^{-1} + \xi^{-\frac{1}{2}}\boldsymbol{\Upsilon}^{\frac{1}{2}}\boldsymbol{\varepsilon}^{\circ}, \quad (2)$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^K$, $\boldsymbol{\Upsilon}$ is a symmetric positive definite matrix of order K , $\boldsymbol{\varepsilon}^{\circ} \sim iid N(\mathbf{0}, \mathbf{I}_K)$ and the positive mixing variable ξ is an independent iid GIG with parameters $-\nu$, γ and δ , or $\xi \sim GIG(-\nu, \gamma, \delta)$ for short, where $\nu \in \mathbb{R}$ and $\gamma, \delta \in \mathbb{R}^+$ (see Jørgensen (1982) and Johnson, Kotz, & Balakrishnan (1994) for details). Since $\boldsymbol{\varepsilon}$ given ξ is Gaussian with conditional mean $\boldsymbol{\alpha} + \boldsymbol{\Upsilon}\boldsymbol{\beta}\xi^{-1}$ and covariance matrix $\boldsymbol{\Upsilon}\xi^{-1}$, it is clear that $\boldsymbol{\alpha}$ and $\boldsymbol{\Upsilon}$ play the roles of location vector and dispersion matrix, respectively. There is a further scale parameter, δ , two other scalars, ν and γ , to allow for flexible tail modelling, and the vector $\boldsymbol{\beta}$, which introduces skewness in this distribution. In this sense, the distribution of $\boldsymbol{\varepsilon}$ becomes a simple scale mixture of normals, and thereby spherical, when $\boldsymbol{\beta}$ is zero. Mencía and Sentana (2012) set $\delta = 1$ and derive restrictions on $\boldsymbol{\alpha}$ and $\boldsymbol{\Upsilon}$ which ensure that the elements of $\boldsymbol{\varepsilon}$ are uncorrelated with zero means and unit variances. They also parametrize $\boldsymbol{\beta}$ as a function of a new vector of parameters \mathbf{b} in the following way:

$$\boldsymbol{\beta}(\boldsymbol{\rho}, \mathbf{b}) = \mathbf{P}^{\frac{1}{2}'}(\boldsymbol{\rho})\mathbf{b}, \quad (3)$$

so that the resulting distribution does not depend on the choice of square root matrix.³ Finally, it is analytically convenient to replace ν and γ by η and ψ , where $\eta = -.5\nu^{-1}$ and $\psi = (1+\gamma)^{-1}$, although we continue to use ν and γ in some equations for notational simplicity.

A rather useful property of the *GH* distributions is that the marginal distributions of linear combinations (including the individual components) also follow univariate *GH* distributions (see Blæsild (1981)). As a result, we can easily construct the *GH* copula by once again combining the joint distribution and its marginals.

Figure 5 displays a bivariate asymmetric Student *t* copula density with negative tail dependence but the same Spearman correlation as in Figure 3 ($\rho = .186$), both with uniform (Figures 5a-b) and Gaussian margins (Figures 5c-d). In turn, Figure 5e presents the conditional densities that arise from the asymmetric Student *t* copula density with Gaussian margins. In this case, all higher moments depend on the values of the conditioning variable even when $\rho = 0$.

2.2.3 Hermite polynomial expansions of the Gaussian copula

Hermite polynomial expansions of the multivariate normal density can be understood as Edgeworth-like expansions of its characteristic function. To keep the algebra manageable, we only consider explicitly fourth order expansions in the bivariate case. We say that $(\varepsilon_1, \varepsilon_2)$ follow a fourth order Hermite expansion of the Gaussian distribution when their joint density function is given by

$$f_2(\varepsilon_1, \varepsilon_2; \rho, \boldsymbol{\varphi}) = \phi_2 \left(\frac{\varepsilon_1^2 + \varepsilon_2^2 - 2\rho\varepsilon_1\varepsilon_2}{1 - \rho^2} \right) P_2(\varepsilon_1, \varepsilon_2; \boldsymbol{\varphi})$$

where $\boldsymbol{\varphi} = (c_1, c_2, c_3, c_4, d_1, d_2, d_3, d_4, d_5)'$ and

$$\begin{aligned} P_2(\varepsilon_1, \varepsilon_2; \boldsymbol{\varphi}) = & 1 + c_1 H_3(\varepsilon_1) + c_2 H_2(\varepsilon_1) H_1(\varepsilon_2) + c_3 H_1(\varepsilon_1) H_2(\varepsilon_2) + c_4 H_3(\varepsilon_2) + d_1 H_4(\varepsilon_1) \\ & + d_2 H_3(\varepsilon_1) H_1(\varepsilon_2) + d_3 H_2(\varepsilon_1) H_2(\varepsilon_2) + d_4 H_1(\varepsilon_1) H_3(\varepsilon_2) + d_5 H_4(\varepsilon_2), \end{aligned} \quad (4)$$

with $H_j(\varepsilon_k)$ denoting the j^{th} standardized Hermite polynomial of the k^{th} variable.

This distribution obviously nests the bivariate normal distribution with zero means, unit variances and correlation coefficient ρ when $\boldsymbol{\varphi} = \mathbf{0}$, but it can depart significantly from Gaussianity otherwise. Nevertheless, by relying on third and fourth order (cross-products of) Hermite polynomials we ensure that the two marginals have zero means and unit variances, and at the same time, the correlation ρ will remain the same under the alternative.

³ $\mathbf{P}^{1/2}(\boldsymbol{\rho})$ denotes some particular “square root” matrix such that $\mathbf{P}^{1/2}(\boldsymbol{\rho})\mathbf{P}^{1/2'}(\boldsymbol{\rho}) = \mathbf{P}(\boldsymbol{\rho})$.

A complication of this distribution is that without further restrictions on $\boldsymbol{\varphi}$, $P_2(\varepsilon_1, \varepsilon_2; \boldsymbol{\varphi})$ cannot be guaranteed to be positive with probability 1. Unfortunately, the required restrictions are non-trivial to find. But since we are interested in testing, as opposed to estimation, we will ignore those positivity restrictions in our analysis, even though this might result in some loss of power.

Figure 6 displays a bivariate Hermite polynomial expansion of the Gaussian copula density with the same Spearman correlation as in Figure 3 ($\rho = .14$), both with uniform (Figures 6a-b) and Gaussian margins (Figures 6c-d). Finally, Figure 6e presents the conditional densities that arise from this copula with Gaussian margins. A noticeable feature of this copula is that it might generate multimodal conditional distributions for certain conditioning levels.

2.3 Score, Hessian and information matrix

Before developing our proposed tests, though, it is convenient to obtain generic expressions for the score, Hessian and information matrix of a K -dimensional copula derived from a multivariate distribution of the same size.

To do so, let $f_K(\boldsymbol{\varepsilon}; \boldsymbol{\rho}, \boldsymbol{\varphi})$ denote the joint density of $\boldsymbol{\varepsilon}$, which we assume is well defined. The $p + q$ parameters of interest are $\boldsymbol{\phi} = (\boldsymbol{\rho}', \boldsymbol{\varphi}')'$, which include the correlation parameters $\boldsymbol{\rho}$ and the shape parameters $\boldsymbol{\varphi}$. Similarly, let $f_{1k}(\varepsilon; \boldsymbol{\varphi})$ denote the marginal density of the k^{th} element of the multivariate nesting distribution, which is such that

$$\int_{-\infty}^{\varepsilon_k} f_{1k}(\varepsilon; \boldsymbol{\varphi}) d\varepsilon = u_k$$

where $u_k(\boldsymbol{\varphi}) = F_{1k}(\varepsilon_k; \boldsymbol{\varphi})$ is the probability integral transform of the k^{th} observed variable with respect to its assumed marginal distribution. Finally, let $\boldsymbol{\phi}_0$ denote the true value of the parameter vector $\boldsymbol{\phi}$.

The log-likelihood function of the copula for a sample of size N for those values of $\boldsymbol{\rho}$ for which $\mathbf{P}(\boldsymbol{\rho})$ has full rank will take the form $L_N(\boldsymbol{\phi}) = \sum_{n=1}^N \ln c(\mathbf{u}_n; \boldsymbol{\phi})$,

$$\ln c(\mathbf{u}; \boldsymbol{\rho}, \boldsymbol{\varphi}) = \ln f_K [F_{11}^{-1}(u_1; \boldsymbol{\varphi}), \dots, F_{1K}^{-1}(u_K; \boldsymbol{\varphi}); \boldsymbol{\rho}, \boldsymbol{\varphi}] - \sum_{k=1}^K \ln f_{1k} [F_{1k}^{-1}(u_k; \boldsymbol{\varphi}); \boldsymbol{\varphi}]. \quad (5)$$

Let $\mathbf{s}(\boldsymbol{\phi})$ denote the score function $\partial \ln c(\mathbf{u}; \boldsymbol{\phi}) / \partial \boldsymbol{\phi}$, and partition it into two blocks, $\mathbf{s}_\rho(\boldsymbol{\phi})$ and $\mathbf{s}_\varphi(\boldsymbol{\phi})$, whose dimensions conform to those of $\boldsymbol{\rho}$ and $\boldsymbol{\varphi}$, respectively. If $\mathbf{P}^{1/2}(\boldsymbol{\rho})$ and $\ln f(\boldsymbol{\varepsilon}; \boldsymbol{\rho}, \boldsymbol{\varphi})$ are differentiable, then

$$\mathbf{s}_\rho(\boldsymbol{\rho}, \boldsymbol{\varphi}) = \frac{\partial \ln f_K [F_{11}^{-1}(u_1; \boldsymbol{\varphi}), \dots, F_{1K}^{-1}(u_K; \boldsymbol{\varphi}); \boldsymbol{\phi}]}{\partial \boldsymbol{\rho}} = -\mathbf{Z}_s(\boldsymbol{\rho}) \mathbf{e}_s(\boldsymbol{\phi})$$

where

$$\begin{aligned}\mathbf{Z}_s(\boldsymbol{\rho}) &= \frac{\partial \text{vec}'[\mathbf{P}^{1/2}(\boldsymbol{\rho})]}{\partial \boldsymbol{\rho}} \cdot [\mathbf{I}_K \otimes \mathbf{P}^{-1/2'}(\boldsymbol{\rho})] \\ \mathbf{e}_s(\boldsymbol{\phi}) &= \text{vec} \left\{ \mathbf{I}_K + \frac{\partial \ln f[\boldsymbol{\varepsilon}^*(\boldsymbol{\rho}); \boldsymbol{\varphi}]}{\partial \boldsymbol{\varepsilon}^*} \cdot \boldsymbol{\varepsilon}^{*'}(\boldsymbol{\rho}) \right\}, \\ \boldsymbol{\varepsilon}^*(\boldsymbol{\rho}) &= \mathbf{P}^{-1/2}(\boldsymbol{\rho}) \boldsymbol{\varepsilon}_t,\end{aligned}$$

because $\boldsymbol{\rho}$ only enters through the joint distribution and not through the marginals or the quantile functions.

On the other hand,

$$\begin{aligned}\mathbf{s}_\varphi(\boldsymbol{\rho}, \boldsymbol{\varphi}) &= \frac{\partial \ln f_K [F_{11}^{-1}(u_1; \boldsymbol{\varphi}), \dots, F_{1K}^{-1}(u_K; \boldsymbol{\varphi}); \boldsymbol{\phi}]}{\partial \boldsymbol{\varphi}} - \sum_{k=1}^K \frac{\partial \ln f_1 [F_1^{-1}(u_k; \boldsymbol{\varphi}); \boldsymbol{\varphi}]}{\partial \boldsymbol{\varphi}} \\ &= \frac{\partial \ln f_K [y_1, \dots, y_K; \boldsymbol{\phi}]}{\partial \boldsymbol{\varphi}} - \sum_{k=1}^K \frac{\partial \ln f_1 (y_k; \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} \\ &\quad + \sum_{k=1}^K \left(\frac{\partial \ln f_K (y_1, \dots, y_K; \boldsymbol{\phi})}{\partial y_k} - \frac{\partial \ln f_1 (y_k; \boldsymbol{\varphi})}{\partial y_k} \right) \frac{\partial F_1^{-1}(u_k; \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}}.\end{aligned}\tag{6}$$

Expression (6) decomposes the copula score into three easy to interpret components. The first one corresponds to the score of the joint distribution. The second one to the scores of the K marginal distributions. Finally, for the third component, we have to multiply the difference between the log-derivatives of the joint and marginal distributions with respect to their arguments by the derivatives of the marginal quantile functions with respect to the shape parameters.

Let $\mathbf{h}(\boldsymbol{\phi})$ denote the Hessian function $\partial \mathbf{s}(\boldsymbol{\phi}) / \partial \boldsymbol{\phi}' = \partial^2 \ln c_n(\mathbf{u}; \boldsymbol{\phi}) / \partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'$. Assuming twice differentiability of the different functions involved, we can show that

$$\begin{aligned}\mathbf{h}_{\varphi\varphi}(\boldsymbol{\phi}) &= \frac{\partial \mathbf{s}_\varphi(\boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}'} \\ &= \frac{\partial^2 \ln f_K [F_{11}^{-1}(u_1; \boldsymbol{\varphi}), \dots, F_{1K}^{-1}(u_K; \boldsymbol{\varphi}); \boldsymbol{\rho}, \boldsymbol{\varphi}]}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} - \sum_{k=1}^K \frac{\partial^2 \ln f_{1k} [F_{1k}^{-1}(u_k; \boldsymbol{\varphi}); \boldsymbol{\varphi}]}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} \\ &= \frac{\partial^2 \ln f_K [y_1, \dots, y_K; \boldsymbol{\rho}, \boldsymbol{\varphi}]}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} - \sum_{k=1}^K \frac{\partial^2 \ln f_{1k} [y_k; \boldsymbol{\varphi}]}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} \\ &\quad + 2 \sum_{k=1}^K \frac{\partial F_1^{-1}(u_k; \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} \left(\frac{\partial^2 \ln f_K [y_1, \dots, y_K; \boldsymbol{\rho}, \boldsymbol{\varphi}]}{\partial x_k \partial \boldsymbol{\varphi}'} - \frac{\partial^2 \ln f_{1k} [y_k; \boldsymbol{\varphi}]}{\partial x_k \partial \boldsymbol{\varphi}'} \right) \\ &\quad + \sum_{k=1}^K \sum_{j=1}^K \frac{\partial F_{1k}^{-1}(u_k; \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} \left(\frac{\partial^2 \ln f_K [y_1, \dots, y_K; \boldsymbol{\rho}, \boldsymbol{\varphi}]}{\partial y_k \partial y_j} - \frac{\partial^2 \ln f_{1k} [y_k; \boldsymbol{\varphi}]}{\partial x_k \partial x_j} \right) \frac{\partial F_{1j}^{-1}(u_j; \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}'} \\ &\quad + \sum_{k=1}^K \frac{\partial^2 F_{1k}^{-1}(u_k; \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} \left(\frac{\partial^2 \ln f_K [x_1, \dots, x_K; \boldsymbol{\rho}, \boldsymbol{\varphi}]}{\partial x_k} - \frac{\partial^2 \ln f_{1k} [x_k; \boldsymbol{\varphi}]}{\partial x_k} \right),\end{aligned}\tag{7}$$

$$\mathbf{h}_{\rho\rho}(\phi) = \mathbf{Z}_s(\rho) \frac{\partial \mathbf{e}_s(\phi)}{\partial \rho'} + [\mathbf{e}'_s(\phi) \otimes \mathbf{I}_p] \frac{\partial \text{vec}[\mathbf{Z}_s(\rho)]}{\partial \rho'},$$

and

$$\mathbf{h}_{\rho\varphi}(\phi) = \mathbf{Z}_s(\theta) \partial \mathbf{e}_s(\phi) / \partial \varphi'.$$

Importantly, while $\mathbf{Z}_s(\rho)$, and $\partial \text{vec}[\mathbf{Z}_s(\rho)] / \partial \rho'$ depend on the specification of the correlation structure, the first and second derivatives of $\ln f_K(\mathbf{x}; \rho, \varphi)$ depend on the specific distribution assumed for testing purposes.

Finally, the (minus) expected value of $\mathbf{h}(\phi)$ will give us the information matrix.

2.3.1 The correlation scores and information matrix under Gaussianity

Maximum likelihood estimation of ρ and φ usually requires a numerical optimization procedure. However, given that in terms of the Gaussian ranks the scores with respect to ρ_{kj} ($k, j = 1, \dots, K$) take the form

$$s_{\rho_{kj}}(y_k, y_j; \rho_{kj}) = \frac{(1 + \rho_{kj}^2)y_k y_j - \rho_{kj}(y_k^2 + y_j^2 - 1) - \rho_{kj}^3}{(1 - \rho_{kj}^2)^2}, \quad (8)$$

there are numerically convenient closed-form expressions for the unrestricted ML estimator of ρ under the assumption of normality. Specifically, $\hat{\rho}_{kj}$ will be the real root to the cubic equation $\sum_{i=1}^N s_{\rho_{kj}}(y_{ki}, y_{ji}; \rho_{kj}) = 0$ that leads to the largest log-likelihood value.⁴

In addition, we can prove the block diagonality of the information matrix:

Proposition 1 *The scores $\mathbf{s}_\rho(\rho, \varphi)$ and $\mathbf{s}_\varphi(\rho, \varphi)$ evaluated at $\varphi = \mathbf{0}$ are orthogonal when the true copula is Gaussian.*

This result, which is the analog for copulas of Lemma 2 in Fiorentini and Sentana (2010), is particularly convenient for our purposes because it allows us to evaluate our tests at any root- N consistent estimator of ρ without having to adjust the asymptotic variance of $\mathbf{s}_\varphi(\phi)$ for parameter uncertainty. Obvious consistent estimators are $\check{\rho}_{kj} = N^{-1} \sum_{i=1}^N y_{ki} y_{ji}$ and $\tilde{\rho}_{kj} = \check{\rho}_{kj} / \sqrt{\tilde{\sigma}_k^2 \tilde{\sigma}_j^2}$, with $\tilde{\sigma}_k^2 = N^{-1} \sum_{i=1}^N y_{ki}^2$, which are the sample covariance and correlation coefficients of the Gaussian ranks y_k and y_j , respectively. Nevertheless, these estimators are generally inefficient relative to the ML estimator:

⁴Algebraic solutions to any cubic equation have been available since at least the early 16th century even though they remain relatively unknown. What is well known, though, is that every cubic equation with real coefficients has at least one real solution, while the other two solutions can be either real or a pair of complex conjugates. In Section 2.5 we revisit the uniqueness of the real root to the cubic equation $\sum_{i=1}^N s_{\rho_{kj}}(y_{ki}, y_{ji}; \rho_{kj}) = 0$.

Proposition 2 *When the copula is Gaussian and the margins are known, the asymptotic variances of $\hat{\rho}_{kj}$, $\tilde{\rho}_{kj}$ and $\check{\rho}_{kj}$, which are the ML estimator of ρ_{kj} , and the sample correlation and covariance coefficients of the Gaussian ranks, respectively, are*

$$AVar(\hat{\rho}_{kj}) = \frac{(1 - \rho_{kj}^2)^2}{1 + \rho_{kj}^2}, \quad AVar(\tilde{\rho}_{kj}) = (1 - \rho_{kj}^2)^2, \quad AVar(\check{\rho}_{kj}) = \frac{1}{1 + \rho_{kj}^2},$$

so that $AVar(\hat{\rho}_{kj}) \leq AVar(\tilde{\rho}_{kj}) \leq AVar(\check{\rho}_{kj})$, with equality if and only if $\rho_{kj} = 0$.

2.4 Score tests

2.4.1 Test against Student t copulas

Under the null hypothesis, $\varsigma(\boldsymbol{\rho}_0)$ will be distributed as a χ^2 random variable with K degrees of freedom (see Malevergne and Sornette (2003)). Let

$$L_2[\varsigma(\boldsymbol{\rho})] = \frac{1}{4}\varsigma^2(\boldsymbol{\rho}) - \frac{K+2}{2}\varsigma(\boldsymbol{\rho}) + \frac{K(K+2)}{4}$$

denote the second Laguerre polynomial associated to this special case of a gamma random variable. Finally, let $\mathbf{y}'_{(k)} = (y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_K)$, $\mathbf{P}_{(kj)}(\boldsymbol{\rho})$ the $(K-1) \times (K-1)$ matrix obtained from $\mathbf{P}(\boldsymbol{\rho})$ after suppressing row k and column j , and $\mathbf{p}_{(k)}(\boldsymbol{\rho})$ the coefficients in the theoretical least squares projection of y_k on to (the linear span of) $\mathbf{y}_{(k)}$.

Proposition 3 *The score of the Student t copula with respect to the reciprocal of the degrees of freedom parameter η when $\eta = 0$ is given by*

$$\begin{aligned} s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, 0) &= \sqrt{\frac{K(K+2)}{2}} L_2[\varsigma(\boldsymbol{\rho})] - \sqrt{\frac{3}{2}} \sum_{k=1}^K H_4(y_k) \\ &\quad + \frac{1}{2} \sqrt{\frac{3}{2}} \sum_{k=1}^K \left[\frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) [\mathbf{y}_{(k)} - \mathbf{p}_{(k)}(\boldsymbol{\rho}) y_k]}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} \right] H_3(y_k). \end{aligned} \quad (9)$$

Therefore, the LM test will simply be given by N times the square of the sample average of (9) evaluated at some consistent estimator of $\boldsymbol{\rho}$ divided by the variance of this score, whose expression can be found in Supplemental Appendix E.3.

It is important to mention that the fact that $\eta = 0$ lies at the boundary of the admissible parameter space invalidates the usual χ_1^2 distribution of the likelihood ratio (LR) and Wald (W) tests, which under the null will be more concentrated towards the origin (see Andrews (2001) and the references therein). The intuition can be perhaps more easily obtained in terms of the W test. Given that $\hat{\eta}_N$ cannot be negative, $\sqrt{N}\hat{\eta}_N$ will have a half-normal asymptotic distribution under

the null (Andrews (1999)). As a result, the W test will be an equally weighted mixture of a chi-square distribution with 0 degrees of freedom (by convention, χ_0^2 is a degenerate random variable that equals zero with probability 1), and a chi-square distribution with 1 degree of freedom. In practice, we simply need to compare the appropriate t -statistic with the appropriate one-sided critical value from the normal tables. For analogous reasons, the asymptotic distribution of the LR test will also be degenerate half the time, and a χ_1^2 the other half.

Although the above argument does not invalidate the distribution of the LM test statistic, intuition suggests that the one-sided nature of the alternative hypothesis should be taken into account to obtain a more powerful test. For that reason, we also propose a simple one-sided version of the LM test for multivariate normality. In particular, since $E[s_\eta(\boldsymbol{\rho}_0, 0)|\boldsymbol{\phi}_0] > 0$ when $\eta_0 > 0$, we suggest to use the LM test statistic when the sample average of the score is positive and 0 otherwise as our one-sided LM test, and to compare it to the same 50:50 mixture of chi-squares 0 and 1. In this context, we would reject H_0 at the $100\kappa\%$ significance level if the average score with respect to η evaluated under the Gaussian null is positive *and* the LM statistic exceeds the $100(1 - 2\kappa)$ percentile of a χ_1^2 distribution. Since the Kuhn-Tucker (KT) multiplier associated with the inequality restriction $\eta \geq 0$ is equal to $\max[-N^{-1} \sum_n s_{\eta n}(\hat{\boldsymbol{\rho}}_N, 0), 0]$, our proposed one-sided LM test is equivalent to the KT multiplier test introduced by Gourieroux, Holly and Monfort (1980), which in turn is equivalent in large samples to the LR and W tests. As we argued before, the reason is that those tests are implicitly one-sided in our context. In this respect, it is important to mention that when there is a single restriction, such as in our case, those one-sided tests would be asymptotically locally more powerful (Andrews (2001)).

It is interesting to compare our score test of the Student t copula to the corresponding test of the Student t distribution. Following Fiorentini, Sentana and Calzolari (2003), the score for η under the null is proportional to $L_2[\varsigma(\boldsymbol{\rho})]$. In the bivariate case with $\boldsymbol{\rho} = 0$ this distributional score becomes

$$H_4(y_1) + H_4(y_2) + H_2(y_1)H_2(y_2)$$

while the copula score reduces to

$$H_2(y_1)H_2(y_2).$$

Given that the Gaussian ranks satisfy the additional moment conditions $E[H_4(y_1)] = 0$ and $E[H_4(y_2)] = 0$ by construction, irrespective of the copula being Gaussian, including those additional marginal terms necessarily reduces (local) power.

Nevertheless, it is also important to emphasize that (9) is computed in terms of the Gaussian ranks y_1, \dots, y_K while the score of the joint Student t distribution is computed in terms of the innovations of a multivariate conditionally heteroskedastic regression model. As a result, the copula test is not generally invariant to affine multivariate transformations of the observed series, unlike the distribution test. The reason is that in the case of a copula, the original variables are of direct interest.

2.4.2 Tests against asymmetric Student t copulas

The asymmetric Student t distribution is a location-scale mixture of normals in which the mixing variable follows a reciprocal gamma distribution, which is a special case of the *GIG*. The derivation of the Lagrange multiplier (LM) test for a multivariate normal copula versus an asymmetric one is complicated by the fact that \mathbf{b} drops out from both the joint and marginal distributions when $\eta \rightarrow 0$ (see Mencía and Sentana (2012)). One standard solution in the literature to deal with testing situations with underidentified parameters under the null involves fixing those parameters to some arbitrary values, and then computing the appropriate test statistic for the chosen values.

Proposition 4 *The score of the asymmetric Student t copula with respect to the reciprocal of the degrees of freedom parameter η when $\eta = 0$ for fixed values of the skewness parameters \mathbf{b} is given by*

$$\begin{aligned} s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, \eta, \mathbf{b}) &= s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, 0) + \mathbf{b}' \mathbf{y} [\varsigma(\boldsymbol{\rho}) - (K + 2)] - \sqrt{6} \sum_{k=1}^K [\beta_k(\boldsymbol{\rho}, \mathbf{b}) H_3(y_k)] \\ &\quad + \sqrt{2} \sum_{k=1}^K \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) [\mathbf{y}_{(k)} - \mathbf{p}_{(k)}(\boldsymbol{\rho}) y_k]}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} \beta_k(\boldsymbol{\rho}, \mathbf{b}) H_2(y_k). \end{aligned} \quad (10)$$

where $\beta(\boldsymbol{\rho}, \mathbf{b}) = \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}$.

On this basis, it would be straightforward to develop the associated test statistic, $LM_N(\mathbf{b})$. However, the previous approach is plausible in situations where there are values of the underidentified parameters that make sense from an economic or statistical point of view. Unfortunately, it is not at all clear a priori what values of \mathbf{b} are likely to prevail under the alternative of *GH* innovations. For that reason, we consider instead a second approach, which consists in computing the LM test for all possible values of \mathbf{b} , and then take the supremum over those parameter values.

It turns out that we can maximize $LM_N(\mathbf{b})$ with respect to \mathbf{b} in closed form, and also obtain the asymptotic distribution of the resulting test statistic:

Proposition 5 *The supremum with respect to \mathbf{b} of the LM tests based on (10) is equal to the sum of two asymptotically independent components: the symmetric Student t LM test based on (9), and a moment test based on the following K influence functions*

$$m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, \eta) = y_k [\varsigma(\boldsymbol{\rho}) - (K + 2)] - \sqrt{6} \sum_{j=1}^K \mathbf{P}_{kj}(\boldsymbol{\rho}) H_3(y_j) + \sqrt{2} \sum_{j=1}^K \frac{\mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho}) [\mathbf{y}_{(j)} - \mathbf{p}_{(j)}(\boldsymbol{\rho}) y_j]}{1 - \mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(j)}(\boldsymbol{\rho})} \mathbf{P}_{kj}(\boldsymbol{\rho}) H_2(y_j). \quad (11)$$

This second moment test is asymptotically distributed as a χ^2 distribution with K degrees of freedom when the true copula is Gaussian.

Given that $s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, \eta)$ is orthogonal to the other K moment conditions, we can conduct a partially one-sided test by combining the KT one-sided version of the Student t test and the moment test based on $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, \eta)$. By analogy with Mencía and Sentana (2012), this one-sided version should be equivalent in large samples to the corresponding LR test. The asymptotic distribution of the joint test under the null will be a 50:50 mixture of χ_K^2 and χ_{K+1}^2 .

Once again, it is interesting to compare the moments of the asymmetric Student t copula associated to b_k under normality to the corresponding moments for the asymmetric Student t distribution in Mencía and Sentana (2012). In the bivariate case with $\rho = 0$ the influence functions of the distribution test corresponding to the asymmetric components are

$$H_3(y_1) + H_2(y_1) H_1(y_2),$$

$$H_3(y_2) + H_2(y_2) H_1(y_1),$$

while in the copula test they reduce to

$$H_2(y_1) H_1(y_2),$$

$$H_2(y_2) H_1(y_1).$$

Given that the Gaussian ranks satisfy the additional moment conditions $E[H_3(y_1)] = 0$ and $E[H_3(y_2)] = 0$, by construction, irrespective of the copula being Gaussian, the inclusion of those terms necessarily reduces (local) power, as in the symmetric component case.

Interestingly, the K moment conditions $E[m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, \eta)] = 0$ can also be used to consistently test the hypothesis of a symmetric Student copula against an asymmetric one because we can show that the expected values of those influence functions would remain zero under this new null. But the test will be incorrectly sized if we used the covariance matrix of (11) under Gaussianity. To avoid size distortions, we can either compute the correct covariance expression by numerical quadrature or Monte Carlo integration for a given value of η , or else run the univariate regression of 1 on $m_{b_1}(\hat{\boldsymbol{\rho}}_T), \dots, m_{b_K}(\hat{\boldsymbol{\rho}}_T)$. We use this second approach in the empirical application.

2.4.3 Tests against Generalized Hyperbolic copulas

As discussed by Mencía and Sentana (2012), there are three different paths along which a symmetric GH distribution converges to a Gaussian distribution. Specifically, the normal distribution can be achieved when (i) $\nu \rightarrow -\infty$ or (ii) $\nu \rightarrow +\infty$, regardless of the value of γ ; and (iii) $\gamma \rightarrow \infty$ irrespective of the value of ν . In addition, one of the shape parameters becomes increasingly underidentified when the other one is on a normality path. Nevertheless, Mencía and Sentana (2012) showed that the score of the remaining identified parameter evaluated under the null of normality is (proportional to) (9) along those three paths. As a result, the LM/KT tests of Gaussian copula against a “symmetric” GH copula are numerically identical to the LM/KT tests against symmetric Student t in Proposition 3.

Since the same is true for asymmetric GH alternatives, the LM/KT tests of Gaussian copula against an “asymmetric” GH copula will also be numerically identical to the LM/KT tests against asymmetric Student t in Proposition 5.

2.4.4 Tests against Hermite expansions of the Gaussian copula

The econometric complication here is that some linear combinations of the scores with respect to the original coefficient vector $\boldsymbol{\varphi} = (c_1, c_2, c_3, c_4, d_1, d_2, d_3, d_4, d_5)'$ in (4) are identically 0 under the null, which means that we cannot compute the usual LM test for $H_0 : \boldsymbol{\varphi} = \mathbf{0}$. To deal with this unusual type of testing situation, Lee and Chesher (1986) propose to replace the LM test by what they call an “extremum test” (see also Bera, Ra, and N. Sarkar (1998)).⁵ Given that the first-order conditions are identically 0, their suggestion is to study the restrictions that the null imposes on higher order conditions. Lee and Chesher (1986) also show that the same

⁵Neyman and Scott (1966) considered similar problems in the context of the $C(\alpha)$ statistic.

test can often be achieved by a suitable reparametrization. Specifically, if we re-write the model in terms of γ_1 , γ_4 and δ_5 , where

$$\gamma_1 = \sqrt{c_1 - \frac{\rho}{\sqrt{3}}c_2}, \quad \gamma_4 = \sqrt{c_4 - \frac{\rho}{\sqrt{3}}c_3}, \quad \delta_5 = \sqrt{d_5 - \frac{\rho}{2}d_4 + \frac{\rho}{2}d_2 - d_1},$$

we can show that the scores with respect to these three new parameters are one-half the second derivatives of the log-likelihood function with respect to γ_1^2 , γ_4^2 and δ_5^2 .

In Proposition 10 in Appendix B we provide the influence functions which are the basis for our LM test of the null hypothesis of a Gaussian copula against a Hermite polynomial expansion.⁶ From a practical point of view, the main problem with this test is that it involves many degrees of freedom: 4 to capture skewness and 5 for kurtosis. Moreover, the number of singularities increases if we maintain the assumption that $\rho = 0$. In particular, 2 additional skewness parameters and 1 kurtosis one become underidentified in that case. For that reason, we suggest a modified test for this situation in Proposition 11 in Appendix B.

2.5 Replacing margins with empirical cdf's

Marginal distributions are rarely known in practice. The most common solution is a two-step estimation procedure, whereby the margins are replaced by their (re-scaled) empirical cdf counterparts. In this manner, the proposed tests can be viewed as functions of the collection of Gaussian ranks obtained from the uniform ranks, where the scaling factor $N/(N+1)$ is only introduced to avoid potential problems with the copula density blowing up at the boundary of $[0, 1]^K$. Smoothed versions of the empirical cdf can also be used, but the effects should be the same (up to first-order).

The use of sample ranks has two implications. First, the exact discrete uniform nature of their distribution simplifies some of the previous expressions. Specifically, the sample averages of all the odd-order Hermite polynomials of the Gaussian ranks will be identically zero, while the sample averages of the even-order ones will converge to zero at faster than square root N rates. Among other things, this in turn implies that the real solution to the cubic equation $\sum_{i=1}^N s_{\rho_{kj}}(y_{ki}, y_{ji}; \hat{\rho}_{kj}) = 0$ which defines the unrestricted ML estimator of ρ_{ij} will be unique for $N > 6$ (see footnote 4).

Second, it effectively transforms the Gaussian ML estimation procedure we have considered so

⁶The Lee and Chesher (1986) test effectively imposes one-sided restrictions on γ_1 , γ_4 and δ_5 because the scores with respect to γ_1 and $-\gamma_1$, say, are identical under the null. Nevertheless, given that we also ignore the positivity restrictions that may affect some of the remaining parameters in our analysis, we focus on two-sided LM tests.

far into a sequential semiparametric procedure, which requires us to take into account the sample uncertainty resulting from its non-parametric first-stage (see Newey and McFadden (1994)). Otherwise, our test statistics will have size distortions even in large samples.

Following Chen and Fan (2006a), we can obtain the variance of a generic influence function $m_\phi(u_1, \dots, u_K)$ adjusted for non-parametric estimation of the margins by computing the variance of the adjusted function

$$m_\phi^c(u_1, \dots, u_K; \boldsymbol{\rho}, \mathbf{0}) = m_\phi(u_1, \dots, u_K; \boldsymbol{\rho}, \mathbf{0}) - n_\phi(u_1, \dots, u_K; \boldsymbol{\rho}),$$

where

$$n_\phi = \sum_{j=1}^K \int_0^1 [1\{U_j \leq u_j\} - u_j] W_{\phi_k}^j du_j,$$

with

$$W_\phi^j = \int \dots \int \frac{\partial m_\phi(u_1, \dots, u_K)}{\partial u_j} c(u_1, \dots, u_K; \boldsymbol{\phi}) du_1 \dots du_{k-1} du_{k+1} \dots du_K.$$

In the case of the correlation parameters, one can capture the inflation in variance by adding linear combinations of second order Hermite polynomials of each of the variables (in Gaussian form) to the original scores $\mathbf{s}_\rho(\boldsymbol{\rho}, \mathbf{0})$, as the following proposition shows:

Proposition 6 *The correction to $s_{\rho_{kj}}(u_k, u_j; \rho_{kj})$ is given by*

$$n_{\rho_{kj}}(u_k, u_j; \rho_{kj}) = -\frac{1}{\sqrt{2}} \frac{\rho_{kj}}{1 - \rho_{kj}^2} [H_2(y_k) + H_2(y_j)].$$

On this basis, we obtain the following modified version of Proposition 2:

Proposition 7 *When the copula is Gaussian and the margins are replaced by their empirical cdfs, the asymptotic variances of $\hat{\rho}_{kj}$, $\tilde{\rho}_{kj}$ and $\check{\rho}_{kj}$, which are the ML estimator of ρ_{kj} , and the sample correlation and covariance coefficients of the Gaussian ranks, respectively, are given by*

$$AVar(\hat{\rho}_{kj}) = AVar(\tilde{\rho}_{kj}) = AVar(\check{\rho}_{kj}) = (1 - \rho_{kj}^2)^2.$$

Similarly, for the distributions that we consider as alternatives, we show that it is possible to capture the variance modification in the scores of the shape parameters, $\mathbf{s}_\varphi(\boldsymbol{\rho}, \mathbf{0})$, by adding linear combinations of third and fourth Hermite polynomials in those variables. The following result provides the expressions for the corrections corresponding to Propositions 3, 4 and 5:

Proposition 8 *The correction of $s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, \mathbf{0})$ is given by*

$$\begin{aligned} n_\eta(u_1, \dots, u_K; \boldsymbol{\rho}) &= -\frac{1}{4}\sqrt{\frac{3}{2}} \sum_{k=1}^K \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(k)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(k)}(\boldsymbol{\rho})} H_4(y_k) \\ &\quad + \frac{1}{4}\sqrt{\frac{3}{2}} \sum_{k=1}^K \sum_{h \neq k} \frac{\mathbf{p}'_{(h)}(\boldsymbol{\rho})\mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(h)}(\boldsymbol{\rho})\mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(h)}(\boldsymbol{\rho})} \mathbf{P}_{kh}^3(\boldsymbol{\rho}) H_4(y_k), \end{aligned}$$

the correction of $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, \eta)$ by

$$\begin{aligned} n_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}) &= -\sqrt{\frac{2}{3}} \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(k)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(k)}(\boldsymbol{\rho})} H_3(y_k) \\ &\quad + \sqrt{\frac{2}{3}} \sum_{j \neq k} \frac{\mathbf{p}'_{(j)}(\boldsymbol{\rho})\mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(j)}(\boldsymbol{\rho})\mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(j)}(\boldsymbol{\rho})} \mathbf{P}_{kj}^2(\boldsymbol{\rho}) H_3(y_k) \\ &\quad - \sqrt{\frac{2}{3}} \sum_{j \neq k} \frac{\mathbf{p}'_{(j)}(\boldsymbol{\rho})\mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(j)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(j)}(\boldsymbol{\rho})\mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(j)}(\boldsymbol{\rho})} \mathbf{P}_{kj}(\boldsymbol{\rho}) H_3(y_j) \\ &\quad + \sqrt{\frac{2}{3}} \sum_{j \neq k} \sum_{h \neq j} \frac{\mathbf{p}'_{(h)}(\boldsymbol{\rho})\mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(h)}(\boldsymbol{\rho})\mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(h)}(\boldsymbol{\rho})} \mathbf{P}_{kh}(\boldsymbol{\rho}) \mathbf{P}_{jh}^2(\boldsymbol{\rho}) H_3(y_j), \end{aligned}$$

and the correction of $s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, \eta, \mathbf{b})$ by

$$n_\eta(u_1, \dots, u_K; \boldsymbol{\rho}) + \sum_{k=1}^K b_k n_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}).$$

For example, for the bivariate asymmetric Student t , the score corrections will be

$$\begin{aligned} n_\eta(y_1, y_2; \rho) &= -\frac{1}{4}\sqrt{\frac{3}{2}}\rho^2 [H_4(y_1) + H_4(y_2)], \\ n_{b_1}(y_1, y_2; \rho) &= \sqrt{\frac{2}{3}}\rho^2 [(6 - \rho^2)H_3(y_1) + 5\rho H_3(y_2)] \end{aligned}$$

and

$$n_{b_1}(y_1, y_2; \rho) = \sqrt{\frac{2}{3}}\rho^2 [(6 - \rho^2)H_3(y_2) + 5\rho H_3(y_1)]$$

In turn, the correction terms for the influence functions of Propositions 10 and 11 are given in Proposition 12 in Appendix B.

Importantly, since Hermite polynomials form an orthonormal basis under Gaussianity, the orthogonality between the original scores for correlation and shape parameters stated in Proposition 1 is preserved in the modified scores.

2.6 Power comparisons

We can assess the power of the different score tests that we have introduced in previous sections by computing the probability of rejecting the null hypothesis when it is false as a

function of the shape parameters φ under the assumption that the asymptotic non-central chi-square distributions of the different LM and KT tests provide reliable rejection probabilities in finite samples. In this regard, we explain in appendix D the way in which we compute the non-centrality parameters of our proposed tests, as well as the non-centrality parameters of distributional tests of Gaussian vs Student t and Gaussian vs asymmetric Student t , which ignore that the margins of the copula are uniform by construction. But given that the degrees of freedom are the same for copula and distributional tests, we can directly compare the non-centrality parameters instead.

Figures 7a-c depict the non-centrality parameters of symmetric Student t tests under asymmetric Student t local alternatives, while Figures 8a-c do the same for asymmetric Student t tests. In those plots, LM and LM^{NP} denote the LM-copula tests when the marginals are known and when they are estimated nonparametrically, respectively, while $Dist^{NP}$ indicates the LM-distributional test when the margins are estimated nonparametrically.

In Figures 7a and 8a we have represented η in the x -axis for fixed values of $\rho = .75$ and $b_i = 0$. As can be seen, the distributional tests have less power than the copula tests when marginals are estimated nonparametrically, which in turn have less power than the copula tests when the margins are known.

We then look at the non-centrality parameters for different values of ρ in the x -axis for fixed values of $\eta = .1$ and $b_i = -.5$ in Figures 7b and 8b. As expected, as ρ approaches zero, LM , LM^{NP} and $Dist^{NP}$ tend to have the same power.

Finally, we plot the non-centrality parameters against asymmetric Student t alternatives with increasing skewness when $\eta = .1$ and $\rho = .75$. Not surprisingly, the Student t tests are not sensitive to the different values of b (Figure 7c), while the asymmetric Student t tests have more power as b moves away from zero.

2.7 Constrained indirect estimation

If a researcher who uses our proposed tests does not reject the null hypothesis, she can rely on the Gaussian copula evaluated at the Gaussian rank correlation coefficients. However, if she rejects, she might be interested in estimating the parameters of the alternative distributions that we have considered. Before doing so, though, it is important to remember that many empirical researchers continue to use the Gaussian pseudo-maximum likelihood (PML) estimators advocated by Bollerslev and Wooldridge (1992) in multivariate location-scale models with non-

normal observations because they remain consistent for the (conditional) mean and variance parameters as long as those moments are correctly specified. The following result characterizes the analogous property for Gaussian copulas:

Proposition 9 *The Gaussian pseudo-ML estimators of the $p \times 1$ vector of parameters $\boldsymbol{\rho}$ characterizing the potentially restricted, but correctly specified correlation matrix $\mathbf{P}(\boldsymbol{\rho})$ of the Gaussian ranks remain consistent even when the copula is not Gaussian.*

The same would be true of the $K(K - 1)/2$ sample Gaussian rank correlation coefficients $\tilde{\rho}_{kj}$ if $\mathbf{P}(\boldsymbol{\rho})$ were unrestricted.

Conceptually, the most straightforward procedure for estimating the parameters of the non-Gaussian copulas we have considered would be maximum likelihood using the analytical expressions for the scores that we have derived. Unfortunately, this is easier said than done because those analytical expressions are computationally involved under the alternative (see the supplemental appendix to Mencía and Sentana (2012) for some of the required expressions).

Nevertheless, it is possible to come up with much simpler consistent estimators of $\boldsymbol{\rho}$ and $\boldsymbol{\varphi}$ along the lines of Calzolari, Fiorentini and Sentana (2004). Specifically, we can estimate those coefficients for a specific parametric copula by generating data from this copula and matching in the simulated data the values in the original data of both the Gaussian rank correlation coefficients and the test statistics we have proposed. Proposition 4 in Calzolari, Fiorentini and Sentana (2004) guarantees the consistency and asymptotic normality of these constrained indirect estimators. Similarly, their Proposition 7 characterizes the efficiency loss of these estimators relative to MLE.

3 Monte Carlo evidence

In this section, we assess the finite sample size and power properties of the testing procedures discussed above by means of several extensive Monte Carlo exercises. Given that all our tests are numerically invariant to strictly monotonic transformations of the data, without loss of generality we can work directly with observations $\boldsymbol{\varepsilon}$ generated from the relevant multivariate distribution regardless of the marginal distribution we would choose, which saves CPU time.

3.1 Design and estimation details

For simplicity, we look at bivariate and trivariate models. In the bivariate case, the correlation coefficient of the Gaussian copula is $\rho = .12$, which matches the one used in Figure 3. In contrast, we use $\rho_{12} = .013$, $\rho_{13} = .041$, and $\rho_{23} = .014$ in the trivariate case, which are motivated by some of the empirical estimates.

For our first alternative, we draw from a (symmetric) bivariate or trivariate Student t distribution with the same correlation matrices. In addition, we generate observations from an asymmetric Student t with identical correlation matrices but negative tail dependence ($b_i = -.5$).

Given that the asymptotic distributions that we have derived in previous sections may be unreliable in finite samples, we also compute bootstrap critical values. Specifically, we employ a parametric bootstrap procedure with 10,000 simulated samples for all tests.⁷ Despite the asymptotic orthogonality of the scores corresponding to correlation and shape parameters, our bootstrap procedure takes into account the sensitivity of the critical values to the values of ρ in order not to rule out higher order refinements (see Supplemental appendix E.1 for further details).

Importantly, we compare our proposed score tests to the Kolmogorov–Smirnov and Cramér–von Mises tests for copula models, which are often reported in empirical work (see Rémillard (2010) for details). Since the asymptotic distributions of these tests in copula models with non-parametric margins are unknown, we must exclusively rely on their bootstrap values.

3.2 Size properties

The first question that we need to address is whether the asymptotic distribution under the null attributed to the test statistics introduced in section 2 is reliable in finite samples.

Figures 9a-b show the dependence of the rejection rates on ρ in the bivariate case. As expected, the asymptotic p-values are very reasonable in samples of size 3,200, when they hardly depend on ρ . For smaller sample sizes, though, this is not the case. In contrast, Table 1 indicates that the parametric bootstrap rejection rates are close to being perfect for all the different samples sizes and significance levels we consider.

⁷For the Hermite expansion based test we use the influence functions given in Proposition 10 whenever $|\hat{\rho}| \geq .02$ and the ones in Proposition 11 otherwise. In our simulations, though, $|\hat{\rho}| < .02$ only happens 2.6% of the time for $N = 200$, and 0.04% for $N = 800$.

3.3 Power properties

Table 2 reports the Monte Carlo rejection rates at the 1%, 5% and 10% significance levels for a Student t alternative with 20 degrees of freedom. By and large, the behavior of the different test statistics is in accordance with what the asymptotic results would suggest. In particular, the Student t tests are the most powerful, with the Kuhn-Tucker versions being more powerful than the Lagrange multiplier’s ones. Interestingly, the test against a Hermite expansion yields decent rejection rates. In contrast, both the Kolmogorov–Smirnov and Cramér–von Mises tests have close to trivial power in samples of 200 observations.

In turn, Table 3 reports the Monte Carlo rejection rates at the same levels for an asymmetric Student t alternative with 20 degrees of freedom and asymmetry parameters $b_i = -.5$. The only surprising result is the fact that in the bivariate case LM- t is slightly more powerful than LM- At in samples of 200 observations when the DGP is an asymmetric Student t . In all other cases, the ranking suggested by our asymptotic results is preserved.

4 Momentum and reversals in stock returns

In this section we apply our Gaussian copula tests to formally analyze the cross-sectional dependence between monthly returns on individual U.S. stocks in the CRSP database and some of their observable characteristics. Given their prominence in the empirical finance literature, we focus on short term reversals, in which the observable characteristic are the individual stock returns over the previous month, and momentum, where the relevant variable are the individual stock returns from month $t - 2$ to month $t - 12$. In line with most previous studies (see e.g. Asness, Moskowitz and Pedersen (2013)), we only consider common equities (CRSP sharecodes 10 and 11) and exclude those stocks with share prices less than \$1 at the beginning of the holding period. We also restrict our analysis to those firms with at least 60 months of return history, so that we focus on liquid stocks with low transaction costs and high tradability.

An important advantage of working with either uniform or Gaussian ranks is that we obtain exactly the same numerical results whether we work with the original returns or with their deviations from the returns on an aggregate stock market index or the level of the risk free rate. Nevertheless, the presence of other time-varying effects that may potentially affect different firms differently could alter the cross-sectional dependence. For that reason, we carry out our analysis

both at the aggregate level using all individual stocks and at the industry level.⁸

Our dataset spans the period from July 1991 to December 2012. For any given month, we have available a cross-section of the form $Y_t = \{(y_{11}^t, y_{21}^t), \dots, (y_{1N_t}^t, y_{2N_t}^t)\}$, where N_t is the number of individual stocks for which we have data on both their return and the relevant observable characteristics. From the point of view of devising trading strategies, though, a period by period analysis is of little interest. For that reason, we pool all $\sum_{t=1}^T N_t$ bivariate observations $\mathbf{Y} = \{Y_1, \dots, Y_T\}$ as follows. First, for each t we compute the cross-sectional average of the log-likelihood score, say $\bar{s}_{\phi t}(Y_t; \rho) = N_t^{-1} \sum_{i=1}^{N_t} s_{\phi}(y_{1i}^t, y_{2i}^t; \rho)$. Then, we time-average those scores, thereby creating a pooled average score $\bar{s}_{\phi}(Y_t; \rho) = T^{-1} \sum_{t=1}^T \bar{s}_{\phi t}(Y_t; \rho)$, on the basis of which we can estimate the correlation coefficient ρ and construct our tests. The only complication is that our pooled procedure requires the computation of robust standard errors to capture the potential time-series dependence in $\bar{s}_{\phi t}(Y_t; \rho)$ for different t 's (see Supplemental Appendix E for details).

Before characterizing dependence through the copula, though, it is convenient to look at correlations. Table 4 presents the parameter estimates and their corresponding asymptotically robust standard errors for the Pearson, Spearman and Gaussian rank correlation coefficients. In this regard, the standard errors have been corrected for heteroskedasticity and autocorrelation (HAC) using Bartlett weights with 5 lags. The left panel corresponds to short term reversals while the right panel contains the correlation estimates for momentum strategies. Both Spearman and Gaussian rank correlations have the expected sign for all the industries when looking at momentum strategies, and the same is true for reversals with the exception of Telecommunications. In contrast, Pearson correlation estimates have the wrong sign for most of the industries, especially for short term reversals.

In Table 5 we report the Gaussian copula test statistics, with KT- t and KT- At denoting the Kuhn-Tucker versions of the tests against Student t and asymmetric Student t copulas,⁹ Skew the Lagrange multiplier test based on the K moment conditions $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, 0)$ in Proposition 4, and LM-HE the Lagrange multiplier test based on the score of the Hermite expansion of the

⁸Industry definitions: Non Durables: Consumer NonDurables – Food, Tobacco, Textiles, Apparel, Leather, Toys; Durables :Consumer Durables – Cars, TV's, Furniture, Household Appliances; Manufacturing: Manufacturing – Machinery, Trucks, Planes, Off Furn, Paper, Com Printing; Energy: Oil, Gas, and Coal Extraction and Products; Chemicals: Chemicals and Allied Products; Business : Business Equipment – Computers, Software, and Electronic Equipment; Telecom: Telephone and Television Transmission; Utilities; Shops: Wholesale, Retail, and Some Services (Laundries, Repair Shops); Healthcare: Healthcare, Medical Equipment, and Drugs; Financials; and Other: Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment. See Ken French's website for details.

⁹We omit the Lagrange multiplier versions since they are numerically identical.

Gaussian copula in Proposition 5. As can be seen, in all cases we reject the null hypothesis of a Gaussian copula for both short term reversals and momentum by a long margin. Importantly, the source of the rejection is not only the “cokurtosis” between the Gaussian ranks, but also their “coskewness”, specially for momentum strategies. In this regard, it is worth emphasizing that the use of the HAC procedure ensures that the asymmetric component of the test is correctly sized under the null of a symmetric Student t copula too, as argued at the end of section 2.4.2.

Given those rejections, the natural next step is to gauge the parameters of the alternative distributions that we have considered. As explained in section 2.7, we can consistently do so by means of an equality constrained indirect estimation procedure which matches the observed tests statistics and the estimated Gaussian rank correlations. We report in Table 6 the resulting estimates of the correlation and shape parameters. We find moderate negative tail dependence but quite substantive “leptokurtosis”, with estimated degrees of freedom in the neighborhood of 5. Therefore, we can safely conclude that it should be possible to devise successful trading strategies that exploit those features of the data.

4.1 Trading implications of a non-Gaussian copula

As mentioned in the introduction, the (rank) correlation is obviously very important in deriving probabilistic predictions about the rank of a stock over the next month given the rank of its characteristic, but it is by no means the only determinant. In general, non-linear tail dependence also matters. To illustrate the importance of looking at the entire copula, we use the parameter estimates for the Gaussian, Student t and asymmetric Student t copulas in Table 6 obtained using the full data set to compute the probabilities that a stock will be in the bottom 30, middle 40 or top 30 percentiles during period t conditional on the same stock being in the bottom 5%, next 25%, middle 40%, next 25% and top 5% according to its short-term reversal or momentum characteristics.¹⁰

Figure 10 presents the results for short-term reversals. As can be observed, the estimated negative correlation is not large enough for the Gaussian copula to suggest any position. In contrast, the non-linear dependence of both the symmetric and asymmetric Student t copulas results in long positions on recent losers (5%) and short positions on recent winners (95%).

Figure 11 contains the result of a similar exercise with momentum strategies. Once again, we

¹⁰See Gagliardini et al. (2014) for a formal discussion of portfolio choice based on the maximization of the expected utility of the ranks.

find that the small positive correlation of the Gaussian copula is too weak to lead to any position. But the non-linear dependence of the symmetric Student t copula changes the probabilities enough to recommend taking short positions on past losers (5%) and long positions on past winners (95%). Somewhat surprisingly, though, the negative tail dependence of the asymmetric Student t in this case, which is higher than for short-term reversals, leads to the opposite trading strategy for the case of winners.

5 Conclusions

We derive computationally simple and intuitive expressions for score tests of Gaussian copulas against GH alternatives, which include the symmetric and asymmetric Student t , and Hermite polynomial expansions. We decompose our tests into moment tests based on linear combinations of third and fourth cross products of Hermite polynomials of the Gaussian ranks. By taking into account the partial one-sided nature of some of the alternative hypotheses, we also obtain more powerful one-sided Kuhn-Tucker versions that are equivalent to the Likelihood Ratio test, whose asymptotic distribution we provide. This equivalence implies that our approach has a likelihood interpretation. Thus, we can learn from our tests in which directions the model is really worth extending.

We conduct detailed Monte Carlo exercises to study our proposed tests in finite samples. We find that the asymptotic p-values are very reasonable in moderately large samples sizes, and that the parametric bootstrap rejection rates are close to being perfect for all samples sizes. Moreover, the finite sample power of the different test statistics agrees with what the asymptotic results would suggest. Importantly, our findings indicate that our parametric tests have substantially more power than the Kolmogorov–Smirnov and Cramér–von Mises tests even for departures for which our procedures are not optimal. This is particularly true of the test against Hermite polynomial expansions, which has a non-parametric flair.

In an empirical application to CRSP data, we re-assess the widely held view that stocks that underperformed in the past month (short term reversals) and stocks that outperformed in previous months (momentum) show superior performance. Our tests indicate that short-term reversals and momentum effects are better captured by non-Gaussian copulas. We devise successful trading strategies based on an asymmetric Student t copula whose parameters we estimate by an indirect inference procedure that matches our test statistics in the simulated and

real data. Although we obtain exactly the same numerical results whether we work with the original returns or with their deviations from the returns on an aggregate stock market index or the level of the risk free rate, we carry out our empirical analysis both at the aggregate level and at the industry level to allow for the presence of time-varying industry effects.

Nevertheless, we could explicitly consider more complex models by conditioning on past values of \mathbf{x} or present and past values of some exogenous variables \mathbf{z} (see e.g. Patton (2006) or Chen and Fan (2006a) for some interesting examples of dynamic copula models).

Another interesting extension would be to develop testing procedures that direct power locally over the third quadrant, say, as opposed to over the entire distribution. In principle, one could do so by using mixtures of normals with the second component located in the third quadrant, or with a GH-based test in which we fix both betas to be big and negative.

It would also be interesting to compare our score tests to information criteria approaches (see e.g. Chen and Fan (2005 and 2006b)), as well as to tests based on non-parametric estimates of the copula density (see Fermanian (2005) and Scaillet (2007)). Alternatively, we could consider as our null hypothesis special cases of the GH distribution, such as the symmetric or asymmetric Student t . All these extensions constitute promising avenues for further research.

References

- Andrews, D.W.K. (1999): “Estimation when a parameter is on the boundary”, *Econometrica*, 67, 1341-1383.
- Andrews, D.W.K. (2001): “Testing when a parameter is on the boundary of the maintained hypothesis”, *Econometrica*, 69, 683-734.
- Asness, C.S., T.J. Moskowitz and L.H. Pedersen (2013): “Value and momentum everywhere”, *Journal of Finance*, 68, 929-985.
- Barndorff-Nielsen, O. (1977): “Exponentially decreasing distributions for the logarithm of particle size”, *Proceedings of the Royal Society*, 353, 401-419.
- Bera, A., S. Ra, and N. Sarkar (1998): “Hypothesis testing for some nonregular cases in econometrics”, in D. Coondoo and R. Mukherjee (eds.), *Econometrics: Theory and Practice*, Buffalo, NY: Allied Publishers.
- Berg, D., J.-F. Quessy (2009): “Local power analyses of goodness-of-fit tests for copulas”, *Scandinavian Journal of Statistics*, 36, 389-412.
- Blæsild, P. (1981): “The two-dimensional hyperbolic distribution and related distributions, with an application to Johanssen’s bean data”, *Biometrika*, 68, 251-263.
- Bollerslev, T., and J. Wooldridge (1992): “Quasi maximum likelihood estimation and inference in dynamic models with time-varying covariances”, *Econometric Reviews*, 11, 143-172.
- Calzolari, G., G. Fiorentini, and E. Sentana (2004): “Constrained indirect estimation”, *Review of Economic Studies*, 71, 945-973.
- Chen, X. and Y. Fan (2005): “Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection”, *Canadian Journal of Statistics*, 33, 389-414.
- Chen, X. and Y. Fan (2006a): “Estimation of copula-based semiparametric time series models”, *Journal of Econometrics*, 130, 307-335.
- Chen, X. and Y. Fan (2006b): “Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification”, *Journal of Econometrics*, 135, 125-154.
- Donnelly, C. and Embrechts, P. (2010): “The devil is in the tails: actuarial mathematics and the subprime mortgage crisis”, *ASTIN Bulletin*, 40, 1-33.
- Fermanian, J.-D. (2005): “Goodness-of-fit tests for copulas”, *Journal of Multivariate Analysis*, 95, 119-152.

Fiorentini, G. and Sentana, E. (2010): “On the efficiency and consistency of likelihood estimation in multivariate conditionally heteroskedastic dynamic regression models”, mimeo.

Fiorentini, G., Sentana, E. and Calzolari, G. (2003): “Maximum likelihood estimation and inference in multivariate conditionally heteroskedastic dynamic regression models with Student t innovations”, *Journal of Business and Economic Statistics*, 21, 532–546.

Gagliardini, P., Gourioux, C. and Rubin, M. (2014): “Positional portfolio management”, mimeo Università della Svizzera Italiana.

Genest, C., B. Rémillard, D. Beaudoin (2009): “Goodness-of-fit tests for copulas: a review and a power study”, *Insurance Mathematics and Economics*, 44, 199–213.

Gourioux C., Holly, A., and Monfort, A. (1980): “Kuhn–Tucker, likelihood ratio and Wald tests for nonlinear models with inequality constraints on the parameters”, Discussion Paper 770, Harvard Institute of Economic Research.

Joe, H. (2015): *Dependence Modeling with Copulas*, CRC Press, Boca Raton.

Jørgensen, B. (1982): *Statistical Properties of the Generalized Inverse Gaussian Distribution*, Springer-Verlag, New York.

Johnson, N., S. Kotz, and N. Balakrishnan (1994): *Continuous Univariate Distributions*, Wiley, New York.

Lee, L. F., and A. Chesher (1986): “Specification testing when score test statistics are identically zero”, *Journal of Econometrics*, 31, 121–149.

Lo, A. (2007): “What happened to the quants in August 2007?”, *Journal of Investment Management*, 5, 5–54.

MacKenzie, D. and T. Spears (2012): “The formula that killed Wall Street? The Gaussian copula and the material cultures of modelling”, mimeo, School of Social & Political Science, University of Edinburgh.

Malevergne, Y. and D. Sornette (2003): “Testing the Gaussian copula hypothesis for financial assets dependencies”, *Quantitative Finance*, 3, 231–250.

Mencía, J., and E. Sentana (2009): “Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation”, *Journal of Econometrics*, 153, 105–121.

Mencía, J., and E. Sentana (2012): “Distributional tests in multivariate dynamic models with Normal and Student t innovations”, *Review of Economics and Statistics*, 94, 133–152.

Newey, W.K. and McFadden, D.L. (1994): “Large sample estimation and hypothesis testing”,

in R.F. Engle and D.L. McFadden (eds.) *Handbook of Econometrics* vol. IV, 2111–2245, Elsevier.

Newey, W.K. and K.D. West (1987): “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix”, *Econometrica*, 55, 703–708.

Neyman, J. and Scott, E.L. (1966): “On the use of optimal C_α -tests of composite hypotheses”, *Bulletin of the International Statistical Institute*, 41, 477–497.

Patton, A.J. (2006): “Modelling asymmetric exchange rate dependence”, *International Economic Review*, 47, 527–56.

Rémillard, B. (2010): “Goodness-of-fit tests for copulas of multivariate time series”, Working Paper, HEC Montréal.

Salmon, F. (2009): “Recipe for disaster: the formula that killed Wall Street”, *Wired*, 23 February, http://www.wired.com/print/techbiz/it/magazine/17-03/wp_quant

Scaillet, O. (2007): “Kernel based goodness-of-fit tests for copulas with fixed smoothing parameters”, *Journal of Multivariate Analysis*, 98, 533–543.

Appendix

A Proofs

Proposition 1

We can use the conditional analogue to the generalized information matrix equality (see e.g. Newey and McFadden (1994)) to show that

$$E \{ \mathbf{s}_\rho(\rho, \mathbf{0}) \mathbf{s}'_\varphi(\rho, \varphi) | \rho, \varphi \} = -E \left\{ \left[\frac{\partial \mathbf{s}_\rho(\rho, \mathbf{0})}{\partial \varphi'} \right] \middle| \rho, \varphi \right\} = \mathbf{0},$$

irrespective of the conditional distribution of ε_t^* , where we have used the fact that $\mathbf{s}_\rho(\rho, \mathbf{0})$ does not vary with φ when regarded as the influence function for $\hat{\rho}_T$. Then, the required result follows from the *i.i.d.* nature of both $\mathbf{e}_s(\rho_0, \mathbf{0})$ and $\mathbf{s}_\varphi(\rho_0, \varphi_0)$. \square

Proposition 2

Under the null of Gaussian copula, it is straightforward to obtain $AVar(\hat{\rho}_{kj})$ using the moments of the bivariate normal. To obtain the asymptotic variances of $\tilde{\rho}_{kj}$ and $\check{\rho}_{kj}$, consider the following vector of influence functions

$$\mathbf{m}(y_k, y_j; \rho_{kj}) = [\mathbf{m}'_{\mu_{kj}}(y_k, y_j; \rho_{kj}), \mathbf{m}'_{\Sigma_{kj}}(y_k, y_j; \rho_{kj})]'$$

with $\mathbf{m}_{\mu_{kj}}(y_k, y_j; \rho_{kj}) = (y_k, y_j)'$ and $\mathbf{m}_{\Sigma_{kj}}(y_k, y_j; \rho_{kj}) = (y_k^2 - \sigma_k^2, y_j^2 - \sigma_j^2, y_k y_j - \rho_{kj} \sigma_k \sigma_j)'$. Since $E [\mathbf{m}_{\mu_{kj}}(y_k, y_j; \rho_{kj}) \mathbf{m}'_{\Sigma_{kj}}(y_k, y_j; \rho_{kj})] = \mathbf{0}$, the relevant quantities are

$$A_{kj} = E [\mathbf{m}_{\Sigma_{kj}}(y_k, y_j; \rho_{kj}) \mathbf{m}'_{\Sigma_{kj}}(y_k, y_j; \rho_{kj})] \quad \text{and} \quad B_{kj} = E [\mathbf{s}_{\Sigma_{kj}}(y_k, y_j; \rho_{kj}) \mathbf{m}'_{\Sigma_{kj}}(y_k, y_j; \rho_{kj})],$$

where

$$\mathbf{s}_{\Sigma_{kj}}(y_k, y_j; \rho_{kj}) = [s_{\sigma_k^2}(y_k, y_j; \rho_{kj}), s_{\sigma_j^2}(y_k, y_j; \rho_{kj}), s_{\rho_{kj}}(y_k, y_j; \rho_{kj})]'$$

This expression can be easily evaluated at $\sigma_k^2 = \sigma_j^2 = 1$ to obtain $AVar(\tilde{\rho})$ as the (3, 3) element of $(B_{kj} A_{kj}^{-1} B'_{kj})^{-1}$. \square

Proposition 3

From (1) we can easily compute the first two terms of (6), which taking limits when $\eta \rightarrow 0$ reduce to

$$\lim_{\eta \rightarrow 0} \frac{\partial \ln f_K(y_1, \dots, y_K; \rho, \eta)}{\partial \eta} = \sqrt{\frac{K(K+2)}{2}} \times L_2(\varsigma)$$

and

$$\lim_{\eta \rightarrow 0} \sum_{k=1}^K \frac{\partial \ln f_1(y_k; \eta)}{\partial \eta} = \sqrt{\frac{3}{2}} \times \sum_{k=1}^K H_4(y_k),$$

where $L_j(\cdot)$ and $H_j(\cdot)$ are the normalized Laguerre and Hermite polynomials of order j , respectively. As for the remaining terms in (6), we can use the fact that for a generic copula density such as the one in (5),

$$\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial \ln f_K(y_1, \dots, y_K; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial \varsigma} = -\frac{1}{2} \quad \text{and} \quad \lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial \ln f_1(y_k; \boldsymbol{\varphi})}{\partial y_k} = -y_k,$$

so that

$$\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \left[\frac{\partial \ln f_K(\varsigma; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial \varsigma} \frac{\partial \varsigma}{\partial y_k} - \frac{\partial \ln f_1(y_k; \boldsymbol{\varphi})}{\partial y_k} \right] = \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) [\mathbf{y} - \mathbf{p}_{(k)}(\boldsymbol{\rho}) y_k]}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})}, \quad (\text{A1})$$

where $\mathbf{y} - \mathbf{p}_{(k)} y_k$ are residuals of univariate simple regressions of \mathbf{y}_{-k} onto y_k because $\mathbf{p}_{(k)}$ are the corresponding OLS coefficients and the denominator is the residual variance in a regression of y_k onto the remaining components of \mathbf{y} . As for $\partial F_1^{-1}(u_k; \eta) / \partial \eta$, differentiating $\int_{-\infty}^{F_1^{-1}(u_k; \eta)} f_1(y_k; \eta) dy_k = u_k$ with respect to η yields

$$\frac{\partial F_1^{-1}(u_k; \eta)}{\partial \eta} = \frac{-1}{f_1[F_1^{-1}(u_k; \eta); \eta]} \int_{-\infty}^{F_1^{-1}(u_k; \eta)} \frac{\partial f_1(y_k; \eta)}{\partial \eta} dy_k.$$

But then, noticing that $\lim_{\eta \rightarrow 0} f_1[F_1^{-1}(u_k; \eta); \eta] = \phi[\Phi^{-1}(u_k)]$ and that

$$\lim_{\eta \rightarrow 0} \frac{\partial f_1(y_k; \eta)}{\partial \eta} = \phi(y_k) \times \sqrt{\frac{3}{2}} \times H_4(y_k)$$

we obtain

$$\lim_{\eta \rightarrow 0} \frac{\partial F_1^{-1}(u_k; \eta)}{\partial \eta} = \frac{1}{2} \sqrt{\frac{3}{2}} \times H_3(y_k).$$

Collecting terms finally yields (9). □

Proposition 4

For fixed \mathbf{b} , the LM test is based on the average score with respect to η evaluated at the limit of $\eta \rightarrow 0$. In this regard, we first obtain the parameters of the corresponding marginal distributions appearing in (5). Specifically, if $\mathbf{y} \sim At(\mathbf{0}, \mathbf{P}(\boldsymbol{\rho}), \eta, \mathbf{b})$ with $\boldsymbol{\beta} = \mathbf{P}^{1/2'}(\boldsymbol{\rho}) \mathbf{b}$, then $y_k \sim At(0, 1, \eta, \mathfrak{b}_k(\eta, \mathbf{b}))$ where

$$\mathfrak{b}_k(\eta, \mathbf{b}) = \frac{c(\mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}, \eta) \iota'_k \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}}{1 + [c(\mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}, \eta) - 1] \iota'_k \mathbf{P}(\boldsymbol{\rho}) \mathbf{b} \mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \iota_k / \mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}},$$

with ι_k denoting a $K \times 1$ vector with 1 in its k 'th position and 0's otherwise, and

$$c(\mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}, \eta) = \frac{-(1 - 4\eta) + \sqrt{(1 - 4\eta)^2 + 8\eta(1 - 4\eta) \mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}}}{4\eta \mathbf{b}' \mathbf{P}(\boldsymbol{\rho}) \mathbf{b}}.$$

In this context, we can write $s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, \eta, \mathbf{b})$ as

$$\begin{aligned} & \frac{\partial \ln f_K(y_1, \dots, y_K; \boldsymbol{\rho}, \eta, \mathbf{b})}{\partial \eta} - \sum_{k=1}^K \frac{\partial \ln f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \eta} - \sum_{k=1}^K \frac{\partial \ln f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \mathfrak{B}_k} \frac{\partial \mathfrak{B}_k(\eta, \mathbf{b})}{\partial \eta} \\ & + \sum_{k=1}^K \left(\frac{\partial \ln f_K(y_1, \dots, y_K; \boldsymbol{\rho}, \eta, \mathbf{b})}{\partial y_k} - \frac{\partial \ln f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial y_k} \right) \frac{\partial F_1^{-1}[u_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \eta}. \end{aligned} \quad (\text{A2})$$

As for the first two terms of (A2), Mencía and Sentana (2012) provide the corresponding expressions, which reduce to

$$\lim_{\eta \rightarrow 0} \frac{\partial \ln f_K(y_1, \dots, y_K; \boldsymbol{\rho}, \eta, \mathbf{b})}{\partial \eta} = \sqrt{\frac{K(K+2)}{2}} \times L_2(\varsigma) + \mathbf{b}'\mathbf{y}[\varsigma - (K+2)]$$

and

$$\lim_{\eta \rightarrow 0} \frac{\partial \ln f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \eta} = \sqrt{\frac{3}{2}} H_4(y_k) + \mathbf{P}_{[k]}(\boldsymbol{\rho}) \mathbf{b} H_3(y_k),$$

where $\mathbf{P}_{[k]}(\boldsymbol{\rho})$ denotes the k 'th row of $\mathbf{P}(\boldsymbol{\rho})$. Regarding the third term of (A2), they also show that

$$\lim_{\eta \rightarrow 0} \frac{\partial \ln f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \mathfrak{B}_k} = 0.$$

As for $\partial F_1^{-1}(u_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b}))/\partial \eta$, differentiating $\int_{-\infty}^{F_1^{-1}(u_k, \eta, \mathfrak{B}_k(\eta, \mathbf{b}))} f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})] dy_k = u_k$ with respect to η yields

$$\frac{\partial F_1^{-1}(u_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b}))}{\partial \eta} = \frac{-1}{f_1[F_1^{-1}(u_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})); \eta]} \int_{-\infty}^{F_1^{-1}(u_k, \eta, \mathfrak{B}_k(\eta, \mathbf{b}))} \frac{\partial f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \eta} dy_k.$$

Then, noticing that $\lim_{\eta \rightarrow 0} f_1[F_1^{-1}(u_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})); \eta] = \phi[\Phi^{-1}(u_k)]$ and that

$$\lim_{\eta \rightarrow 0} \frac{\partial f_1[y_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b})]}{\partial \eta} = \phi(y_k) \left[\sqrt{\frac{3}{2}} H_4(y_k) + \mathfrak{B}_k(\eta, \mathbf{b}) \sqrt{6} H_3(y_k) \right]$$

we obtain

$$\lim_{\eta \rightarrow 0} \frac{\partial F_1^{-1}(u_k; \eta, \mathfrak{B}_k(\eta, \mathbf{b}))}{\partial \eta} = \mathfrak{B}_k(\eta, \mathbf{b}) \sqrt{2} H_2(y_k) + \frac{1}{2} \sqrt{\frac{3}{2}} H_3(y_k).$$

Collecting terms and using (A1) yields (10). \square

Proposition 5

Under normality, the score with respect to \mathbf{b} is 0, while the score with respect to η for fixed values of \mathbf{b} is given in Proposition 4. Now consider a reparameterization in terms of η^\dagger and \mathbf{b}^\dagger , where $\eta^\dagger = \eta$ and $\mathbf{b}^\dagger = \mathbf{b}\eta$. This reparameterization is such that under normality, both η^\dagger and \mathbf{b}^\dagger will be zero, while under local alternatives of the form $\eta_T^\dagger = T^{-1/2} \bar{\eta}^\dagger$ and $\mathbf{b}_T^\dagger = T^{-1/2} \bar{\mathbf{b}}^\dagger$, we

will have an asymmetric Student t distribution with parameters $\eta_T = T^{-1/2}\bar{\eta}$ and $\mathbf{b}_T = \bar{\mathbf{b}}$. As for the score test, we start by defining

$$c^\dagger(\mathbf{u}; \boldsymbol{\rho}, \eta^\dagger, \mathbf{b}^\dagger) = c\left(\mathbf{u}; \boldsymbol{\rho}, \eta, \frac{\mathbf{b}^\dagger}{\eta^\dagger}\right).$$

We can then expand $\ln c(\mathbf{u}; \boldsymbol{\rho}, \eta, \mathbf{b})$ around $\eta = 0$ as follows

$$\ln c(\mathbf{u}; \boldsymbol{\rho}, \eta, \mathbf{b}) = \ln c(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{b}) + s_\eta(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{b})\eta + O(\eta^2),$$

and similarly, we can also expand $\ln c^\dagger(\mathbf{u}; \boldsymbol{\rho}, \eta^\dagger, \mathbf{b}^\dagger)$ as

$$\begin{aligned} \ln c^\dagger(\mathbf{u}; \boldsymbol{\rho}, \eta^\dagger, \mathbf{b}^\dagger) &= \ln c^\dagger(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{0}) + s_\eta^\dagger(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{0})\eta \\ &\quad + \mathbf{s}_{\mathbf{b}^\dagger}^\dagger(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{0})\mathbf{b}^\dagger + O(\eta^2) + O(\mathbf{b}^\dagger'\mathbf{b}^\dagger) + O(\mathbf{b}^\dagger\eta^\dagger). \end{aligned}$$

Since $\ln c(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{b})$ does not depend on \mathbf{b} and

$$s_\eta(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{b})\eta = J_0(\mathbf{u}; \boldsymbol{\rho})\eta + \sum_{k=1}^K J_k(\mathbf{u}; \boldsymbol{\rho})b_k\eta$$

in light of Proposition 4, we can identify $J_0(\mathbf{u}; \boldsymbol{\rho})$ with $s_\eta^\dagger(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{0})$ and, for $k = 1, \dots, K$, $J_k(\mathbf{u}; \boldsymbol{\rho})$ with $\mathbf{s}_{\mathbf{b}^\dagger}^\dagger(\mathbf{u}; \boldsymbol{\rho}, 0, \mathbf{0})$ as $\mathbf{b}^\dagger = \mathbf{b}\eta$. \square

Proposition 6

As in Chen and Fan (2006a), to obtain the correction for non-parametric estimation of the marginals for the element $s_{\rho_{kj}}(u_k, u_j; \rho_{kj})$, we need to compute

$$n_{\rho_{kj}} = \int_0^1 [1\{U_k \leq u_k\} - u_k] W_{\rho_{kj}}^k du_k + \int_0^1 [1\{U_j \leq u_j\} - u_j] W_{\rho_{kj}}^j du_j,$$

with $W_{\rho_{kj}}^j = \int [\partial s_{\rho_{kj}}(u_k, u_j; \rho_{kj}) / \partial u_j] c(u_k, u_j; \rho_{kj}) du_k$. Then, the result follows from

$$W_{\rho_{kj}}^j = \int \left[\frac{1 + \rho_{kj}^2}{(1 - \rho_{kj}^2)^2} y_k - \frac{2\rho_{kj}}{(1 - \rho_{kj}^2)^2} y_j \right] \phi(y_j) du_k = \frac{1 + \rho_{kj}^2}{(1 - \rho_{kj}^2)^2} y_j$$

and the fact that

$$\int_{-\infty}^y H_1(x) \Phi(x) dx = \frac{H_1(y)}{2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) + \frac{1}{2\sqrt{2}} H_2(y) \left[1 + \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right) \right]$$

and

$$\int_y^\infty H_1(x) [1 - \Phi(x)] dx = \frac{H_1(y)}{2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) - \frac{1}{2\sqrt{2}} H_2(y) \left[\operatorname{erfc}\left(\frac{y}{\sqrt{2}}\right) \right],$$

where $\operatorname{erf}(z) = 2\pi^{-1/2} \int_0^z e^{-t^2} dt$ and $\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$. \square

Proposition 7

Analogous calculations to the ones provided in the proof of Proposition 6 allow us to obtain

$$n_{\rho_{kj}}(y_k, y_j; \rho_{kj}) = -\frac{\rho_{kj}}{2} [H_2(y_k) + H_2(y_j)],$$

$$n_{\sigma_k^2}(y_k, y_j; \rho_{kj}) = \sqrt{2}H_2(y_k) \quad \text{and} \quad n_{\sigma_j^2}(y_k, y_j; \rho_{kj}) = \sqrt{2}H_2(y_j).$$

In this context, the asymptotic variances of the estimators of ρ_{kj} can be obtained as

$$AVar(\hat{\rho}_{kj}^{np}) = \frac{cov[s_{\rho_{kj}}^{np}(y_k, y_j; \rho_{kj}), s_{\rho_{kj}}(y_k, y_j; \rho_{kj})]}{V[s_{\rho_{kj}}^{np}(y_k, y_j; \rho_{kj})]} = (1 - \rho_{kj}^2)^2,$$

which, interestingly, coincides with the common asymptotic variance of $\tilde{\rho}_{kj}^{np}$ and $\hat{\rho}_{kj}^{np}$ given by

$$AVar(\tilde{\rho}_{kj}^{np}) = AVar(\hat{\rho}_{kj}^{np}) = (1 - \rho_{kj}^2)^2$$

since $E[\mathbf{m}_{\Sigma_{kj}}^{np}(y_k, y_j; \rho_{kj})\mathbf{m}_{\Sigma_{kj}}^{np'}(y_k, y_j; \rho_{kj})]$ and $E[\mathbf{s}_{\Sigma_{kj}}(y_k, y_j; \rho_{kj})\mathbf{m}_{\Sigma_{kj}}^{np'}(y_k, y_j; \rho_{kj})]$ have all the elements equal to zero except the $(3, 3)$ ones, which are equal to $(1 - \rho_{kj}^2)^2$ and 1, respectively. \square

Proposition 8

Following Chen and Fan (2006a), to obtain the correction for non-parametric estimation of the marginals for a generic score $s_\phi(u_1, \dots, u_K)$, we need to compute

$$n_\phi = \sum_{j=1}^K \int_0^1 [1\{U_j \leq u_j\} - u_j] W_{\phi_k}^j du_j$$

with

$$W_\phi^j = \int \dots \int \frac{\partial s_\phi(u_1, \dots, u_K)}{\partial u_j} c(u_1, \dots, u_K; \phi) du_1 \dots du_{k-1} du_{k+1} \dots du_K.$$

To do so, we can exploit the fact that $\mathbf{z}_k = \mathbf{y}_{(k)} - \mathbf{p}_{(k)}(\boldsymbol{\rho}) y_k \sim N(\mathbf{0}, \boldsymbol{\Upsilon}_k)$ with $\boldsymbol{\Upsilon}_k = \mathbf{P}_{(kk)}(\boldsymbol{\rho}) - \mathbf{p}_{(k)}(\boldsymbol{\rho})\mathbf{p}_{(k)}'(\boldsymbol{\rho})$, and that

$$\varsigma(\boldsymbol{\rho}) = y_k^2 + \mathbf{z}_k' \left[\mathbf{P}_{(kk)}(\boldsymbol{\rho}) - \mathbf{p}_{(k)}(\boldsymbol{\rho})\mathbf{p}_{(k)}'(\boldsymbol{\rho}) \right]^{-1} \mathbf{z}_k, \quad (\text{A3})$$

$$y_k - \mathbf{p}_{(k)}'(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{y}_{(k)} = \left[1 - \mathbf{p}_{(k)}'(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{p}_{(k)}(\boldsymbol{\rho}) \right] y_k + \mathbf{p}_{(k)}'(\boldsymbol{\rho})\mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho})\mathbf{z}_k. \quad (\text{A4})$$

We can also make use of the fact that

$$c(\mathbf{u}_{(k)}; \phi) d\mathbf{u}_{(k)} = c(u_1, \dots, u_K; \phi) du_1 \dots du_{k-1} du_{k+1} \dots du_K$$

involves integrating with respect to

$$f(\mathbf{z}_k; \phi) = \frac{(2\pi)^{-(K-1)/2}}{|\boldsymbol{\Upsilon}_k|^{1/2}} \times \exp\left(-\frac{1}{2}\mathbf{z}_k' \boldsymbol{\Upsilon}_k^{-1} \mathbf{z}_k\right).$$

Specifically, for the first term of (9), using the fact that

$$\frac{\partial}{\partial \varsigma} \left[\sqrt{\frac{K(K+2)}{2}} L_2(\varsigma) \right] = \frac{\varsigma - (K+2)}{2} \quad \text{and} \quad \frac{\partial \varsigma(\boldsymbol{\rho})}{\partial y_k} = 2 \times \frac{y_k - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{y}_{(k)}}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})},$$

and substituting (A3) and (A4), we obtain

$$\int \frac{\partial}{\partial y_k} \left[\sqrt{\frac{K(K+2)}{2}} L_2(\varsigma) \right] c(\mathbf{u}_{(k)}; \boldsymbol{\phi}) d\mathbf{u}_{(k)} = \sqrt{6} H_3(y_k) \quad (\text{A5})$$

where the last equality follows from

$$\int \mathbf{z}_k \mathbf{z}'_k \boldsymbol{\Upsilon}_k^{-1} \mathbf{z}_k f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \mathbf{0} \quad \text{and} \quad \int \mathbf{z}'_k \boldsymbol{\Upsilon}_k^{-1} \mathbf{z}_k f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = K - 1.$$

Similarly, for the second term of (9), using $\partial H_j(y)/\partial y = \sqrt{j} H_{j-1}(y)$ we notice that

$$\int \frac{\partial}{\partial y_k} \left[\sqrt{\frac{3}{2}} \sum_{h=1}^K H_4(y_h) \right] f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \sqrt{6} H_3(y_k),$$

which cancels with (A5). Regarding the final term of (9), given that $\int \mathbf{z}_k f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \mathbf{0}$ and that

$$\int H_3(y_j) f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \int H_3(y_j) \frac{1}{\sqrt{1 - \rho_{jk}^2}} \phi \left(\frac{y_j - \rho_{jk} y_k}{\sqrt{1 - \rho_{jk}^2}} \right) f(\mathbf{z}_k; \boldsymbol{\phi}) dy_j = \rho_{jk}^3 H_3(y_k),$$

we can show that

$$\int \frac{\partial}{\partial y_k} \left\{ \sqrt{\frac{3}{8}} \sum_{h=1}^K \left[\frac{\mathbf{p}'_{(h)}(\boldsymbol{\rho}) \mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho}) \mathbf{z}_h}{1 - \mathbf{p}'_{(h)}(\boldsymbol{\rho}) \mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(h)}(\boldsymbol{\rho})} \right] H_3(y_h) \right\} f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k$$

can be written as

$$-\sqrt{\frac{3}{8}} \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} H_3(y_k) + \sqrt{\frac{3}{8}} \sum_{h \neq k} \frac{\mathbf{p}'_{(h)}(\boldsymbol{\rho}) \mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(h)}(\boldsymbol{\rho}) \mathbf{P}_{(hh)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(h)}(\boldsymbol{\rho})} \mathbf{P}_{kj}^3(\boldsymbol{\rho}) H_3(y_k).$$

As for $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, \eta)$, we can use (A3) to rewrite its first term as

$$\frac{\partial \{y_k [\varsigma(\boldsymbol{\rho}) - (K+2)]\}}{\partial y_k} = 3y_k^2 + \mathbf{z}'_k \boldsymbol{\Upsilon}_k^{-1} \mathbf{z}_k - (K+2) - 2 \times \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{z}_k}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})},$$

so that

$$\int \frac{\partial \{y_k [\varsigma(\boldsymbol{\rho}) - (K+2)]\}}{\partial y_k} f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = 3(y_k^2 - 1) = 3\sqrt{2} H_2(y_k),$$

where we have used the fact that $\int \mathbf{z}'_k \boldsymbol{\Upsilon}_k^{-1} \mathbf{z}_k f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = K - 1$ and $\int \mathbf{z}_k f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \mathbf{0}$,

which again cancels with the correction corresponding to the second term since

$$\int \frac{\partial}{\partial y_k} \left[\sqrt{6} \sum_{j=1}^K \mathbf{P}_{kj}(\boldsymbol{\rho}) H_3(y_j) \right] f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = 3\sqrt{2} H_2(y_k)$$

because $\mathbf{P}_{kk}(\boldsymbol{\rho}) = 1$ and $\int f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = 1$. Finally, as for

$$\frac{\partial}{\partial y_k} \left\{ \sqrt{2} \sum_{j=1}^K \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) [\mathbf{y}_{(k)} - \mathbf{p}_{(k)}(\boldsymbol{\rho}) y_k]}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} \mathbf{P}_{kj}(\boldsymbol{\rho}) H_2(y_j) \right\},$$

we have to deal with the following two terms:

$$-\sqrt{2} \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} H_2(y_k) + \sqrt{2} \sum_{j \neq k} \frac{\mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(j)}(\boldsymbol{\rho})} \mathbf{P}_{kj}(\boldsymbol{\rho}) H_2(y_j)$$

and

$$2 \frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{z}_k}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} \mathbf{P}_{kk}(\boldsymbol{\rho}) H_1(y_k).$$

The integral of the last term is zero since $\int \mathbf{z}_k f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \mathbf{0}$. As for the first one, noticing that

$$\int H_2(y_j) f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k = \int H_2(y_j) \frac{1}{\sqrt{1 - \rho_{jk}^2}} \phi \left(\frac{y_j - \rho_{jk} y_k}{\sqrt{1 - \rho_{jk}^2}} \right) f(\mathbf{z}_k; \boldsymbol{\phi}) dy_j = \rho_{jk}^2 H_2(y_k)$$

we obtain that

$$-\int \sqrt{2} \sum_{j=1}^K \frac{\mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(j)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(j)}(\boldsymbol{\rho})} \mathbf{P}_{kj}(\boldsymbol{\rho}) H_2(y_j) f(\mathbf{z}_k; \boldsymbol{\phi}) d\mathbf{z}_k$$

is equal to

$$-\sqrt{2} \left[\frac{\mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(k)}(\boldsymbol{\rho}) \mathbf{P}_{(kk)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(k)}(\boldsymbol{\rho})} + \sum_{j \neq k} \frac{\mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho})}{1 - \mathbf{p}'_{(j)}(\boldsymbol{\rho}) \mathbf{P}_{(jj)}^{-1}(\boldsymbol{\rho}) \mathbf{p}_{(j)}(\boldsymbol{\rho})} \mathbf{P}_{kj}^2(\boldsymbol{\rho}) \right] H_2(y_k).$$

Analogous calculations allow us to obtain the relevant quantities for $\partial m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, \boldsymbol{\eta}) / \partial y_j$.

Finally, the results stated in the proposition are obtained by collecting terms and integrating y_k out using the fact that

$$\begin{aligned} \int_{-\infty}^y H_3(x) \Phi(x) dx &= \frac{H_3(y)}{4} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) + \frac{1}{4} H_4(y) \left[1 + \operatorname{erfc}\left(\frac{y}{\sqrt{2}}\right)\right], \\ \int_y^{\infty} H_3(x) [1 - \Phi(x)] dx &= \frac{H_3(y)}{4} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) - \frac{1}{4} H_4(y) \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right) \end{aligned}$$

so that

$$\int_{-\infty}^y H_3(x) \Phi(x) dx - \int_y^{\infty} H_3(x) [1 - \Phi(x)] dx = \frac{1}{2} H_4(y),$$

and

$$\begin{aligned} \int_{-\infty}^y H_2(x) \Phi(x) dx &= \frac{H_2(y)}{3} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) + \frac{1}{2\sqrt{3}} H_3(y) \left[1 + \operatorname{erfc}\left(\frac{y}{\sqrt{2}}\right)\right], \\ \int_y^{\infty} H_2(x) [1 - \Phi(x)] dx &= \frac{H_2(y)}{3} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) - \frac{1}{2\sqrt{3}} H_3(y) \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right) \end{aligned}$$

so that

$$\int_{-\infty}^y H_2(x) \Phi(x) dx - \int_y^{\infty} H_2(x) [1 - \Phi(x)] dx = \frac{1}{\sqrt{3}} H_3(y).$$

□

Proposition 9

If $\mathbf{P}(\boldsymbol{\rho})$ were unrestricted i.e. when $\boldsymbol{\rho}$ contains the off-diagonal elements of $\mathbf{P}(\boldsymbol{\rho})$, $p = K(K-1)/2$ and $\hat{\rho}_{kj} \xrightarrow{p} \rho_{kj\infty}$ where

$$E \left[s_{\rho_{kj}}(y_k, y_j; \rho_{kj\infty}) \middle| c_0(\cdot, \phi_0) \right] = E \left[\frac{(1 + \rho_{kj\infty}^2) y_k y_j - \rho_{kj\infty} (y_k^2 + y_j^2 - 1) - \rho_{kj\infty}^3}{(1 - \rho_{kj\infty}^2)^2} \middle| c_0(\cdot, \phi_0) \right] = 0.$$

But since $E(y_k^2) = E(y_j^2) = 1$, the previous expression reduces to

$$E \left[(1 + \rho_{kj\infty}^2) y_k y_j - \rho_{kj\infty} (1 + \rho_{kj\infty}^2) \middle| c_0(\cdot, \phi_0) \right] = (1 + \rho_{kj\infty}^2) E \left[y_k y_j - \rho_{kj\infty} \middle| c_0(\cdot, \phi_0) \right] = 0$$

so that $\rho_{kj\infty} = E[y_k y_j | c_0(\cdot, \phi_0)]$. More generally, we could consider $\mathbf{P}[\boldsymbol{\rho}(\boldsymbol{\theta})]$, where $\boldsymbol{\rho}(\boldsymbol{\theta})$ is a $K(K-1)/2 \times 1$ vector and $\boldsymbol{\theta}$ a $p \times 1$ vector with $p < K(K-1)/2$ so that

$$0 = E[s_{\boldsymbol{\theta}}(\mathbf{y}; \boldsymbol{\theta}) | c_0(\cdot, \phi_0)] = \frac{\partial \boldsymbol{\rho}'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} E[s_{\boldsymbol{\rho}}(\mathbf{y}; \boldsymbol{\rho}) | c_0(\cdot, \phi_0)]$$

as desired. \square

B Hermite expansion-based copula tests

Proposition 10 *The LM test for a Gaussian copula against a Hermite polynomial expansion copula with $\rho \neq 0$ coincides with a moment test based on the following influence functions:*

$$\mathbf{m}_{\boldsymbol{\varphi}}(u_1, u_2; \boldsymbol{\rho}) = \begin{pmatrix} s_{\gamma_1}(y_1, y_2; \rho) \\ s_{c_2}(y_1, y_2; \rho) \\ s_{c_3}(y_1, y_2; \rho) \\ s_{\gamma_4}(y_1, y_2; \rho) \\ s_{d_1}(y_1, y_2; \rho) \\ s_{d_2}(y_1, y_2; \rho) \\ s_{d_3}(y_1, y_2; \rho) \\ s_{d_4}(y_2, y_1; \rho) \\ s_{\delta_5}(y_1, y_2; \rho) \end{pmatrix} = \begin{pmatrix} s_{\gamma_1}(y_1, y_2; \rho) \\ s_{c_2}(y_1, y_2; \rho) \\ s_{c_2}(y_2, y_1; \rho) \\ s_{\gamma_1}(y_2, y_1; \rho) \\ s_{d_1}(y_1, y_2; \rho) \\ s_{d_2}(y_1, y_2; \rho) \\ s_{d_3}(y_1, y_2; \rho) \\ s_{d_2}(y_2, y_1; \rho) \\ s_{\delta_5}(y_1, y_2; \rho) \end{pmatrix}$$

where

$$\begin{aligned} s_{\gamma_1}(y_1, y_2; \rho) &= \rho^6 H_3^2(y_2) + \frac{2}{\sqrt{3}} \rho H_3(y_1) H_2(y_1) [\rho H_1(y_1) - H_1(y_2)] - \frac{2}{\sqrt{3}} \rho H_3(y_2) H_2(y_2) H_1(y_2) \\ &\quad + \frac{1}{3} \rho H_2^2(y_1) \{ \rho [1 - H_1^2(y_2)] + \rho^3 [H_1^2(y_1) - 1] \} + \frac{1}{3} \rho^6 H_2^2(y_2) H_1^2(y_2) \\ &\quad + \frac{2\sqrt{2}}{3} \rho H_2(y_1) H_1(y_1) (1 - \rho^2) [\rho H_1(y_1) - H_1(y_2)] \\ &\quad + \frac{2}{3} \rho^3 [2H_1(y_1) H_1(y_2) - 2H_1^2(y_1) - \rho^3] \end{aligned}$$

$$\begin{aligned} s_{c_2}(y_1, y_2; \rho) &= H_2(y_1) H_1(y_2) - H_2(y_1) H_1(y_1) \rho - H_1(y_2) H_2(y_2) \rho^2 \\ &\quad + \frac{\rho(y_2 - \rho y_1)}{1 - \rho^2} \rho [H_2(y_1) + \sqrt{2}] + \frac{\rho(y_1 - \rho y_2)}{1 - \rho^2} \rho^2 [H_2(y_2) + \sqrt{2}] \end{aligned}$$

$$s_{d_1}(y_1, y_2, \rho) = -H_4(y_2)\rho^4 + \frac{\rho(y_2 - \rho y_1)}{1 - \rho^2} \frac{1}{2} H_3(y_1) + \frac{\rho(y_1 - \rho y_2)}{1 - \rho^2} \frac{\rho^4}{2} H_3(y_2)$$

$$\begin{aligned} s_{d_2}(y_1, y_2, \rho) &= H_3(y_1)H_1(y_2) - H_3(y_1)H_1(y_1)\rho - H_3(y_2)H_1(y_2)\rho^3 \\ &+ \frac{\rho(y_2 - \rho y_1)}{1 - \rho^2} \rho \left[H_3(y_1) + \sqrt{\frac{3}{2}} H_1(y_1) \right] + \frac{\rho(y_1 - \rho y_2)}{1 - \rho^2} \rho^3 \left[H_3(y_2) + \sqrt{\frac{3}{2}} H_1(y_2) \right] \end{aligned}$$

$$\begin{aligned} s_{d_3}(y_1, y_2, \rho) &= H_2(y_1)H_2(y_2) - \rho^2 [H_2^2(y_1) + H_2^2(y_2) - 1] \\ &+ \frac{\rho(y_2 - \rho y_1)}{1 - \rho^2} \rho^2 \left[\sqrt{\frac{3}{2}} H_3(y_1) + 2H_1(y_1) \right] + \frac{\rho(y_1 - \rho y_2)}{1 - \rho^2} \rho^2 \left[\sqrt{\frac{3}{2}} H_3(y_2) + 2H_1(y_2) \right] \end{aligned}$$

and

$$\begin{aligned} s_{\delta_5}(y_1, y_2, \rho) &= 2H_4(y_1)H_4(y_2) - (2 - \rho^4)\rho^4 [H_4^2(y_1) + H_4^2(y_2)] \\ &+ \rho H_4(y_1) \left[H_3(y_1)H_1(y_2) - H_3(y_2)H_1(y_1) - \rho H_3(y_1)H_1(y_1)(1 + \rho^2)(1 - \rho^2)^2 \right] \\ &+ \rho H_4(y_2) \left[H_3(y_2)H_1(y_1) - H_3(y_1)H_1(y_3) - \rho H_3(y_2)H_1(y_2)(1 + \rho^2)(1 - \rho^2)^2 \right] \\ &+ \frac{1}{4}\rho^2 H_3^2(y_1) \left\{ 1 - H_1^2(y_2) + \rho^2 [5 - 6\rho^2 + H_1^2(y_1)(1 - \rho^2)^2] \right\} \\ &+ \frac{1}{4}\rho^2 H_3^2(y_2) \left\{ 1 - H_1^2(y_1) + \rho^2 [5 - 6\rho^2 + H_1^2(y_2)(1 - \rho^2)^2] \right\} \\ &- \frac{\sqrt{3}}{4}\rho(1 - \rho^2)H_3(y_1) \left\{ 2H_2(y_1)H_1(y_2) + \rho H_1(y_1) [3\sqrt{2}\rho^4 - 2H_2(y_1)(1 - \rho^2)] \right\} \\ &- \frac{\sqrt{3}}{4}\rho(1 - \rho^2)H_3(y_2) \left\{ 2H_2(y_2)H_1(y_1) + \rho H_1(y_2) [3\sqrt{2}\rho^4 - 2H_2(y_2)(1 - \rho^2)] \right\} \\ &+ \frac{3}{2\sqrt{2}}\rho^3(1 - \rho^2)H_2(y_1)H_1(y_1) [H_1(y_2) + H_1(y_1)\rho(1 - \rho^2)] \\ &+ \frac{3}{2\sqrt{2}}\rho^3(1 - \rho^2)H_2(y_2)H_1(y_2) [H_1(y_1) + H_1(y_2)\rho(1 - \rho^2)] \\ &+ \frac{3}{8}\rho^5(1 - \rho^2) \{ 2H_1(y_1)H_1(y_2) + \rho [H_1^2(y_1) + H_1^2(y_2)] \} \\ &+ \frac{1}{2}\rho [1 - \rho^2 + \rho H_1(y_1)H_1(y_2)] H_1(y_3)H_3(y_2) \end{aligned}$$

Proof. Under the alternative hypothesis, (y_1, y_2) follow a fourth order Hermite expansion of the Gaussian distribution so that their joint density is

$$f_2(y_1, y_2; \rho, \boldsymbol{\varphi}) = \phi_2 \left(\frac{y_1^2 + y_2^2 - 2\rho y_1 y_2}{1 - \rho^2} \right) P_2(y_1, y_2; \rho, \boldsymbol{\varphi}),$$

where $\boldsymbol{\varphi} = (c_1, c_2, c_3, c_4, d_1, d_2, d_3, d_4, d_5)'$ and $P_2(y_1, y_2; \rho, \boldsymbol{\varphi})$ is given in (4), so that the marginals become

$$f_i(y_i; \rho, \boldsymbol{\varphi}) = \frac{1}{1 + d_3\rho^2} \phi_1(y_i) P_{y_i}(y_i; \rho, \boldsymbol{\varphi}),$$

where

$$\begin{aligned} P_{y_j}(y_j; \rho, \boldsymbol{\varphi}_j) &= 1 + (\varphi_1^j + \varphi_4^j \rho^3) H_3(y_j) + (\varphi_2^j \rho + \varphi_3^j \rho^2) H_1(y_j) H_2(y_j) \\ &+ (\varphi_5^j + \varphi_9^j \rho^4) H_4(y_j) + (\varphi_6^j \rho + \varphi_8^j \rho^3) H_1(y_j) H_3(y_j) + \varphi_7^j \rho^2 H_2^2(y_j), \end{aligned}$$

with $\boldsymbol{\varphi}_j = (c_1, c_2, c_3, c_4, d_1, d_2, d_3, d_4, d_5)'$ and $\boldsymbol{\varphi}_2 = (c_4, c_3, c_2, c_1, d_5, d_4, d_3, d_2, d_1)'$.

We can easily obtain the relevant quantities for the joint score by noticing that

$$\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial \ln f_2(y_1, y_2; \rho, \boldsymbol{\varphi})}{\partial \varphi_i}$$

coincides with the Hermite Polynomial associated to the coefficient φ_i , with the exception of $\varphi_7 = d_3$, in which case $\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \partial \ln f_2(y_1, y_2; \rho, \boldsymbol{\varphi}) / \partial d_3 = H_2(y_1)H_2(y_2) - \rho^2$. Analogously, we can obtain the corresponding expressions for the score of the marginal densities, which yield $\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \partial \ln f_1(y; \boldsymbol{\varphi}) / \partial d_3 = [H_2^2(y) - 1] \rho^2$, while for the remaining elements of $\boldsymbol{\varphi}$ coincide with the Hermite Polynomial associated to the coefficient φ_i of $P_{y_i}(y_i; \rho, \boldsymbol{\varphi})$.

The remaining quantities involved in the expressions for the score are the quantile derivatives under Gaussianity. Specifically, if we define

$$F_{1(j,i)}^{-1}(u; \rho, \mathbf{0}) = \lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial F_1^{-1}(u; \rho, \boldsymbol{\varphi}_j)}{\partial \varphi_i} = \frac{-1}{\phi[\Phi^{-1}(u)]} \int_{-\infty}^{\Phi^{-1}(u)} \phi(z) \frac{\partial P_j(z; \rho, \boldsymbol{\varphi})}{\partial \varphi_i} dz$$

we will have that

$$\begin{aligned} F_{1(1,1)}^{-1}(u; \rho, \mathbf{0}) &= \frac{1}{\rho^3} F_{1(1,4)}^{-1}(u; \rho, \mathbf{0}) = \frac{1}{\sqrt{3}} H_2(y_1), \\ F_{1(1,2)}^{-1}(u; \rho, \mathbf{0}) &= \frac{1}{\rho} F_{1(1,3)}^{-1}(u; \rho, \mathbf{0}) = \rho \left[H_2(y_1) + \sqrt{2} H_0(y_1) \right], \\ F_{1(1,5)}^{-1}(u; \rho, \mathbf{0}) &= \frac{1}{\rho^4} F_{1(1,9)}^{-1}(u; \rho, \mathbf{0}) = \frac{1}{2} H_3(y_1), \\ F_{1(1,2)}^{-1}(u; \rho, \mathbf{0}) &= \frac{1}{\rho^3} F_{1(1,3)}^{-1}(u; \rho, \mathbf{0}) = \rho \left[H_3(y_1) + \sqrt{\frac{3}{2}} H_1(y_1) \right], \end{aligned}$$

and

$$F_{1(1,7)}^{-1}(u; \rho, \mathbf{0}) = \rho^2 \left[\sqrt{\frac{3}{2}} H_3(y_1) + 2 H_1(y_1) \right].$$

Collecting terms we then obtain the score vector $\mathbf{s}_{\boldsymbol{\varphi}}(y_1, y_2; \rho)$. It is then straightforward to see that

$$\begin{aligned} s_{c_1}(y_1, y_2, \rho) - \frac{\rho}{\sqrt{3}} s_{c_2}(y_1, y_2, \rho) &= 0, \\ s_{c_4}(y_1, y_2, \rho) - \frac{\rho}{\sqrt{3}} s_{c_3}(y_1, y_2, \rho) &= 0, \end{aligned}$$

and

$$s_{d_1}(y_1, y_2, \rho) - \frac{\rho}{2} s_{d_2}(y_1, y_2, \rho) + \frac{\rho}{2} s_{d_4}(y_1, y_2, \rho) - s_{d_5}(y_1, y_2, \rho) = 0.$$

But if we reparametrize the expansion so that

$$\gamma_1 = \sqrt{c_1 - \frac{\rho}{\sqrt{3}} c_2}, \quad \gamma_4 = \sqrt{c_4 - \frac{\rho}{\sqrt{3}} c_3} \quad \text{and} \quad \delta_5 = \sqrt{d_5 - \frac{\rho}{2} d_4 + \frac{\rho}{2} d_2 - d_1},$$

we can show that the scores with respect to these three new parameters are one-half the second derivatives of the log-likelihood function with respect to γ_1^2 , γ_4^2 and δ_5^2 . Specifically,

$$\begin{aligned} s_{\gamma_1}(y_1, y_2, \rho) &= h_{c_1 c_1}(y_1, y_2, \rho) - \frac{\rho}{\sqrt{3}} h_{c_2 c_2}(y_1, y_2, \rho), \\ s_{\gamma_4}(y_1, y_2, \rho) &= h_{c_4 c_4}(y_1, y_2, \rho) - \frac{\rho}{\sqrt{3}} h_{c_3 c_3}(y_1, y_2, \rho), \end{aligned}$$

and

$$s_{\delta_5}(y_1, y_2, \rho) = h_{d_5 d_5}(y_1, y_2, \rho) - \frac{\rho}{2} h_{d_4 d_4}(y_1, y_2, \rho) + \frac{\rho}{2} h_{d_2 d_2}(y_1, y_2, \rho) - h_{d_1 d_1}(y_1, y_2, \rho),$$

where $h_{\varphi_i \varphi_i}(y_1, y_2, \rho)$ gets slightly simpler than in (7) since the last two terms in that expression disappears because $\partial^2 F_1^{-1}(u_i; \boldsymbol{\varphi}) / \partial \varphi_i^2 = 0$ for all i in the case of an Hermite expansion. Therefore

$$\begin{aligned} h_{\varphi_i \varphi_i}(y_1, y_2, \rho) &= \frac{\partial^2 \ln f_2(y_1, y_2; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial \varphi_i^2} - \frac{\partial^2 \ln f_1(y_1; \boldsymbol{\varphi})}{\partial \varphi_i^2} - \frac{\partial^2 \ln f_1(y_2; \boldsymbol{\varphi})}{\partial \varphi_i^2} \\ &+ 2 \sum_{k=1}^2 \frac{\partial F_1^{-1}(u_k; \boldsymbol{\varphi})}{\partial \varphi_i} \left(\frac{\partial^2 \ln f_2(y_1, y_2; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial y_k \partial \varphi_i} - \frac{\partial^2 \ln f_1(y_k; \boldsymbol{\varphi})}{\partial y_k \partial \varphi_i} \right) \\ &+ \sum_{k=1}^2 \sum_{j=1}^2 \frac{\partial F_1^{-1}(u_k; \boldsymbol{\varphi})}{\partial \varphi_i} \left(\frac{\partial^2 \ln f_2(y_1, y_2; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial y_k \partial y_j} - \frac{\partial^2 \ln f_1(y_k; \boldsymbol{\varphi})}{\partial y_k \partial y_j} \right) \frac{\partial F_1^{-1}(u_j; \boldsymbol{\varphi})}{\partial \varphi_i}. \end{aligned}$$

We can easily obtain the relevant quantities for the joint score by noticing that

$$\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial^2 \ln f_2(y_1, y_2; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial \varphi_i \partial \varphi_j}$$

coincides with minus the corresponding product of Hermite polynomials, except again for d_3 , where $\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \partial^2 \ln f_2(y_1, y_2; \boldsymbol{\rho}, \boldsymbol{\varphi}) / \partial d_3^2 = \rho^4 - H_2^2(y_1) H_2^2(y_2)$. Analogously, we can obtain the corresponding expressions for the marginal Hessian components, which are given by minus the corresponding product of Hermite polynomials. For instance, the relevant quantities in the case of γ_1^2 are

$$\begin{aligned} \lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial^2 \ln f_2(y_1, y_2; \boldsymbol{\rho}, \boldsymbol{\varphi})}{\partial (\gamma_1^2)^2} &= - \left[H_3(y_1) - \frac{\rho}{\sqrt{3}} H_2(y_1) H_1(y_2) \right]^2, \\ \lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial^2 \ln f_1(y_1; \boldsymbol{\varphi})}{\partial (\gamma_1^2)^2} &= - \left[H_3(y_1) - \frac{\rho^2}{\sqrt{3}} H_2(y_1) H_1(y_1) \right]^2, \end{aligned}$$

and

$$\lim_{\boldsymbol{\varphi} \rightarrow \mathbf{0}} \frac{\partial^2 \ln f_1(y_2; \boldsymbol{\varphi})}{\partial (\gamma_1^2)^2} = - \left[H_3(y_2) \rho^3 - \frac{\rho^3}{\sqrt{3}} H_2(y_2) H_1(y_2) \right]^2.$$

The expressions for $\partial^2 \ln f / \partial y_1 \partial \varphi_i$ can be easily obtained using the recursion

$$\frac{\partial H_n(y_1)}{\partial y_1} = \frac{2^{-(n-1)/2} n}{\sqrt{n!}} H_{n-1} \left(\frac{y_1}{\sqrt{2}} \right).$$

Specifically, for the terms involving derivatives with respect to \tilde{c}_1 , we obtain

$$\begin{aligned}\lim_{\varphi \rightarrow 0} \frac{\partial^2 \ln f_2(y_1, y_2; \rho, \varphi)}{\partial(\gamma_1^2) \partial y_1} &= \sqrt{3} H_2(y_1) - \rho \frac{2}{3} \sqrt{\frac{3}{2}} H_1(y_1) H_1(y_2), \\ \lim_{\varphi \rightarrow 0} \frac{\partial^2 \ln f_2(y_1, y_2; \rho, \varphi)}{\partial(\gamma_1^2) \partial y_2} &= -\frac{1}{\sqrt{3}} \rho H_2(y_1), \\ \lim_{\varphi \rightarrow 0} \frac{\partial^2 \ln f_1(y_1; \varphi)}{\partial(\gamma_1^2) \partial y_1} &= \sqrt{3} H_3(y_1) - \frac{\rho^2}{\sqrt{3}} \left[H_2(y_1) + \sqrt{2} H_1^2(y_1) \right],\end{aligned}$$

and

$$\lim_{\varphi \rightarrow 0} \frac{\partial^2 \ln f_1(y_2; \varphi)}{\partial(\gamma_1^2) \partial y_1} = \rho^3 \left\{ \sqrt{3} H_3(y_1) - \frac{1}{\sqrt{3}} \left[H_2(y_1) + \sqrt{2} H_1^2(y_1) \right] \right\}.$$

Symmetric expressions apply to the terms $\partial^2./\partial(\gamma_4^2)\partial y_1$, and $\partial^2./\partial(\gamma_4^2)\partial y_2$. Analogous calculations, omitted for the sake of brevity, deliver the required expressions for the terms involving derivatives with respect to δ_5 . Finally, we obtain the desired result by collecting terms appropriately and using the fact that

$$\frac{\partial^2 \ln f_2(y_1, y_2; \rho, \varphi)}{\partial y_j^2} - \frac{\partial^2 \ln f_1(y_k; \varphi)}{\partial y_k^2} = -\frac{\rho^2}{1 - \rho^2} \quad \text{for } j, k = 1, 2,$$

and

$$\frac{\partial^2 \ln f_2(y_1, y_2; \rho, \varphi)}{\partial y_1 \partial y_2} = \frac{\rho}{1 - \rho^2}.$$

□

Proposition 11 *The LM test for a Gaussian copula against a Hermite polynomial expansion copula under the maintained assumption that $\rho = 0$ coincides with a moment test based on the following subset of influence functions:*

$$\mathbf{p}_\varphi(u_1, u_2; \rho) = \begin{pmatrix} s_{c_2}(y_1, y_2; \rho) \\ s_{c_3}(y_1, y_2; \rho) \\ s_{d_1}(y_1, y_2; \rho) \\ s_{d_2}(y_1, y_2; \rho) \\ s_{d_3}(y_1, y_2; \rho) \\ s_{d_4}(y_1, y_2; \rho) \end{pmatrix} = \begin{pmatrix} s_{c_2}(y_1, y_2; \rho) \\ s_{c_2}(y_2, y_1; \rho) \\ s_{d_1}(y_1, y_2; \rho) \\ s_{d_2}(y_1, y_2; \rho) \\ s_{d_3}(y_1, y_2; \rho) \\ s_{d_2}(y_2, y_1; \rho) \end{pmatrix}.$$

Proof. It follows directly from Proposition 10 and the fact that $\rho = 0$. □

Proposition 12 *The correction terms of the elements of $\mathbf{m}_\varphi(u_1, u_2; \rho)$ and $\mathbf{p}_\varphi(u_1, u_2; \rho)$ are given by*

$$\begin{aligned}n_{\gamma_1}(y_1, y_2; \rho) &= \frac{1}{3} \rho^2 (1 - \rho^2) [1 - H_3^2(y_1)] + \sqrt{\frac{2}{3}} \rho^2 [H_4(y_1) - H_4(y_2)] \\ &\quad + \frac{2\sqrt{2}}{3} \rho^2 (1 - \rho^2) H_2(y_1) + \frac{2\sqrt{2}}{3} \rho^6 H_2(y_2),\end{aligned}$$

$$\begin{aligned}
n_{c_2}(y_1, y_2; \rho) &= -\frac{\rho(1+\rho^2)}{3}H_3(y_1) - \frac{2}{\sqrt{3}}\rho^2H_3(y_2), \\
n_{d_1}(y_1, y_2; \rho) &= -\frac{1}{4}\rho^2(1+\rho^2+\rho^4)H_4(y_1) - \frac{3}{4}\rho^4H_4(y_2), \\
n_{d_2}(y_1, y_2; \rho) &= -\frac{1}{2}\rho(1+\rho^2+\rho^4)H_4(y_1) - \frac{3}{2}\rho^3H_4(y_2) - \sqrt{\frac{3}{2}}\rho^3[H_2(y_1) - H_2(y_2)], \\
n_{d_3}(y_1, y_2; \rho) &= -\frac{1}{4}\rho^2(2+\rho^2)[H_2^2(y_1) + H_2^2(y_2) - 2] - \frac{1}{\sqrt{2}}\rho^4[H_2(y_1) + H_2(y_2)],
\end{aligned}$$

and

$$\begin{aligned}
n_{\delta_5}(y_1, y_2; \rho) &= -\frac{1}{4}\rho^2(1-\rho^4)[H_4^2(y_1) + H_4^2(y_2) - 2] \\
&\quad + \frac{1}{4}(\rho^2 + 3\rho^4 - 3\rho^6 - \rho^8)[H_3^2(y_1) + H_3^2(y_2) - 2] \\
&\quad - \frac{9}{4}\sqrt{\frac{3}{2}}\rho^2(1-\rho^4)[H_4(y_1) + H_4(y_2)] \\
&\quad + \frac{3}{2\sqrt{2}}\rho^4(2\rho^4 + \rho^2 - 3)[H_2(y_1) + H_2(y_2)].
\end{aligned}$$

Proof. Analogous calculations to the ones performed to prove Proposition 8 applied to the corresponding influence functions. \square

C Tables and Figures

Table 1: Rejection rates under the null at 1%, 5%, and 10% significance levels:

Panel A: Bivariate copula ($K = 2$)									
	$N = 200$			$N = 800$			$N = 3,200$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
LM- t	.096	.043	.008	.100	.054	.012	.100	.054	.012
Skew	.102	.052	.011	.103	.053	.011	.103	.053	.011
LM- At	.101	.054	.008	.107	.053	.012	.106	.053	.012
KT- t	.100	.048	.009	.103	.050	.011	.103	.050	.011
KT- At	.100	.053	.008	.105	.054	.013	.105	.054	.013
LM-HE	.107	.054	.009	.105	.056	.011	.105	.056	.010
KS	.100	.053	.011	.104	.053	.011	.104	.053	.011
CvM	.100	.053	.010	.101	.051	.008	.101	.051	.009
Panel B: Trivariate copula ($K = 3$)									
	$N = 200$			$N = 800$			$N = 3,200$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
LM- t	.096	.049	.009	.103	.051	.010	.102	.052	.010
Skew	.109	.058	.012	.102	.051	.011	.101	.051	.011
LM- At	.108	.056	.011	.099	.050	.010	.099	.046	.010
KT- t	.095	.048	.010	.096	.049	.009	.098	.052	.012
KT- At	.099	.052	.011	.099	.050	.011	.099	.049	.010
KS	.108	.057	.011	.109	.052	.012	.104	.051	.011
CvM	.100	.054	.013	.103	.055	.008	.101	.051	.010

Notes: DGP: Gaussian with correlation parameter .12 in Panel A, while in Panel B $\rho_{12} = .013$, $\rho_{13} = -.041$, and $\rho_{23} = -.014$. Critical values are computed using parametric bootstrap. LM- t and LM- At are the Lagrange multiplier tests based on the score of the symmetric and asymmetric Student t copula, respectively; while KT- t and KT- At are the corresponding Kuhn-Tucker versions (see Section 2.3 for details). Skew corresponds to the Lagrange multiplier test based on the K moment conditions $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, 0)$ of Proposition 4 while LM-HE corresponds to the Lagrange multiplier test based on the score of the Hermite expansion of the Gaussian copula of Proposition 10. KS and CvM denote the Kolmogorov-Smirnov and the Cramér-von Mises tests for copula models, see Rémillard (2010) for details.

Table 2: Monte Carlo rejection rates under the null at 1%, 5%, and 10% significance levels under the Student t alternative:

Panel A: Bivariate copula ($K = 2$)									
	$N = 200$			$N = 800$			$N = 3,200$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
LM- t	.191	.125	.058	.403	.308	.159	.860	.788	.622
Skew	.136	.076	.021	.151	.087	.025	.144	.086	.022
LM- At	.187	.118	.046	.343	.251	.116	.757	.663	.454
KT- t	.277	.174	.074	.527	.386	.183	.923	.853	.660
KT- At	.228	.139	.052	.433	.301	.143	.836	.740	.489
LM-HE	.199	.130	.047	.330	.242	.108	.687	.582	.376
KS	.089	.050	.016	.101	.047	.008	.170	.089	.019
CvM	.095	.052	.010	.125	.061	.009	.282	.142	.029
Panel B: Trivariate copula ($K = 3$)									
	$N = 200$			$N = 800$			$N = 3,200$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
LM - t	.287	.203	.089	.725	.631	.427	.998	.995	.979
Skew	.159	.095	.025	.171	.101	.025	.176	.104	.027
LM - At	.254	.171	.061	.584	.469	.271	.987	.975	.926
KT - t	.411	.278	.099	.827	.723	.476	.999	.998	.987
KT - At	.317	.202	.066	.690	.547	.309	.995	.986	.944
KS	.076	.034	.006	.075	.033	.005	.129	.060	.013
CvM	.112	.059	.014	.146	.078	.013	.391	.217	.050

Notes: DGP: Student t with 20 degrees of freedom. The correlation parameter is .12 in Panel A, while in Panel B $\rho_{12} = .013$, $\rho_{13} = -.041$, and $\rho_{23} = -.014$. Critical values are computed using parametric bootstrap. LM- t and LM- At are the Lagrange multiplier tests based on the score of the symmetric and asymmetric Student t copula, respectively; while KT- t and KT- At are the corresponding Kuhn-Tucker versions (see Section 2.3 for details). Skew corresponds to the Lagrange multiplier test based on the K moment conditions $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, 0)$ of Proposition 4 while LM-HE corresponds to the Lagrange multiplier test based on the score of the Hermite expansion of the Gaussian copula of Proposition 10. KS and CvM denote the Kolmogorov-Smirnov and the Cramér-von Mises tests for copula models, see Rémillard (2010) for details.

Table 3: Monte Carlo rejection rates under the null at 1%, 5%, and 10% significance levels under the Asymmetric t alternative:

Panel A: Bivariate copula ($K = 2$)									
	$N = 200$			$N = 800$			$N = 3,200$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
LM- t	.190	.123	.060	.405	.311	.163	.871	.806	.635
Skew	.177	.106	.038	.344	.239	.104	.779	.694	.451
LM- At	.219	.147	.058	.488	.375	.198	.943	.905	.787
KT- t	.276	.175	.072	.533	.391	.189	.929	.864	.681
KT- At	.260	.166	.065	.568	.427	.230	.966	.934	.812
LM-HE	.233	.154	.061	.449	.341	.168	.900	.846	.693
KS	.150	.081	.029	.260	.156	.038	.593	.432	.196
CvM	.100	.051	.012	.177	.093	.019	.559	.383	.152
Panel B: Trivariate copula ($K = 3$)									
	$N = 200$			$N = 800$			$N = 3,200$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
LM- t	.321	.240	.114	.798	.714	.523	.999	.999	.993
Skew	.533	.425	.213	.972	.954	.869	1.000	1.000	1.000
LM- At	.572	.465	.252	.985	.971	.918	1.000	1.000	1.000
KT- t	.449	.312	.123	.877	.795	.571	1.000	.999	.997
KT- At	.592	.477	.257	.982	.970	.926	1.000	1.000	1.000
KS	.236	.130	.030	.494	.325	.099	.941	.869	.603
CvM	.149	.086	.029	.495	.374	.145	.989	.968	.864

Notes: DGP: asymmetric Student t with 20 degrees of freedom, and $b_i = -1$ for all i . The correlation parameter is .12 in Panel A, while in Panel B $\rho_{12} = .013$, $\rho_{13} = -.041$, and $\rho_{23} = -.014$. Critical values are computed using parametric bootstrap. LM- t and LM- At are the Lagrange multiplier tests based on the score of the symmetric and asymmetric Student t copula, respectively; while KT- t and KT- At are the corresponding Kuhn-Tucker versions (see Section 2.3 for details). Skew corresponds to the Lagrange multiplier test based on the K moment conditions $m_{b_k}(u_1, \dots, u_K; \boldsymbol{\rho}, 0)$ of Proposition 4 while LM-HE corresponds to the Lagrange multiplier test based on the score of the Hermite expansion of the Gaussian copula of Proposition 10. KS and CvM denote the Kolmogorov-Smirnov and the Cramér-von Mises tests for copula models, see Rémillard (2010) for details.

Table 4: Correlation parameter estimates

Sector	N.Obs	Short-term reversal strategies				Momentum strategies			
		Beta OLS	Correlation parameter			Beta OLS	Correlation parameter		
			Pearson	Spearman	Copula		Pearson	Spearman	Copula
Non Durables	33,215	.013 (.010)	.013 (.010)	-.028 (.004)	-.030 (.005)	.000 (.002)	.001 (.008)	.029 (.004)	.033 (.005)
Durables	6,816	.014 (.019)	.015 (.020)	-.038 (.009)	-.044 (.012)	.004 (.003)	.022 (.015)	.050 (.009)	.048 (.012)
Manufacturing	62,609	.014 (.007)	.015 (.007)	-.040 (.003)	-.038 (.004)	-.001 (.001)	-.003 (.006)	.023 (.003)	.024 (.004)
Energy	23,868	.005 (.009)	.006 (.009)	-.023 (.005)	-.027 (.006)	.000 (.000)	-.001 (.002)	.026 (.005)	.025 (.006)
Chemicals	11,030	-.013 (.016)	-.014 (.017)	-.051 (.007)	-.046 (.009)	-.001 (.004)	-.003 (.016)	.004 (.007)	.008 (.009)
Business	64,962	.010 (.008)	.010 (.008)	-.029 (.003)	-.025 (.004)	-.002 (.001)	-.011 (.007)	.017 (.003)	.012 (.004)
Telecom	12,357	.041 (.017)	.044 (.018)	.018 (.007)	.024 (.009)	.003 (.002)	.015 (.011)	.062 (.007)	.059 (.009)
Utilities	21,168	.019 (.014)	.019 (.014)	-.036 (.005)	-.037 (.007)	-.001 (.003)	-.003 (.013)	.004 (.005)	.012 (.007)
Shops	65,864	.019 (.011)	.019 (.011)	-.025 (.003)	-.025 (.004)	-.001 (.001)	-.004 (.004)	.034 (.003)	.033 (.004)
Healthcare	25,691	.018 (.012)	.019 (.013)	-.027 (.005)	-.026 (.006)	-.002 (.002)	-.008 (.010)	.032 (.005)	.028 (.006)
Financials	123,850	-.008 (.006)	-.009 (.006)	-.062 (.002)	-.055 (.003)	.001 (.002)	.016 (.018)	.031 (.002)	.034 (.003)
Others	155,624	.004 (.004)	.004 (.004)	-.024 (.002)	-.024 (.003)	.000 (.001)	.002 (.004)	.040 (.002)	.037 (.003)
All									
	607,054	.008 (.003)	.009 (.003)	-.025 (.001)	-.022 (.001)	.000 (.000)	.002 (.003)	.037 (.001)	.035 (.001)

Notes: The data is collected from CRSP and contain monthly series from July 1991 to December 2012. Industry definitions from Ken French's website. Numbers in parenthesis correspond to Newey and West (1987) standard errors; variances of ρ are corrected for heteroskedasticity and autocorrelation using 5 lags. KT- t and KT- At are the Kuhn-Tucker tests based on the score of the symmetric and asymmetric Student t copula, respectively (see Section 2.3 for details). Skew corresponds to the Lagrange multiplier test based on the K moment conditions $m_{b_k}(u_1, \dots, u_K; \rho, 0)$ of Proposition 4 while LM-HE corresponds to the Lagrange multiplier test based on the score of the Hermite expansion of the Gaussian copula of Proposition 10.

Table 5: Test statistics and p-values

Sector	N.Obs	Short-term reversal strategies			Momentum strategies				
		KT- <i>t</i>	Skew	KT- <i>At</i>	LM-H	KT- <i>t</i>	Skew	KT- <i>At</i>	LM-H
Non Durables	33,215	885.7 (.000)	40.3 (.000)	926.0 (.000)	989.4 (.000)	1,238.0 (.000)	184.8 (.000)	1,422.8 (.000)	1,526.9 (.000)
Durables	6,816	61.5 (.000)	11.4 (.003)	72.9 (.000)	93.3 (.000)	82.7 (.000)	22.0 (.000)	104.7 (.000)	139.4 (.000)
Manufacturing	62,609	1,437.5 (.000)	23.4 (.000)	1,461.0 (.000)	1,534.9 (.000)	2,245.6 (.000)	133.3 (.000)	2,378.9 (.000)	2,471.8 (.000)
Energy	23,868	510.2 (.000)	21.1 (.000)	531.3 (.000)	581.2 (.000)	563.1 (.000)	61.1 (.000)	624.2 (.000)	679.3 (.000)
Chemicals	11,030	168.4 (.000)	2.8 (.247)	171.1 (.000)	208.8 (.000)	209.7 (.000)	15.5 (.000)	225.2 (.000)	285.1 (.000)
Business	64,962	810.0 (.000)	34.3 (.000)	844.3 (.000)	883.6 (.000)	1,443.7 (.000)	120.6 (.000)	1,564.2 (.000)	1,651.2 (.000)
Telecom	12,357	236.0 (.000)	25.9 (.000)	261.9 (.000)	306.1 (.000)	211.9 (.000)	98.6 (.000)	310.6 (.000)	365.5 (.000)
Utilities	21,168	500.6 (.000)	34.6 (.000)	535.3 (.000)	581.8 (.000)	559.5 (.000)	34.1 (.000)	593.7 (.000)	635.9 (.000)
Shops	65,864	1,562.2 (.000)	54.8 (.000)	1,617.0 (.000)	1,657.8 (.000)	2,260.6 (.000)	356.0 (.000)	2,616.6 (.000)	2,743.7 (.000)
Healthcare	25,691	429.8 (.000)	19.7 (.000)	449.5 (.000)	496.2 (.000)	564.5 (.000)	120.0 (.000)	684.5 (.000)	763.4 (.000)
Financials	123,850	6,152.6 (.000)	369.4 (.000)	6,522.1 (.000)	6,872.8 (.000)	6,476.7 (.000)	1,238.4 (.000)	7,715.0 (.000)	7,966.5 (.000)
Others	155,624	3,053.2 (.000)	209.5 (.000)	3,262.7 (.000)	3,411.7 (.000)	4,339.6 (.000)	935.8 (.000)	5,275.5 (.000)	5,460.7 (.000)
All	607,054	24,333.7 (.000)	1,086.0 (.000)	25,419.7 (.000)	26,736.9 (.000)	32,408.0 (.000)	4,258.7 (.000)	36,666.7 (.000)	38,200.6 (.000)

Notes: The data is collected from CRSP and contain monthly series from July 1991 to December 2012. Industry definitions from Ken French's website. Numbers in parenthesis correspond to asymptotic p-values. Both, variances of the test moment functions are corrected for heteroskedasticity and autocorrelation using 5 lags. KT- t and KT- At are the Kuhn-Tucker tests based on the score of the symmetric and asymmetric Student t copula, respectively (see Section 2.3 for details). Skew corresponds to the Lagrange multiplier test based on the K moment conditions $m_{b_k}(u_1, \dots, u_K; \rho, 0)$ of Proposition 4 while LM-HE corresponds to the Lagrange multiplier test based on the score of the Hermite expansion of the Gaussian copula of Proposition 10.

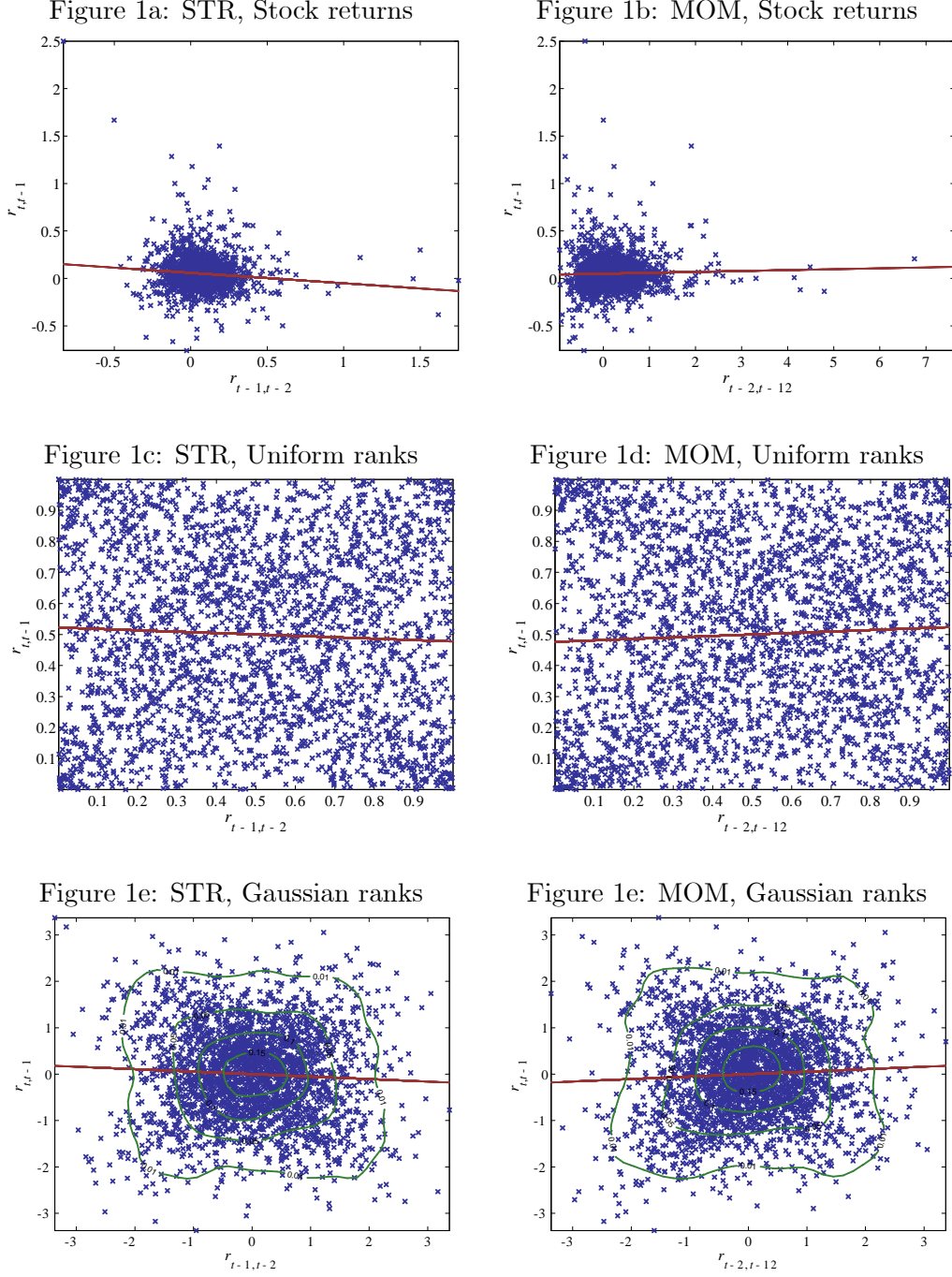
Table 6: Constrained estimates of the shape parameters:

Panel A: Short term reversals strategies						
Sector	Student t		Asymmetric Student t			
	$\hat{\rho}$	$\hat{\eta}$	$\hat{\rho}$	$\hat{\eta}$	\hat{b}_1	\hat{b}_2
Non Durables	-.032	.154	-.029	.155	-.135	-.049
Durables	-.047	.093	-.045	.093	-.156	-.219
Manufacturing	-.040	.144	-.039	.144	-.065	-.045
Energy	-.029	.139	-.027	.139	-.076	-.091
Chemicals	-.049	.119	-.048	.120	-.103	-.014
Business	-.028	.108	-.027	.109	-.125	-.064
Telecom	.022	.131	.027	.130	-.154	-.142
Utilities	-.040	.146	-.036	.146	-.124	-.108
Shops	-.028	.146	-.026	.146	-.109	-.057
Healthcare	-.028	.124	-.026	.124	-.117	-.075
Financials	-.058	.209	-.042	.207	-.117	-.091
Others	-.027	.133	-.024	.134	-.163	-.069
All	-.025	.187	-.018	.187	-.112	-.069

Panel B: Momentum strategies						
Sector	Student t		Asymmetric Student t			
	$\hat{\rho}$	$\hat{\eta}$	$\hat{\rho}$	$\hat{\eta}$	\hat{b}_1	\hat{b}_2
Non Durables	.032	.179	.051	.178	-.170	-.176
Durables	.046	.105	.051	.104	-.262	-.225
Manufacturing	.022	.176	.028	.176	-.087	-.114
Energy	.023	.144	.028	.143	-.131	-.159
Chemicals	.006	.131	.007	.130	-.054	-.140
Business	.010	.141	.014	.140	-.146	-.113
Telecom	.057	.123	.062	.118	-.113	-.467
Utilities	.010	.152	.014	.152	-.117	-.102
Shops	.032	.172	.046	.170	-.136	-.208
Healthcare	.026	.139	.034	.137	-.162	-.247
Financials	.033	.211	.080	.209	-.139	-.252
Others	.035	.155	.047	.153	-.150	-.259
All	.034	.213	.074	.212	-.124	-.190

Notes: The data is collected from CRSP and contain monthly series from July 1991 to December 2012. Estimates are obtained by generating data from this copula and matching in the simulated data the values in the original data of both the Gaussian rank correlation coefficients and the corresponding test statistics.

Figure 1: Short term reversal and momentum, March 1998



Notes: The data is collected from CRSP. STR refers to short term reversal and MOM to momentum. Red lines in the top panels represent the regression lines of the original data, with beta coefficients: $-.109$ in Figure 1a and $.009$ in Figure 1b; red lines in the middle panels correspond to the Spearman rank correlation: $-.045$ and $.023$ in Figures 1c and 1d, respectively; and, red lines in the bottom panels represent the Gaussian rank correlation: $-.053$ in Figure 1e and $.053$ in Figure 1f. Contours in the bottom panels are computed through non-parametric estimation of the bivariate density using Silverman's rule of thumb $N^{-1/6}$ for the standardized Gaussian ranks.

Figure 2: Short term reversal and momentum, August 2007

Figure 2a: STR, Stock returns

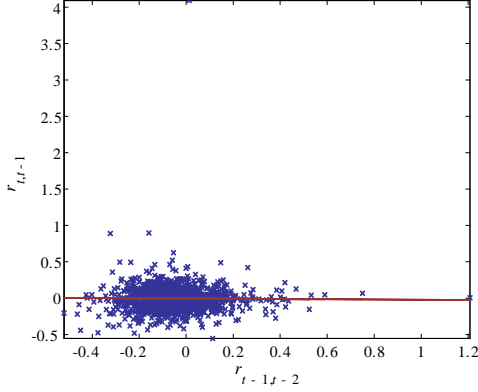


Figure 2b: MOM, Stock returns

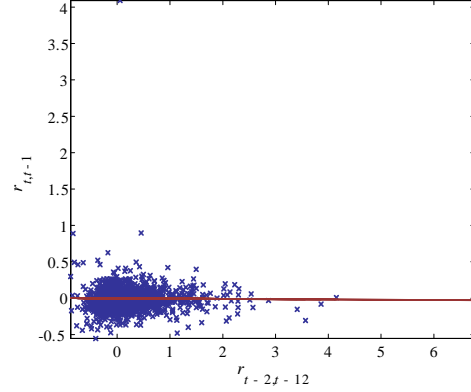


Figure 2c: STR, Uniform ranks

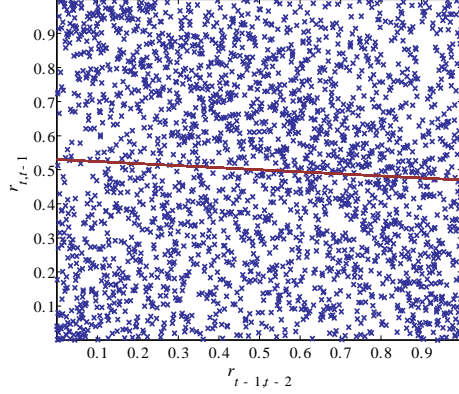


Figure 2d: MOM, Uniform ranks

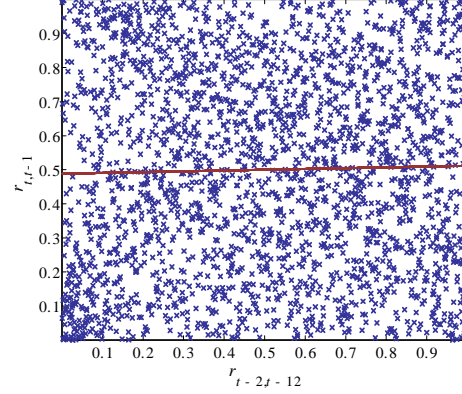


Figure 2e: STR, Gaussian ranks

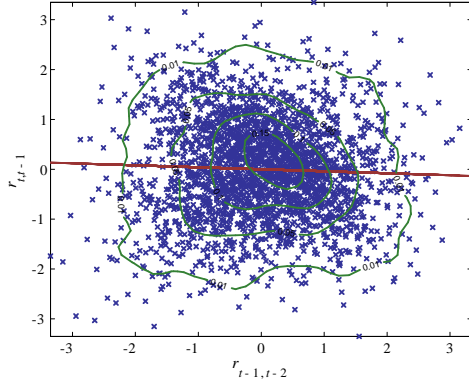
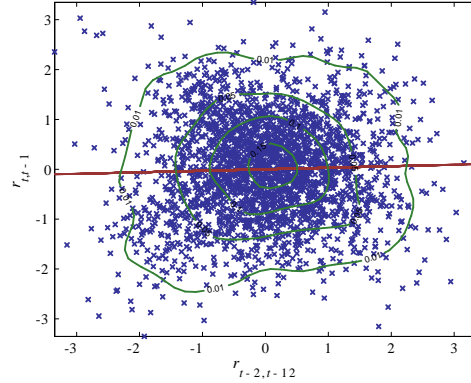


Figure 2f: MOM, Gaussian ranks



Notes: The data is collected from CRSP. STR refers to short term reversal and MOM to momentum. Red lines in the top panels represent the regression lines of the original data, with beta coefficients: -.019 in Figure 2a and -.004 in Figure 2b; red lines in the middle panels correspond to the Spearman rank correlation: -.060 and .023 in Figures 2c and 2d, respectively; and, red lines in the bottom panels represent the Gaussian rank correlation: -.040 in Figure 2e and .030 in Figure 2f. Contours in the bottom panels are computed through non-parametric estimation of the bivariate density using Silverman's rule of thumb $N^{-1/6}$ for the standardized Gaussian ranks.

Figure 3: Gaussian copula

Figure 3a: Bivariate Gaussian copula density with uniform margins

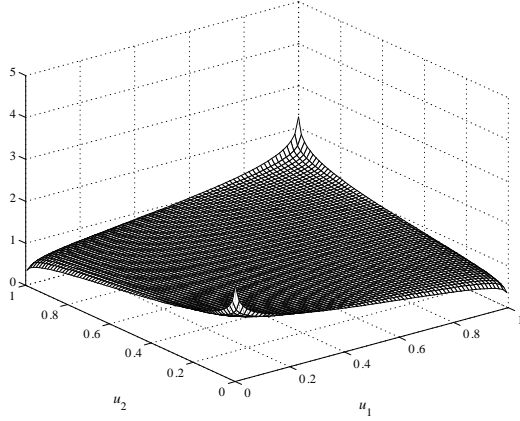


Figure 3b: Contours of a bivariate Gaussian copula with uniform margins

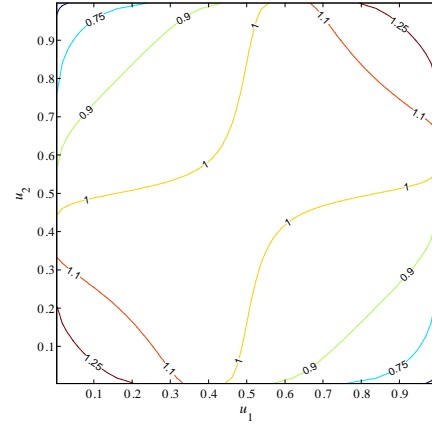


Figure 3c: Bivariate Gaussian copula with Gaussian margins

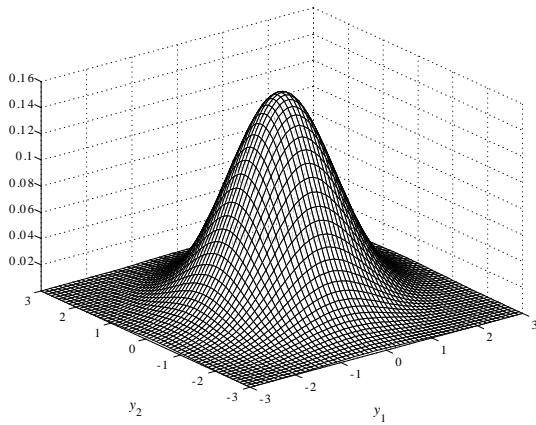


Figure 3d: Contours of a bivariate Gaussian copula with Gaussian margins

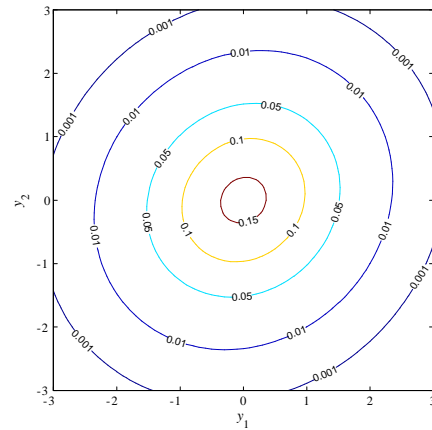
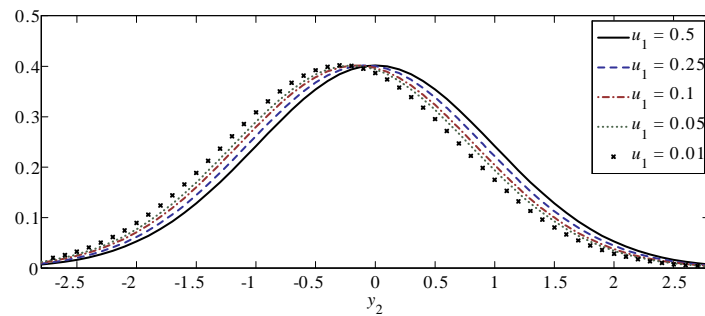


Figure 3e: Conditional density from Gaussian copula with Gaussian marginals



Notes: Gaussian copula with correlation coefficient $\rho = .120$ (Spearman correlation $\rho_S = .115$). Top panels plot density (left) and contours (right) of the corresponding copula on $[0, 1]^2$ while middle panels do the same when marginals are standard normal.

Figure 4: Student t copula

Figure 4a: Bivariate Student t copula density with uniform margins

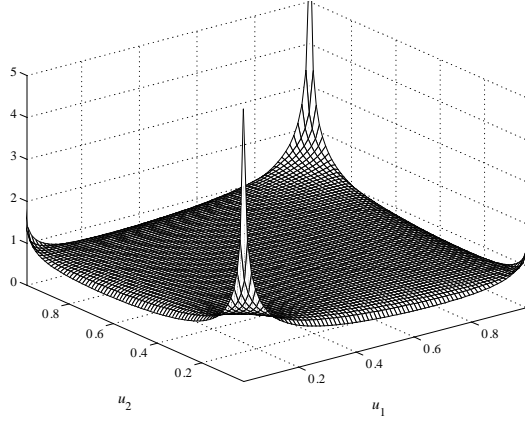


Figure 4b: Contours of a bivariate Student t copula with uniform margins

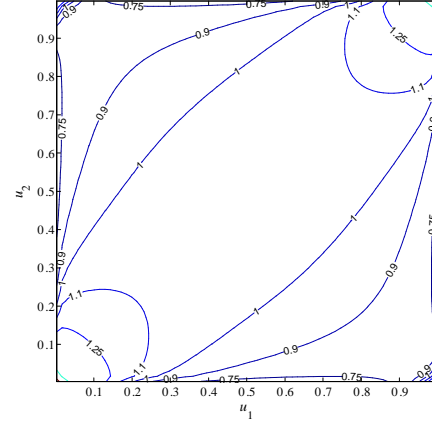


Figure 4c: Bivariate Student t copula density with Gaussian margins

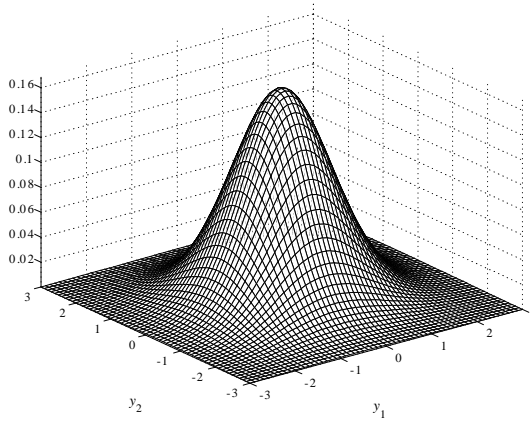


Figure 4d: Contours of a bivariate Student t copula with Gaussian margins

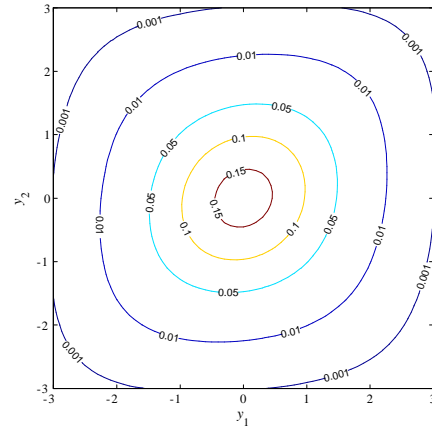
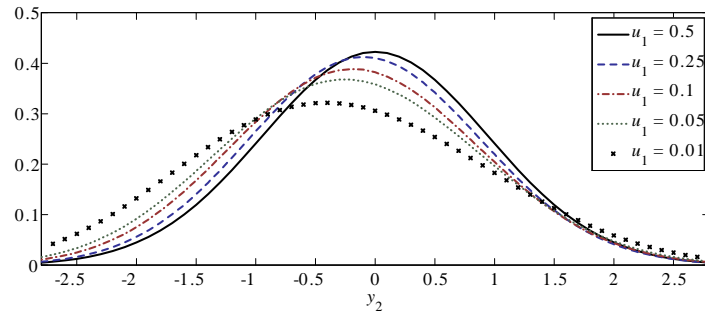


Figure 4e: Conditional density from Student t copula with Gaussian marginals



Notes: Student t copula with 10 degrees of freedom and correlation coefficient $\rho = .122$ (Spearman correlation $\rho_S = .115$). Top panels plot density (left) and contours (right) of the corresponding copula on $[0, 1]^2$ while middle panels do the same when marginals are standard normal.

Figure 5: Asymmetric Student t copula

Figure 5a: Bivariate asymmetric Student t copula density with uniform margins

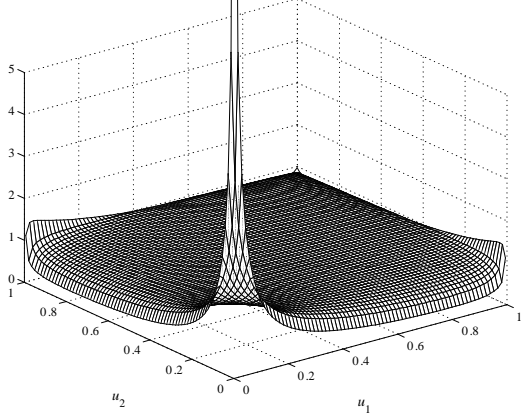


Figure 5b: Contours of a bivariate asymmetric Student t copula with uniform margins

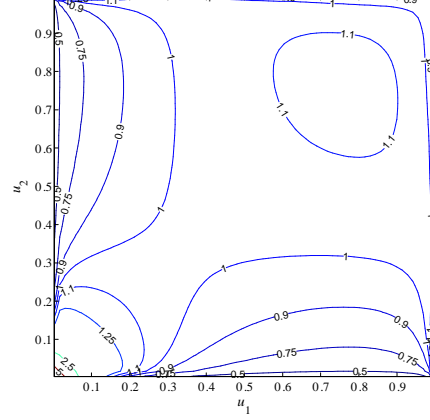


Figure 5c: Bivariate asymmetric Student t copula density with Gaussian margins

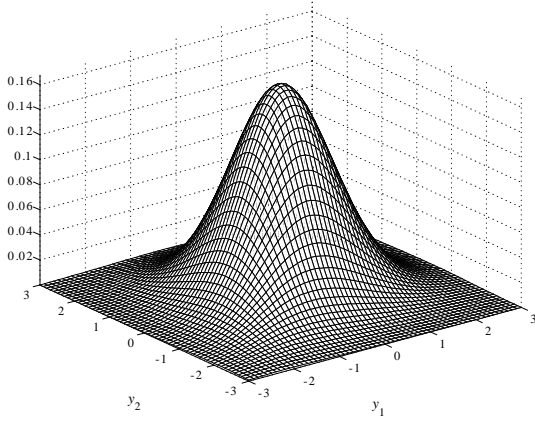


Figure 5d: Contours of a bivariate asymmetric Student t copula with Gaussian margins

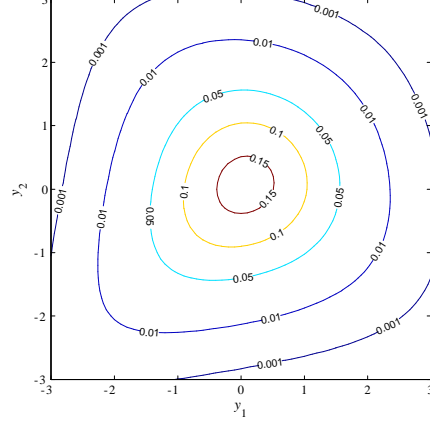
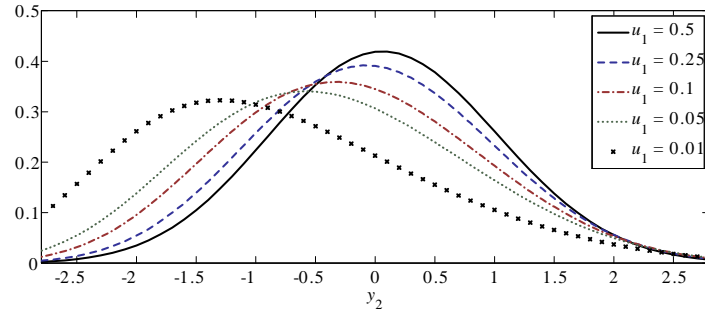


Figure 5e: Conditional density from asymmetric Student t copula with Gaussian marginals



Notes: Asymmetric Student t copula with 10 degrees of freedom, skewness parameters $b_i = -.5$ and correlation coefficient $\rho = .186$ (Spearman correlation $\rho_S = .115$). Top panels plot density (left) and contours (right) of the corresponding copula on $[0, 1]^2$ while middle panels do the same when marginals are standard normal.

Figure 6: Hermite expansion copula

Figure 6a: Bivariate Hermite expansion copula density with uniform margins

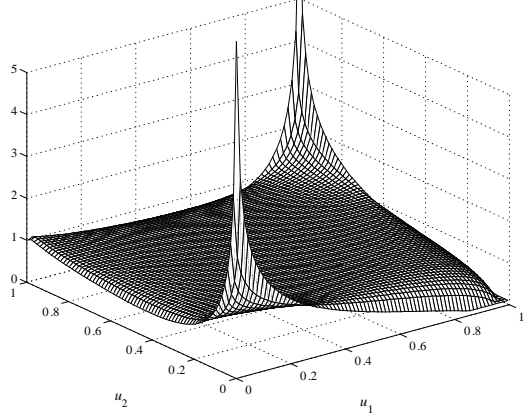


Figure 6b: Contours of a bivariate Hermite expansion copula with uniform margins

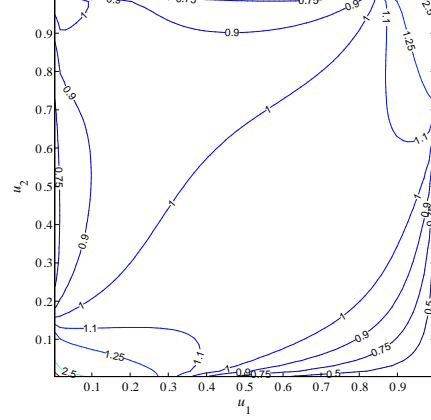


Figure 6c: Bivariate Hermite expansion copula density with Gaussian margins

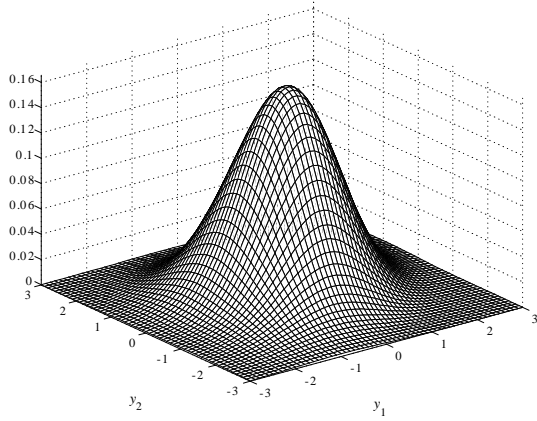


Figure 6d: Contours of a bivariate Hermite expansion copula with Gaussian margins

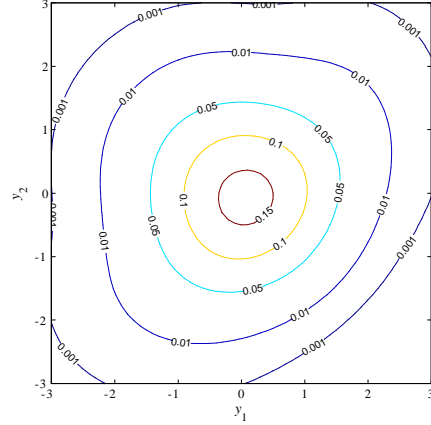
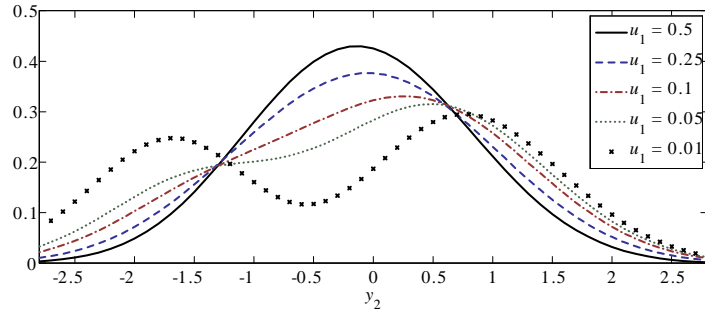


Figure 6e: Conditional density from Hermite expansion copula with Gaussian marginals



Notes: Hermite expansion copula with $\varphi = (0, -1, 1, 0, 0, .05, .05, .05, 0)'$ and correlation coefficient $\rho = .140$ (Spearman correlation $\rho_S = .115$). Top panels plot density (left) and contours (right) of the corresponding copula on $[0, 1]^2$ while middle panels do the same for the originating distribution when marginals are standard normal.

Figure 7: Power of Student t -based tests under asymmetric Student t local alternatives

Figure 7a: Non-centrality parameter for different kurtosis parameter values

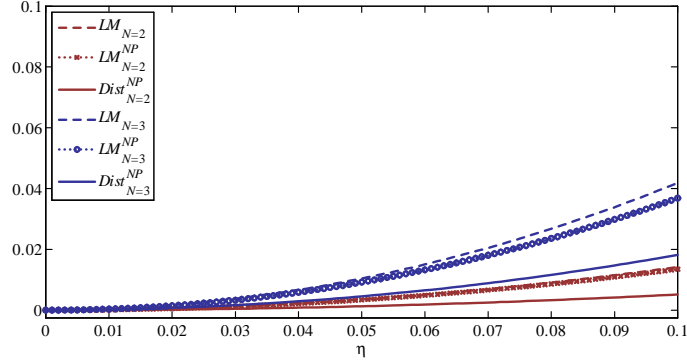


Figure 7b: Non-centrality parameter for different correlation parameter values

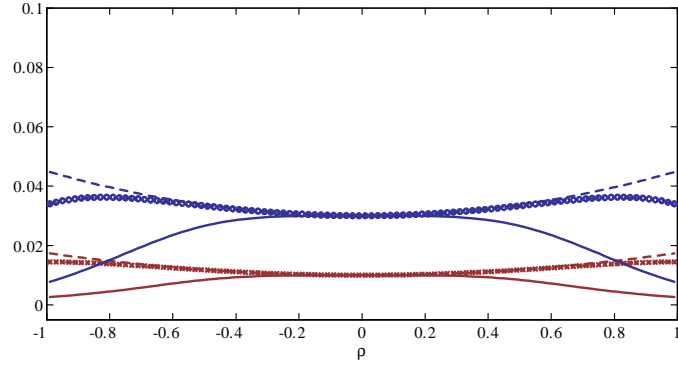
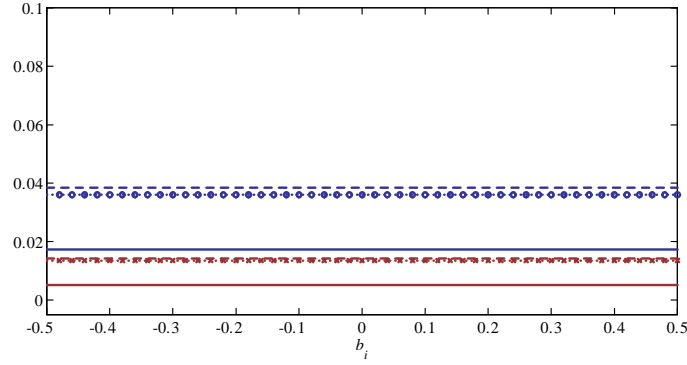


Figure 7c: Non-centrality parameter for different skewness parameter values



Notes: Non-centrality parameters of the Student t -based LM-copula and LM-distributional tests under asymmetric Student t alternatives. LM and LM^{NP} denote the LM-copula tests when marginals are known and when they are estimated nonparametrically, respectively; while $Dist^{NP}$ denotes the LM-distributional test when marginals are estimated nonparametrically. Figure 7a, power against asymmetric Student t alternatives with $\rho = .75$ and $b_i = -.5$. Figure 7b, power against asymmetric Student t alternatives with different correlation parameter and $\eta = .1$, $b_i = -.5$ for $i = 1, 2$. Figure 7c, power against asymmetric Student t alternatives with increasing skewness and $\eta = .1$, $\rho = .75$. Figures 7b-c share the legend of Figure 7a.

Figure 8: Power of asymmetric Student t -based tests under asymmetric Student t local alternatives

Figure 8a: Non-centrality parameter for different kurtosis parameter values

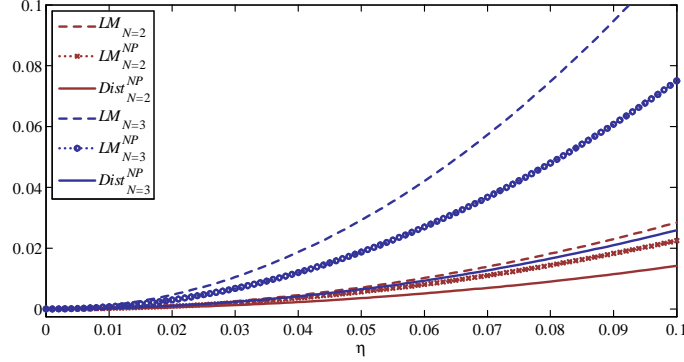


Figure 8b: Non-centrality parameter for different correlation parameter values

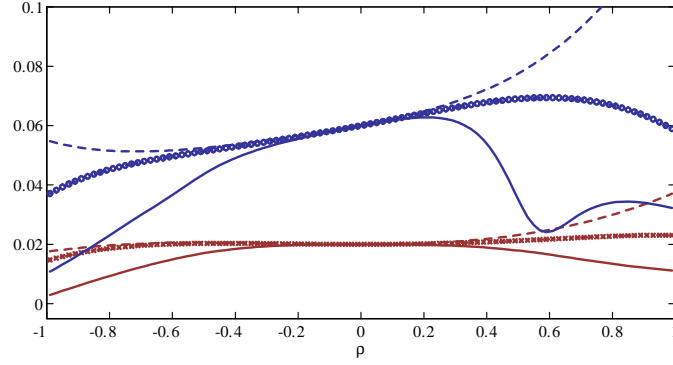
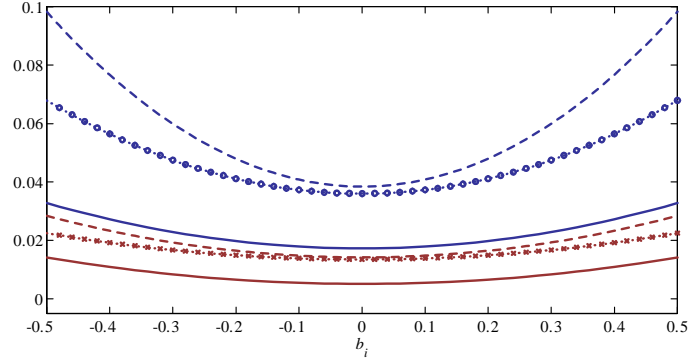


Figure 8c: Non-centrality parameter for different skewness parameter values



Notes: Non-centrality parameters of the Student t -based LM-copula and LM-distributional tests under asymmetric Student t alternatives. LM and LM^{NP} denote the LM-copula tests when marginals are known and when they are estimated nonparametrically, respectively; while $Dist^{NP}$ denotes the LM-distributional test when marginals are estimated nonparametrically. Figure 8a, power against asymmetric Student t alternatives with $\rho = .75$ and $b_i = -.5$. Figure 8b, power against asymmetric Student t alternatives with different correlation parameter and $\eta = .1$, $b_i = -.5$ for $i = 1, 2$. Figure 8c, power against asymmetric Student t alternatives with increasing skewness and $\eta = .1$, $\rho = .75$. Figures 8b-c share the legend of panel Figure 8a.

Figure 9: Size distortion of Student t -based and asymmetric Student t -based tests at the 5% level

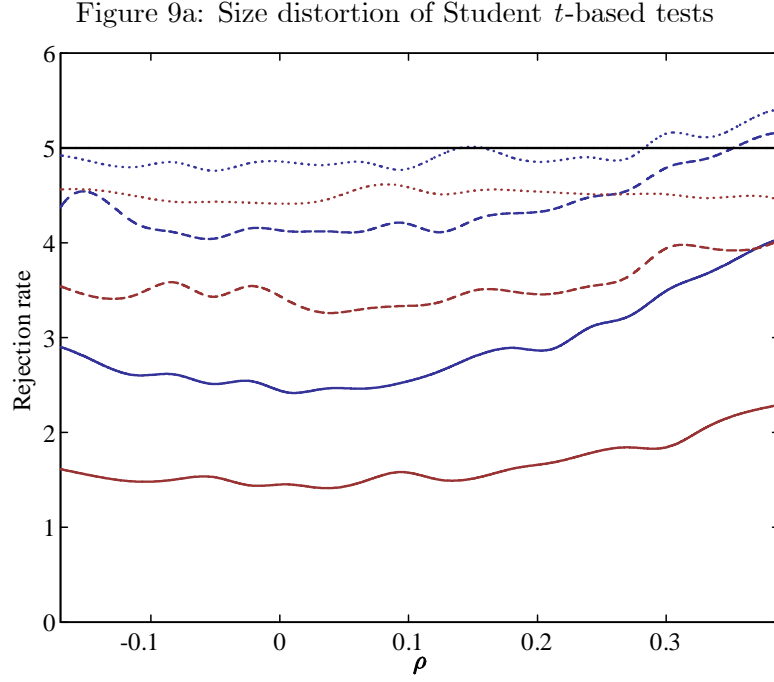
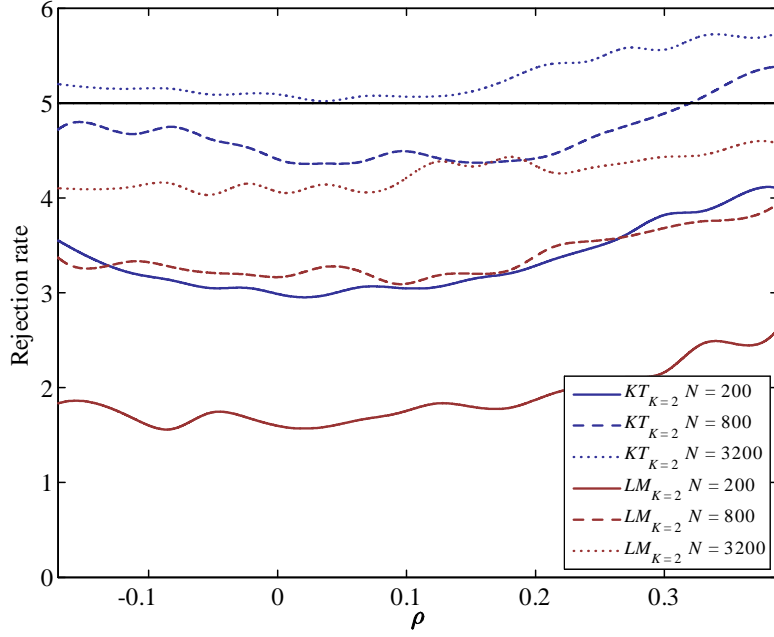
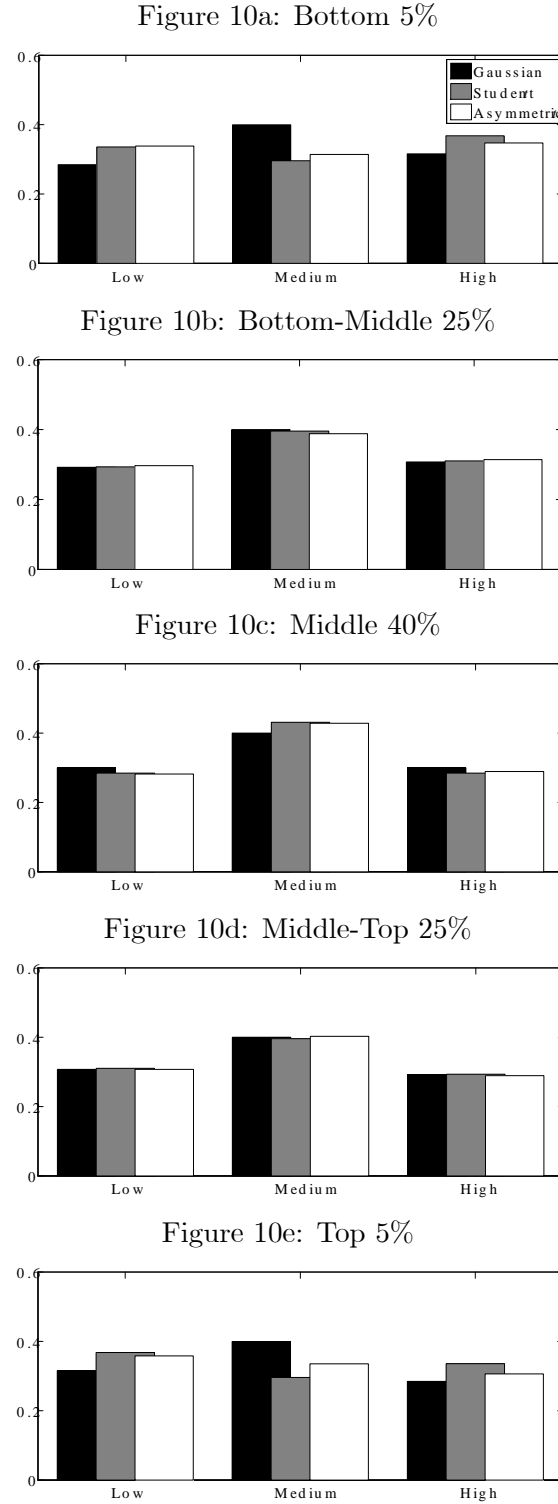


Figure 9b: Size distortion of asymmetric Student t -based tests



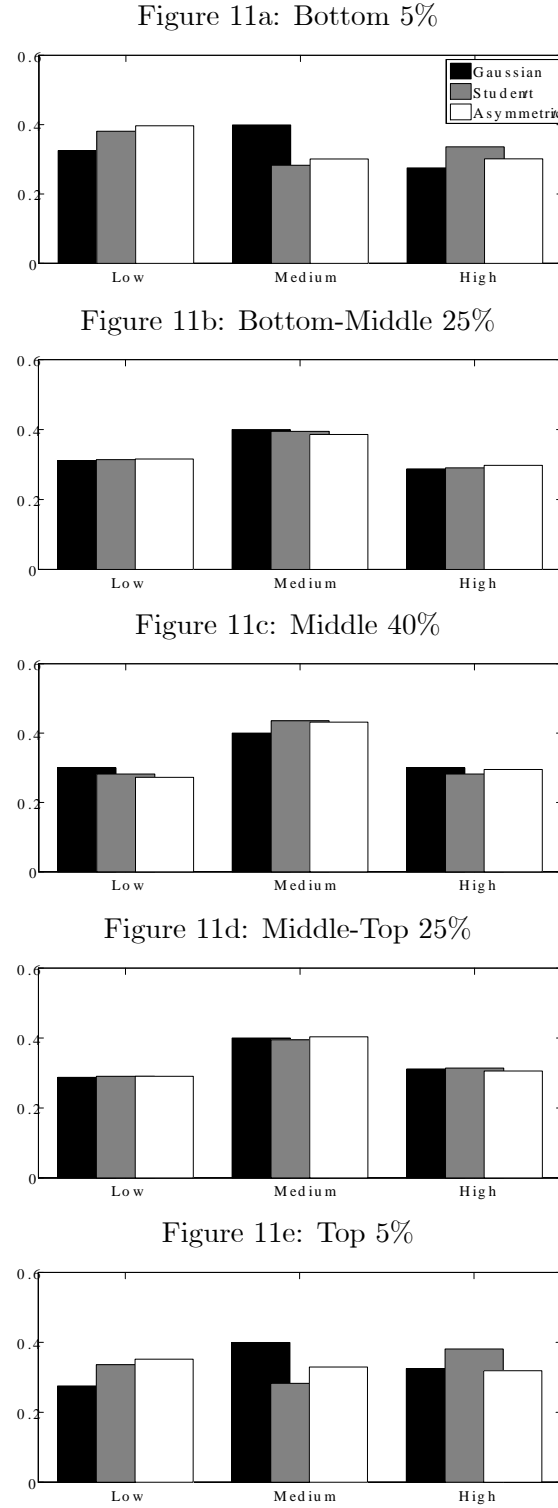
Notes: DGP: Gaussian. Based on 10,000 replications. In Figure 9a, $LM_{K=2}$ and $KT_{K=2}$ are the Lagrange multiplier test and its Kuhn-Tucker version based on the score of the Student t copula; while in Figure 9b, $LM_{K=2}$ and $KT_{K=2}$ are the Lagrange multiplier test and its Kuhn-Tucker version based on the score of the asymmetric Student t copula (see Section 2.3 for details).

Figure 10: Transition probabilities for short term reversals strategies



Notes: The data is collected from CRSP and contain monthly series from July 1991 to December 2012. Gaussian copula with correlation coefficient $\rho = -.022$. For the Student t copula, $\rho = -.025$ and $\eta = .187$; while for the asymmetric Student t copula, $\rho = -.018$, $\eta = .187$, $b_1 = -.112$ and $b_2 = -.069$ (obtained by restricted indirect estimation).

Figure 11: Transition probabilities for momentum strategies



Notes: The data is collected from CRSP and contain monthly series from July 1991 to December 2012. Gaussian copula with correlation coefficient $\rho = .035$. For the Student t copula, $\rho = .034$ and $\eta = .213$; while for the asymmetric Student t copula, $\rho = .074$, $\eta = .212$, $b_1 = -.124$ and $b_2 = -.190$ (obtained by restricted indirect estimation).

Supplemental Appendices for
Is a normal copula the right copula?

Dante Amengual

CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain

<amengual@cemfi.es>

Enrique Sentana

CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain

<sentana@cemfi.es>

August 2015

D Local power calculations

Let $\mathbf{m}_{\varphi t}(\boldsymbol{\rho}, \boldsymbol{\varphi})$ denote the h influence functions used to develop the following moment test of $H_0 : \boldsymbol{\varphi} = \mathbf{0}$:

$$M_T = T \bar{\mathbf{m}}'_{\varphi t}(\boldsymbol{\rho}_0, \mathbf{0}) \boldsymbol{\Psi}^{-1} \bar{\mathbf{m}}_{\varphi t}(\boldsymbol{\rho}_0, \mathbf{0}), \quad (\text{D6})$$

where $\bar{\mathbf{m}}_{\varphi t}(\boldsymbol{\rho}_0, \mathbf{0})$ is the sample average of $\mathbf{m}_{\varphi t}(\boldsymbol{\rho}, \boldsymbol{\varphi})$ evaluated under the null, and $\boldsymbol{\Psi}$ is the corresponding asymptotic covariance matrix. In order to obtain the non-centrality parameter of this test under Pitman sequences of local alternatives of the form $H_l : \boldsymbol{\varphi}_T = \bar{\boldsymbol{\varphi}}/\sqrt{T}$, it is convenient to linearize $\mathbf{m}_{\varphi t}(\boldsymbol{\rho}_0, \mathbf{0})$ with respect to $\boldsymbol{\varphi}$ around its true value $\boldsymbol{\varphi}_T$. This linearization yields

$$\sqrt{T} \bar{\mathbf{m}}_{\varphi t}(\boldsymbol{\rho}_0, \mathbf{0}) = \sqrt{T} \bar{\mathbf{m}}_{\varphi t}(\boldsymbol{\rho}_0, \boldsymbol{\varphi}_T) + \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial \mathbf{m}_{\varphi t}(\boldsymbol{\rho}, \boldsymbol{\varphi}^*)}{\partial \boldsymbol{\varphi}'} \bar{\boldsymbol{\varphi}},$$

where $\boldsymbol{\varphi}^*$ is some “intermediate” value between $\boldsymbol{\varphi}_T$ and $\mathbf{0}$. As a result,

$$\sqrt{T} \bar{\mathbf{m}}_{\varphi t}(\boldsymbol{\rho}_0, \mathbf{0}) \xrightarrow{d} N[\mathbf{M}(\boldsymbol{\rho}_0, \mathbf{0}) \bar{\boldsymbol{\varphi}}, \boldsymbol{\Psi}],$$

under standard regularity conditions, where

$$\mathbf{M}(\boldsymbol{\rho}_0, \mathbf{0}) = E \left[\frac{\partial \mathbf{m}_{\varphi t}(\boldsymbol{\rho}, \mathbf{0})}{\partial \boldsymbol{\varphi}'} \right],$$

so that the non-centrality parameter of the moment test (D6) will be

$$\bar{\boldsymbol{\varphi}}' \mathbf{M}'(\boldsymbol{\rho}_0, \mathbf{0}) \boldsymbol{\Psi}^{-1} \mathbf{M}(\boldsymbol{\rho}_0, \mathbf{0}) \bar{\boldsymbol{\varphi}}. \quad (\text{D7})$$

On this basis, we can easily obtain the limiting probability of M_T exceeding some prespecified quantile of a central χ_h^2 distribution from the cdf of a non-central χ^2 distribution with h degrees of freedom and non-centrality parameter (D7).

Finally, note that (D7) remains valid when we replace $\boldsymbol{\rho}_0$ by its ML estimator under the null if $\mathbf{m}_{\varphi t}(\boldsymbol{\rho}, \mathbf{0})$ and the scores corresponding to $\boldsymbol{\rho}$ are asymptotically uncorrelated when H_0 is true, as in all our tests. In addition, both $\mathbf{M}(\boldsymbol{\rho}_0, \mathbf{0})$ and $\boldsymbol{\Psi}$ coincide with the (2, 2) block of the information matrix when $\mathbf{m}_{\varphi t}(\boldsymbol{\rho}, \boldsymbol{\varphi})$ are the scores with respect to $\boldsymbol{\varphi}$.

D.1 Student t alternatives

Expressions for $s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, 0)$ as well as for $n_\eta(u_1, \dots, u_K; \boldsymbol{\rho})$, which allows to compute

$$s_\eta^c(u_1, \dots, u_K; \boldsymbol{\rho}, 0) = s_\eta(u_1, \dots, u_K; \boldsymbol{\rho}, 0) - n_\eta(u_1, \dots, u_K; \boldsymbol{\rho}),$$

are given in Propositions 3 and 8, respectively. For instance, in the bivariate case, since both $E[\partial m_{\eta t}(\rho, 0)/\partial \eta]$ and $V[m_{\eta t}(\rho, 0)]$ coincide with the (2, 2) block of the information matrix, we only need to compute

$$V[s_{\eta}(u_1, u_2; \rho, 0)] = 1 + \frac{3}{4}\rho^2$$

and

$$V[s_{\eta}^c(u_1, u_2; \rho, 0)] = 1 + \frac{3}{4}\rho^2 + \frac{3}{16}(\rho^4 + \rho^8)$$

in order to obtain the corresponding non-centrality parameters. Similarly, for the distributional version of the test, we have that $m_{\eta t}(\rho, 0) = d^c(u_1, u_2; \rho, 0)$ with

$$d^c(u_1, u_2; \rho, 0) = 2L_2[\varsigma(\rho)] - \sqrt{\frac{3}{2}}[H_4(y_1) + H_4(y_2)],$$

and hence

$$V[d^c(u_1, u_2; \rho, 0)] = 1 + 3\rho^4$$

and

$$\text{cov}[d^c(u_1, u_2; \rho, 0), s_{\eta}^c(u_1, u_2; \rho, 0)] = 1 - \frac{3}{4}\rho^6.$$

D.2 Asymmetric Student t alternatives

The required quantities to compute the non-centrality parameters of the score test in the bivariate case are

$$V[s_{b_k}(u_1, u_2; \rho, 0)] = 26 + 24\rho^2 + 48\rho^4,$$

$$\text{cov}[s_{b_1}(u_1, u_2; \rho, 0), s_{b_2}(u_1, u_2; \rho, 0)] = 48\rho + 26\rho^3 + 24\rho^5,$$

$$V[s_{b_k}^c(u_1, u_2; \rho, 0)] = 2 + \frac{2}{3}(\rho^2 + \rho^4) + \frac{4}{3}\rho^6,$$

and

$$\text{cov}[s_{b_1}^c(u_1, u_2; \rho, 0), s_{b_2}^c(u_1, u_2; \rho, 0)] = \frac{10}{3}\rho^3 + \frac{2}{3}(\rho^5 + \rho^7),$$

while $\text{cov}[s_{\eta}(u_1, u_2; \rho, 0), s_{b_k}(u_1, u_2; \rho, 0)] = \text{cov}[s_{\eta}^c(u_1, u_2; \rho, 0), s_{b_k}^c(u_1, u_2; \rho, 0)] = 0$. The same argument can be applied to the distributional test yielding

$$d_{b_1}^c(u_1, u_2; \rho, 0) = -2 \left[\sqrt{\frac{3}{2}}H_3(y_1) + \rho\sqrt{\frac{2}{3}}H_3(y_2) \right] + y_1[\varsigma(\rho) - 4]$$

and

$$d_{b_1}^c(u_1, u_2; \rho, 0) = -2 \left[\sqrt{\frac{3}{2}}H_3(y_2) + \rho\sqrt{\frac{2}{3}}H_3(y_1) \right] + y_2[\varsigma(\rho) - 4].$$

As in the case of the score test, $d_{b_k}^c(u_1, u_2; \rho, 0)$ for $k = 1, 2$ is orthogonal to $d_{b_1}^c(u_1, u_2; \rho, 0)$. Therefore, the additional quantities required to compute the corresponding non-centrality parameters are

$$\begin{aligned} V[d_{b_k}^c(u_1, u_2; \rho, 0)] &= 2 - \frac{16}{3}\rho^2 + 8\rho^4, \\ \text{cov}[d_{b_1}^c(u_1, u_2; \rho, 0), m_{b_2}^c(u_1, u_2; \rho, 0)] &= -4\rho + 6\rho^3 + \frac{8}{3}\rho^5, \\ \text{cov}[m_{b_k}^c(u_1, u_2; \rho, 0), s_{b_k}^c(u_1, u_2; \rho, 0)] &= 2 - \frac{10}{3}\rho^2 - 2\rho^4 - \frac{4}{3}\rho^6, \end{aligned}$$

and

$$\text{cov}[m_{b_1}^c(u_1, u_2; \rho, 0), s_{b_2}^c(u_1, u_2; \rho, 0)] = -2\rho + \frac{2}{3}\rho^3 - \frac{10}{3}\rho^5.$$

D.3 Interpretation of copula and distributional tests (the bivariate case)

D.3.1 When marginals are known

We can easily express both score copula tests as well as distributional LM tests in terms of Hermite polynomials of the marginal Gaussian ranks. Taking into account that $m_{b_2}(y_1, y_2; \rho) = m_{b_1}(y_2, y_1; \rho)$ and $d_{b_2}(y_1, y_2; \rho) = d_{b_1}(y_2, y_1; \rho)$, the relevant coefficients are in Table D1.

In order to characterize the loss of power of the distributional version of the test we could write

$$d_\varphi(y_1, y_2; \rho) = \beta_\varphi s_\varphi(y_1, y_2; \rho) + u_\varphi$$

where

$$\beta_\varphi = \frac{\text{cov}[d_\varphi(y_1, y_2; \rho), s_\varphi(y_1, y_2; \rho)]}{V[s_\varphi(y_1, y_2; \rho)]}$$

so that the non-centrality parameter of $d_\varphi(y_1, y_2; \rho)$ under a sequence of local alternatives $H_l : \varphi_T = \bar{\varphi}/\sqrt{T}$ can be written as

$$\frac{\beta_\varphi^2 V[s_\varphi(y_1, y_2; \rho)]}{\beta_\varphi^2 V[s_\varphi(y_1, y_2; \rho)] + V(u_\varphi)}.$$

For instance, when $\varphi = \eta$ we have that

$$V[s_\eta(y_1, y_2; \rho)] = 1 + \frac{3}{4}\rho^2$$

and

$$\text{cov}[d_\eta(y_1, y_2; \rho), s_\eta(y_1, y_2; \rho)] = 1,$$

so that the reduction in power of the distributional LM test relative to the score test is

$$V(u_\eta) = 4 - \frac{4}{4 + 3\rho^2}$$

where we have used the fact that $V[d_\eta(y_1, y_2; \rho)] = 4$.

Table D1: Hermite polynomial coefficients for bivariate score copula tests and distributional LM tests when marginals are known

Hermite polynomial	Copula LM test		Distributional LM test	
	$s_\eta(y_1, y_2; \rho)$	$m_{b_1}(y_1, y_2; \rho)$	$d_\eta(y_1, y_2; \rho)$	$d_{b_1}(y_1, y_2; \rho)$
1	$\frac{2\rho^4 + \rho^2}{(1-\rho^2)^2}$	0	$\frac{2\rho^4 + \rho^2}{(1-\rho^2)^2}$	0
$H_1(y_1)$	0	$\frac{2\rho^2}{1-\rho^2}$	0	$\frac{4\rho^2}{1-\rho^2}$
$H_1(y_2)$	0	$-\frac{2(\rho^3 + \rho)}{1-\rho^2}$	0	$-\frac{2\rho}{1-\rho^2}$
$H_2(y_1)$	$\frac{3\rho^2(\rho^2 + 3)}{2\sqrt{2}(1-\rho^2)^2}$	0	$\frac{3\sqrt{2}\rho^2}{(1-\rho^2)^2}$	0
$H_1(y_1)H_1(y_2)$	$-\frac{2(2\rho^3 + \rho)}{(1-\rho^2)^2}$	0	$-\frac{2(2\rho^3 + \rho)}{(1-\rho^2)^2}$	0
$H_2(y_2)$	$\frac{3\rho^2(\rho^2 + 3)}{2\sqrt{2}(1-\rho^2)^2}$	0	$\frac{3\sqrt{2}\rho^2}{(1-\rho^2)^2}$	0
$H_3(y_1)$	0	0	0	$\frac{\sqrt{6}}{1-\rho^2}$
$H_2(y_1)H_1(y_2)$	0	$-\frac{\sqrt{2}\rho}{1-\rho^2}$	0	$-\frac{2\sqrt{2}\rho}{1-\rho^2}$
$H_1(y_1)H_2(y_2)$	0	$\frac{\sqrt{2}(\rho^2 + 1)}{1-\rho^2}$	0	$\frac{\sqrt{2}}{1-\rho^2}$
$H_3(y_2)$	0	$-\frac{\sqrt{6}\rho}{1-\rho^2}$	0	0
$H_4(y_1)$	$\frac{\sqrt{\frac{3}{2}}\rho^2}{(1-\rho^2)^2}$	0	$\frac{\sqrt{\frac{3}{2}}}{(1-\rho^2)^2}$	0
$H_3(y_1)H_1(y_2)$	$-\frac{\sqrt{\frac{3}{2}}\rho(\rho^2 + 3)}{2(1-\rho^2)^2}$	0	$-\frac{\sqrt{6}\rho}{(1-\rho^2)^2}$	0
$H_2(y_1)H_2(y_2)$	$\frac{2\rho^2 + 1}{(1-\rho^2)^2}$	0	$\frac{2\rho^2 + 1}{(1-\rho^2)^2}$	0
$H_1(y_1)H_3(y_2)$	$-\frac{\sqrt{\frac{3}{2}}\rho(\rho^2 + 3)}{2(1-\rho^2)^2}$	0	$-\frac{\sqrt{6}\rho}{(1-\rho^2)^2}$	0
$H_4(y_2)$	$\frac{\sqrt{\frac{3}{2}}\rho^2}{(1-\rho^2)^2}$	0	$\frac{\sqrt{\frac{3}{2}}}{(1-\rho^2)^2}$	0

Similarly, doing the same calculations for $\varphi = b_i$ we have that

$$V \left[\begin{pmatrix} m_{b_1}(y_1, y_2; \rho) \\ m_{b_2}(y_1, y_2; \rho) \end{pmatrix} \right] = \begin{bmatrix} 2 & 2\rho^3 \\ 2\rho^3 & 2 \end{bmatrix}, \quad V \left[\begin{pmatrix} d_{b_1}(y_1, y_2; \rho) \\ d_{b_2}(y_1, y_2; \rho) \end{pmatrix} \right] = \begin{bmatrix} 8 & 8\rho \\ 8\rho & 8 \end{bmatrix}$$

and

$$\text{cov} \left[\begin{pmatrix} m_{b_1}(y_1, y_2; \rho) \\ m_{b_2}(y_1, y_2; \rho) \end{pmatrix}, \begin{pmatrix} d_{b_1}(y_1, y_2; \rho) \\ d_{b_2}(y_1, y_2; \rho) \end{pmatrix} \right] = \begin{bmatrix} 2 - 4\rho^2 & -2\rho \\ -2\rho & 2 - 4\rho^2 \end{bmatrix}.$$

In this way, it is clear that for b_1 ,

$$d_{b_1}(y_1, y_2; \rho) = \frac{1 - \rho^2}{1 + \rho^2 + \rho^4} m_{b_1}(y_1, y_2; \rho) - \frac{\rho + 2\rho^3}{1 + \rho^2 + \rho^4} m_{b_2}(y_1, y_2; \rho) + u_{b_1},$$

so that the reduction in power of the distributional LM test relative to the corresponding score test is characterized by

$$V(u_{b_i}) = \frac{6(1 + 2\rho^2)}{1 + \rho^2 + \rho^4}$$

because $V[d_\eta(y_1, y_2; \rho)] = 4$.

D.3.2 Accounting for marginals uncertainty

Direct application of Proposition 8 yields

$$n_\eta(y_1, y_2; \rho) = \frac{1}{4} \sqrt{\frac{3}{2}} \rho^2 [H_4(y_1) + H_4(y_2)], \quad n_{b_i}(y_1, y_2; \rho) = \sqrt{\frac{2}{3}} \rho [\rho H_3(y_1) + H_3(y_2)]$$

and $n_{b_2}(y_1, y_2; \rho) = n_{b_1}(y_2, y_1; \rho)$. Analogous calculations for the distributional test moments deliver

$$n_\eta^d(y_1, y_2; \rho) = \sqrt{\frac{3}{2}} [H_4(y_1) + H_4(y_2)], \quad n_{b_1}(y_1, y_2; \rho) = \sqrt{6} H_3(y_1) + 2\rho \sqrt{\frac{2}{3}} H_3(y_2),$$

and $n_{b_2}(y_1, y_2; \rho) = n_{b_1}(y_2, y_1; \rho)$. In the following table we summarize the modified moments that account for non parametric estimation of the marginals.

Table D2: Hermite polynomial coefficients for bivariate score copula tests and distributional LM tests when marginals are estimated nonparametrically

Hermite polynomial	Copula LM test		Distributional LM test	
	$s_\eta^{np}(y_1, y_2; \rho)$	$m_{b_1}^{np}(y_1, y_2; \rho)$	$d_\eta^{np}(y_1, y_2; \rho)$	$d_{b_1}^{np}(y_1, y_2; \rho)$
1	$\frac{2\rho^4 + \rho^2}{(1-\rho^2)^2}$	0	$\frac{2\rho^4 + \rho^2}{(1-\rho^2)^2}$	0
$H_1(y_1)$	0	$\frac{2\rho^2}{1-\rho^2}$	0	$\frac{4\rho^2}{1-\rho^2}$
$H_1(y_2)$	0	$-\frac{2(\rho^3 + \rho)}{1-\rho^2}$	0	$-\frac{2\rho}{1-\rho^2}$
$H_2(y_1)$	$\frac{3\rho^2(\rho^2 + 3)}{2\sqrt{2}(1-\rho^2)^2}$	0	$\frac{3\sqrt{2}\rho^2}{(1-\rho^2)^2}$	0
$H_1(y_1)H_1(y_2)$	$-\frac{2(2\rho^3 + \rho)}{(1-\rho^2)^2}$	0	$-\frac{2(2\rho^3 + \rho)}{(1-\rho^2)^2}$	0
$H_2(y_2)$	$\frac{3\rho^2(\rho^2 + 3)}{2\sqrt{2}(1-\rho^2)^2}$	0	$\frac{3\sqrt{2}\rho^2}{(1-\rho^2)^2}$	0
$H_3(y_1)$	0	$\sqrt{\frac{2}{3}}\rho^2$	0	$\frac{\sqrt{6}\rho^2}{1-\rho^2}$
$H_2(y_1)H_1(y_2)$	0	$-\frac{\sqrt{2}\rho}{1-\rho^2}$	0	$-\frac{2\sqrt{2}\rho}{1-\rho^2}$
$H_1(y_1)H_2(y_2)$	0	$\frac{\sqrt{2}(\rho^2 + 1)}{1-\rho^2}$	0	$\frac{\sqrt{2}}{1-\rho^2}$
$H_3(y_2)$	0	$-\frac{\sqrt{\frac{2}{3}}\rho(\rho^2 + 2)}{1-\rho^2}$	0	$-2\sqrt{\frac{2}{3}}\rho$
$H_4(y_1)$	$\frac{\sqrt{\frac{3}{2}}\rho^2(\rho^4 - 2\rho^2 + 5)}{4(1-\rho^2)^2}$	0	$-\frac{\sqrt{\frac{3}{2}}\rho^2(\rho^2 - 2)}{(1-\rho^2)^2}$	0
$H_3(y_1)H_1(y_2)$	$-\frac{\sqrt{\frac{3}{2}}\rho(\rho^2 + 3)}{2(1-\rho^2)^2}$	0	$-\frac{\sqrt{6}\rho}{(1-\rho^2)^2}$	0
$H_2(y_1)H_2(y_2)$	$\frac{2\rho^2 + 1}{(1-\rho^2)^2}$	0	$\frac{2\rho^2 + 1}{(1-\rho^2)^2}$	0
$H_1(y_1)H_3(y_2)$	$-\frac{\sqrt{\frac{3}{2}}\rho(\rho^2 + 3)}{2(1-\rho^2)^2}$	0	$-\frac{\sqrt{6}\rho}{(1-\rho^2)^2}$	0
$H_4(y_2)$	$\frac{\sqrt{\frac{3}{2}}\rho^2(\rho^4 - 2\rho^2 + 5)}{4(1-\rho^2)^2}$	0	$-\frac{\sqrt{\frac{3}{2}}\rho^2(\rho^2 - 2)}{(1-\rho^2)^2}$	0

Again, in order to characterize the loss of power of the distributional version of the test we could write

$$d_\varphi^{np}(y_1, y_2; \rho) = \beta_\varphi^{np} s_\varphi^{np}(y_1, y_2; \rho) + u_\varphi^{np}$$

where

$$\beta_{\varphi}^{np} = \frac{\text{cov}[d_{\varphi}^{np}(y_1, y_2; \rho), s_{\varphi}(y_1, y_2; \rho)]}{\text{cov}[s_{\varphi}^{np}(y_1, y_2; \rho), s_{\varphi}(y_1, y_2; \rho)]}$$

so that the non-centrality parameter of $d_{\varphi}(y_1, y_2; \rho)$ under a sequence of local alternatives H_l :

$\varphi_T = \bar{\varphi}/\sqrt{T}$ can be written as

$$\frac{\beta_{\varphi}^2 \text{cov}[s_{\varphi}^{np}(y_1, y_2; \rho), s_{\varphi}(y_1, y_2; \rho)]}{\beta_{\varphi}^2 \text{cov}[s_{\varphi}^{np}(y_1, y_2; \rho), s_{\varphi}(y_1, y_2; \rho)] + V(u_{\varphi}^{np})}$$

since $\text{cov}[s_{\varphi}^{np}(y_1, y_2; \rho), u_{\varphi}^{np}] = 0$. For instance, when $\varphi = \eta$ we have that

$$\text{cov}[s_{\eta}^{np}(y_1, y_2; \rho), s_{\eta}(y_1, y_2; \rho)] = 1 + \frac{3}{4}\rho^2$$

and

$$\text{cov}[d_{\eta}^{np}(y_1, y_2; \rho), s_{\eta}(y_1, y_2; \rho)] = 1,$$

so that the reduction in power of the distributional LM test relative to the score test is

$$V(u_{\eta}) = \frac{12(1 + \rho^2)(\rho + 2\rho^3)^2}{(4 + 3\rho^2)^2},$$

where we have used the fact that $V[d_{\eta}^{np}(y_1, y_2; \rho)] = 1 + 3\rho^4$.

E Computational details

E.1 Monte Carlo details

The Monte Carlo study about the properties of the test when the critical values are bootstrapped is divided in three main blocks:

1. Estimation of the correlation parameters.
2. Construction of the table with critical values.
3. Evaluation of the test size and power.

E.1.1 Estimation of the correlation parameters

To obtain the range of parameter values from which we calculate the table of critical values utilized for the Monte Carlo evaluation of the tests, the steps are the following:

1. Fix seed s_1 .
2. Simulate data from the joint distribution F .

3. Compute the Gaussian ranks of the simulated data.
4. Estimate the parameter vector $\boldsymbol{\rho}$ by ML using the Gaussian rank correlations as initial values.

Steps 2 and 3 are repeated 10,000 times, saving the parameter estimates of each iteration ($\hat{\boldsymbol{\rho}}$'s). After this block has finished we save the estimates of the 10,000 simulations. In block 3 we will use the same seed, thus we will not need to estimate again.

E.1.2 Construction of the table with critical values

To obtain the distribution of the test as a function of the $\boldsymbol{\rho}$'s estimated in the previous section, the steps of the code are the following:

1. Load the estimated $\boldsymbol{\rho}$'s and create a grid of H points, $\mathcal{H} = \{\boldsymbol{\rho}^{(1)}, \dots, \boldsymbol{\rho}^{(h)}, \dots, \boldsymbol{\rho}^{(H)}\}$, that covers all $\hat{\boldsymbol{\rho}}$'s: Φ , say.
2. Fix seed $s_2 \neq s_1$.
3. Simulate an $N \times H$ matrix of independent Gaussian random numbers ϵ .
4. For each point $h = \{1, \dots, H\}$:
 - (a) Simulate Y from a Gaussian copula: $Y = \epsilon \mathbf{P}^{1/2}(\boldsymbol{\rho}^{(h)})$ where $\mathbf{P}^{1/2}(\boldsymbol{\rho})$ is the Cholesky decomposition of $\mathbf{P}(\boldsymbol{\rho})$.
 - (b) Compute the Gaussian ranks for the columns of Y .
 - (c) Estimate the correlation parameters by ML using the Gaussian rank correlations as initial values.
 - (d) Compute the tests evaluated at the parameter estimates in step c: $Test(s; h)$, say.

Steps 3 to 5 are repeated 10,000 times, the test statistics for any $\boldsymbol{\rho}$ in Φ of each iteration are saved.

E.1.3 Evaluation of the test size and power

To obtain the size or power of the tests, the steps of the code are the following:

1. Load the results of sections E.1.1 and E.1.2.

2. For each test, compute the relevant $(1 - \alpha)$ quantiles of the $Test(s, h)$ for each h : Q^α .
3. Fix seed s_1 .
4. Simulate data from the relevant joint distribution F (Gaussian for size).
5. Compute the Gaussian ranks of the data.
6. Compute the test evaluated at $\hat{\rho}$: $Test(s)$, say.
7. Find the critical value (c^α) of the test at significance level α through a linear interpolation of the quantiles.

Steps 4 to 7 are repeated 10,000 times and the number of times $Test(s) > c^\alpha$ is recorded for each test to compute size and power.

E.2 Pooled estimation and testing

Recall that under the null

$$\mathbf{s}_\rho(\rho, \mathbf{0}) = \frac{\partial \ln f_K [F_1^{-1}(u_1; \mathbf{0}), \dots, F_1^{-1}(u_K; \mathbf{0}); \rho, \mathbf{0}]}{\partial \rho} = -\mathbf{Z}_s(\rho) \mathbf{e}_s(\mathbf{0})$$

where

$$\mathbf{Z}_s(\rho) \mathbf{e}_s(\phi, \mathbf{0}) = \frac{1}{2} \frac{\partial vec'[\mathbf{P}(\rho)]}{\partial \rho} \cdot [\mathbf{P}^{-1}(\rho) \otimes \mathbf{P}^{-1}(\rho)] vec \{ \mathbf{y}_t \mathbf{y}_t' - \mathbf{P}(\rho) \}.$$

For a given cross-section, we have $Y_t = \{(y_{11}^t, y_{21}^t), \dots, (y_{1n}^t, y_{2n}^t), \dots, (y_{1N_t}^t, y_{2N_t}^t)\}$. The full sample would then consist of $\sum_{t=1}^T N_t$ bivariate observations $\mathbf{Y} = \{Y_1, \dots, Y_T\}$. At each t , we can compute the average score,

$$\bar{\mathbf{s}}_{\phi t}(Y_t; \rho) = \frac{1}{N_t} \sum_{n=1}^{N_t} \begin{pmatrix} s_\rho(Y_t; \rho) \\ \mathbf{s}_\varphi(Y_t; \rho) \end{pmatrix},$$

which is the basis for the pooled average score $\bar{\mathbf{s}}_\phi(Y_t; \rho) = T^{-1} \sum_{t=1}^T \bar{\mathbf{s}}_{\phi t}(Y_t; \rho)$. As for Spearman correlation coefficient, we can simplify our calculations by noticing that for large N , $\sum_{n=1}^{N_t} \Phi(y_{in}) \approx 1/2$ and $\sum_{n=1}^{N_t} \Phi^2(y_{in}) \approx 1/3$ so that

$$\frac{\sqrt{N_t} \sum_{n=1}^{N_t} \Phi(y_{1n}) \Phi(y_{2n}) - 1/4}{N_t \cdot 1/12}$$

is the relevant moment function required to compute robust standard errors. Finally, to estimate Pearson's correlation and its corresponding robust standard error we can consider the following moment functions

$$\mathbf{m}(X_t) = \frac{1}{N_t} \sum_{n=1}^{N_t} [x_{1n}^t, x_{2n}^t, (x_{1n}^t)^2, (x_{2n}^t)^2, x_{1n}^t x_{2n}^t]'$$

as our basis for computing the covariance matrix of the sample average of the moments through a standard HAC estimator. Next, we introduce $g : \mathbb{R}^5 \rightarrow \mathbb{R}^3$,

$$\mathbf{g}[\mathbf{m}(X_t)] = \begin{bmatrix} \mathbf{m}_3(X_t) - \mathbf{m}_1^2(X_t) \\ \mathbf{m}_4(X_t) - \mathbf{m}_2^2(X_t) \\ \mathbf{m}_5(X_t) - \mathbf{m}_1(X_t)\mathbf{m}_2(X_t) \end{bmatrix} \text{ so that } \frac{\partial \mathbf{g}}{\partial \mathbf{m}} = \begin{bmatrix} -2m_1 & 0 & 1 & 0 & 0 \\ 0 & -2m_2 & 0 & 1 & 0 \\ -m_2 & -m_1 & 0 & 0 & 1 \end{bmatrix}$$

and $h : \mathbb{R}^3 \rightarrow [-1, 1]$,

$$h\{\mathbf{g}[\mathbf{m}(X_t)]\} = \frac{g_3}{\sqrt{g_1 g_2}} \text{ so that } \frac{\partial h}{\partial \mathbf{g}} = \left[\frac{-g_3}{2g_1 \sqrt{g_1 g_2}}, \frac{-g_3}{2g_2 \sqrt{g_1 g_2}}, \frac{1}{\sqrt{g_1 g_2}} \right]$$

and apply the Delta method twice to obtain the asymptotic variance of ρ .

E.3 Variances of the moment functions

E.3.1 The bivariate case: Known marginals

Generalized Hyperbolic The variances are

$$V[s_\eta(\mathbf{y}, \rho)] = 1 + \frac{3}{4}\rho^2,$$

and

$$V[m_{b_k}(\mathbf{y}, \rho)] = 2, \quad \text{for } k = 1, 2,$$

while the covariances are

$$\text{cov}[m_{b_1}(\mathbf{y}, \rho), m_{b_2}(\mathbf{y}, \rho)] = 2\rho^3$$

and

$$\text{cov}[s_\eta(\mathbf{y}, \rho), m_{b_k}(\mathbf{y}, \rho)] = 0, \quad \text{for } k = 1, 2.$$

Hermite expansion The variances are

$$\begin{aligned} V[s_{\tilde{c}_1}(\mathbf{y}, \rho)] &= V[s_{\tilde{c}_4}(\mathbf{y}, \rho)] \\ &= \frac{2}{9}\rho^2 (110 - 363\rho^2 + 408\rho^4 - 151\rho^6 - 8\rho^8 + 8\rho^{10}), \end{aligned}$$

$$V[s_{c_2}(\mathbf{y}, \rho)] = V[s_{c_3}(\mathbf{y}, \rho)] = 1 + \rho^2 + 2\rho^4,$$

$$V[s_{d_1}(\mathbf{y}, \rho)] = \frac{1}{4}\rho^2 (1 + \rho^2 + \rho^4 + 3\rho^6),$$

$$\begin{aligned} V[s_{d_2}(\mathbf{y}, \rho)] &= V[s_{d_4}(\mathbf{y}, \rho)] \\ &= 1 + \rho^2 + \frac{5}{2}\rho^4 + \frac{9}{2}\rho^6, \end{aligned}$$

$$V[s_{d_3}(\mathbf{y}, \rho)] = 1 + 8\rho^2 + 5\rho^4 + 3\rho^6,$$

and

$$\begin{aligned} V[s_{d_5}(\mathbf{y}, \rho)] &= \frac{1}{4} (1 - \rho^2)^2 (16 + 1053\rho^2 - 1148\rho^4 + 1080\rho^6 - 1791\rho^8) \\ &\quad + \frac{1}{16} (1 - \rho^2)^2 (6677\rho^{10} - 3250\rho^{12} + 801\rho^{14} + 108\rho^{16}). \end{aligned}$$

As for the covariances, they are

$$\text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{c_2}(\mathbf{y}, \rho)] = \text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{c_3}(\mathbf{y}, \rho)] = 0,$$

$$\text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{\tilde{c}_4}(\mathbf{y}, \rho)] = \frac{2}{9}\rho^2 (2 + 50\rho^2 - 201\rho^4 + 244\rho^6 - 17\rho^8 - 168\rho^{10} + 94\rho^{12}),$$

$$\text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{d_1}(\mathbf{y}, \rho)] = \sqrt{\frac{2}{3}}\rho^2 (2 - 3\rho^2 + \rho^8),$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{d_2}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}(\mathbf{y}, \rho), s_{d_4}(\mathbf{y}, \rho)] \\ &= \sqrt{\frac{2}{3}}\rho (4 - 5\rho^2 - \rho^4 + 2\rho^8), \end{aligned}$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{d_3}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}(\mathbf{y}, \rho), s_{d_3}(\mathbf{y}, \rho)] \\ &= \frac{2}{3}\rho^2 (12 - 22\rho^2 + 7\rho^4 + 3\rho^6), \end{aligned}$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{d_4}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}(\mathbf{y}, \rho), s_{d_2}(\mathbf{y}, \rho)] \\ &= \sqrt{\frac{2}{3}}\rho^3 (9 - 19\rho^2 + 10\rho^4), \end{aligned}$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] \\ &= \frac{1}{12}\rho^2 (1 - \rho^2)^2 \\ &\quad \times (512 - 590\rho^2 + 344\rho^4 - 761\rho^6 + 759\rho^8 - 210\rho^{10}), \end{aligned}$$

$$\text{cov}[s_{c_2}(\mathbf{y}, \rho), s_{c_3}(\mathbf{y}, \rho)] = \rho(2 + \rho^2 + \rho^4),$$

$$\begin{aligned} \text{cov}[s_{c_j}(\mathbf{y}, \rho), s_{\tilde{c}_4}(\mathbf{y}, \rho)] &= \text{cov}[s_{c_j}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] \\ &= \text{cov}[s_{c_j}(\mathbf{y}, \rho), s_{d_i}(\mathbf{y}, \rho)] \\ &= 0, \end{aligned}$$

for $i = 1, \dots, 4$ and $j = 2, 3$,

$$\text{cov}[s_{\tilde{c}_4}(\mathbf{y}, \rho), s_{d_1}(\mathbf{y}, \rho)] = \sqrt{\frac{2}{3}}\rho^4 (4 - 9\rho^2 + 5\rho^4),$$

$$\begin{aligned}
\text{cov} [s_{d_1}(\mathbf{y}, \rho), s_{d_2}(\mathbf{y}, \rho)] &= \frac{1}{2}\rho (1 + \rho^2 + \rho^4 + 3\rho^6), \\
\text{cov} [s_{d_1}(\mathbf{y}, \rho), s_{d_3}(\mathbf{y}, \rho)] &= \frac{1}{2}\sqrt{\frac{3}{2}}\rho^2(2 + \rho^2)(1 + \rho^4), \\
\text{cov} [s_{d_1}(\mathbf{y}, \rho), s_{d_4}(\mathbf{y}, \rho)] &= \frac{1}{2}\rho^3(3 + \rho^2 + \rho^4 + \rho^6), \\
\text{cov} [s_{d_1}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] &= \frac{3}{4}\sqrt{\frac{3}{2}}\rho^2(2 + \rho^2 - 8\rho^4 + 10\rho^6 - 6\rho^8 + \rho^{10}), \\
\text{cov} [s_{d_2}(\mathbf{y}, \rho), s_{d_3}(\mathbf{y}, \rho)] &= \text{cov} [s_{d_3}(\mathbf{y}, \rho), s_{d_4}(\mathbf{y}, \rho)] \\
&= \sqrt{\frac{3}{2}}\rho(2 + 3\rho^2 + 4\rho^4 + \rho^6), \\
\text{cov} [s_{d_2}(\mathbf{y}, \rho), s_{d_4}(\mathbf{y}, \rho)] &= \frac{1}{2}\rho^2(2 + \rho^2)(3 + \rho^2 + 2\rho^4), \\
\text{cov} [s_{d_2}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] &= \text{cov} [s_{d_4}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] \\
&= \frac{1}{4}\sqrt{\frac{3}{2}}\rho^2(1 - \rho^2)(1 + 3\rho^2)(12 - 14\rho^2 + 14\rho^4 - \rho^6),
\end{aligned}$$

and

$$\text{cov} [s_{d_3}(\mathbf{y}, \rho), s_{\tilde{d}_5}(\mathbf{y}, \rho)] = \frac{1}{2}\rho^2(1 - \rho^2)(40 - 49\rho^2 + 53\rho^4 - 6\rho^6).$$

E.3.2 The bivariate case: Accounting for non-parametric estimation of the marginals

Generalized Hyperbolic

The variances are

$$V[s_\eta^{np}(\mathbf{y}, \rho)] = 1 + \frac{3}{4}\rho^2 + \frac{3}{16}(\rho^4 + \rho^8),$$

and

$$V[m_{b_k}^{np}(\mathbf{y}, \rho)] = 2 + \frac{2}{3}(\rho^2 + \rho^4 + 2\rho^6), \quad \text{for } k = 1, 2,$$

while the covariances are

$$\text{cov} [m_{b_1}^{np}(\mathbf{y}, \rho), m_{b_2}^{np}(\mathbf{y}, \rho)] = 2\rho^3 + \frac{2}{3}\rho^3(2 + \rho^2 + \rho^4),$$

and

$$\text{cov} [s_\eta^{np}(\mathbf{y}, \rho), m_{b_k}^{np}(\mathbf{y}, \rho)] = 0, \quad \text{for } k = 1, 2.$$

Hermite expansion

The variances are

$$\begin{aligned} V[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho)] &= V[s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho)] \\ &= \frac{2}{9}\rho^2(1-\rho^2)^2[110 + 3\rho^2(10\rho^6 + \rho^4 - 40)], \end{aligned}$$

$$\begin{aligned} V[s_{\tilde{c}_2}^{np}(\mathbf{y}, \rho)] &= V[s_{\tilde{c}_3}^{np}(\mathbf{y}, \rho)] \\ &= 1 + \frac{1}{3}(4\rho^2 + 12\rho^4 + 5\rho^6 + 4\rho^8), \end{aligned}$$

$$V[s_{d_1}^{np}(\mathbf{y}, \rho)] = \frac{1}{16}\rho^2(4 + 5\rho^2 + 6\rho^4 + 24\rho^6 + 8\rho^8 + 7\rho^{10} + 6\rho^{12}),$$

$$\begin{aligned} V[s_{d_2}^{np}(\mathbf{y}, \rho)] &= V[s_{d_4}^{np}(\mathbf{y}, \rho)] \\ &= 1 + \frac{1}{4}(5\rho^2 + 12\rho^4 + 36\rho^6 + 14\rho^8 + 7\rho^{10} + 6\rho^{12}), \end{aligned}$$

$$V[s_{d_3}^{np}(\mathbf{y}, \rho)] = 1 + 8\rho^2 + 12\rho^4 + 10\rho^6 + \frac{15}{4}\rho^8 + 3\rho^{10} + \frac{3}{4}\rho^{12},$$

and

$$\begin{aligned} V[s_{d_5}^{np}(\mathbf{y}, \rho)] &= \frac{1}{16}(1-\rho^2)^2(64 + 4212\rho^2 - 3972\rho^4 + 4936\rho^6 - 5877\rho^8) \\ &\quad + \frac{1}{16}(1-\rho^2)^2(5719\rho^{10} - 3076\rho^{12} + 1299\rho^{14} - 313\rho^{16} + 40\rho^{18}), \end{aligned}$$

As for the covariances, they are

$$\text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{c}_2}^{np}(\mathbf{y}, \rho)] = \text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{c}_3}^{np}(\mathbf{y}, \rho)] = 0,$$

$$\text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho)] = \frac{2}{9}\rho^2(1-\rho^2)^2(2 + 56\rho^2 - 48\rho^4 + 36\rho^6 + 46\rho^8 + 3\rho^{10}),$$

$$\text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{d_1}^{np}(\mathbf{y}, \rho)] = \frac{1}{2\sqrt{6}}\rho^2(1-\rho^2)(8 - 2\rho^2 - 3\rho^4 + 8\rho^6 - 18\rho^8 - 2\rho^{10} - \rho^{12}),$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{d_2}^{np}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho), s_{d_4}^{np}(\mathbf{y}, \rho)] \\ &= \frac{1}{\sqrt{6}}\rho(1-\rho^2)(8 - \rho^4 + 8\rho^6 - 18\rho^8 - 2\rho^{10} - \rho^{12}), \end{aligned}$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{d_3}^{np}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho), s_{d_3}^{np}(\mathbf{y}, \rho)] \\ &= \frac{1}{6}\rho^2(1-\rho^2)\{48 - \rho^2[20 + 3\rho^2(-4 + 9\rho^2 + 8\rho^4 + \rho^6)]\}, \end{aligned}$$

$$\begin{aligned} \text{cov}[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{d_4}^{np}(\mathbf{y}, \rho)] &= \text{cov}[s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho), s_{d_2}^{np}(\mathbf{y}, \rho)] \\ &= \sqrt{\frac{2}{3}}\rho^3(9 - 13\rho^2 + 2\rho^4 - 2\rho^8 + 4\rho^{10}), \end{aligned}$$

$$\begin{aligned}
\text{cov} \left[s_{\tilde{c}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] &= \text{cov} \left[s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] \\
&= \frac{1}{12} \rho^2 (1 - \rho^2)^2 \\
&\quad \times (512 - 488\rho^2 + 454\rho^4 - 707\rho^6 + 579\rho^8 - 291\rho^{10} + 49\rho^{12}),
\end{aligned}$$

$$\text{cov} \left[s_{\tilde{c}_2}^{np}(\mathbf{y}, \rho), s_{\tilde{c}_3}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{3} \rho (6 + 7\rho^2 + 8\rho^4 + 6\rho^6 + \rho^8),$$

$$\text{cov} \left[s_{\tilde{c}_j}^{np}(\mathbf{y}, \rho), s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho) \right] = \text{cov} \left[s_{\tilde{c}_j}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] = \text{cov} \left[s_{\tilde{c}_j}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_i}^{np}(\mathbf{y}, \rho) \right] = 0,$$

for $i = 1, \dots, 4$ and $j = 2, 3$,

$$\text{cov} \left[s_{\tilde{c}_4}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_1}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{\sqrt{6}} \rho^4 (8 - 13\rho^2 + 3\rho^4 - 2\rho^8 + 4\rho^{10}),$$

$$\text{cov} \left[s_{\tilde{d}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_2}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{8} \rho (4 + 5\rho^2 + 6\rho^4 + 24\rho^6 + 8\rho^8 + 7\rho^{10} + 6\rho^{12}),$$

$$\text{cov} \left[s_{\tilde{d}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_3}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{8} \sqrt{\frac{3}{2}} (\rho + \rho^5)^2 (8 + 6\rho^2 + \rho^4),$$

$$\text{cov} \left[s_{\tilde{d}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_4}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{8} \rho^3 [12 + (1 + \rho^2)(1 + \rho^4)(10 + \rho^2 + \rho^4)\rho^2],$$

$$\text{cov} \left[s_{\tilde{d}_1}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] = \frac{3}{16} \sqrt{\frac{3}{2}} \rho^2 (1 - \rho^2)(1 + \rho^4)(4 + \rho^2)(2 + 3\rho^2 + 5\rho^4 + 5\rho^6 - \rho^8),$$

$$\begin{aligned}
\text{cov} \left[s_{\tilde{d}_2}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_3}^{np}(\mathbf{y}, \rho) \right] &= \text{cov} \left[s_{\tilde{d}_3}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_4}^{np}(\mathbf{y}, \rho) \right] \\
&= \frac{1}{4} \sqrt{\frac{3}{2}} \rho [8 + \rho^2(7 + 6\rho^2 + \rho^4)(2 + 3\rho^2 + \rho^6)],
\end{aligned}$$

$$\text{cov} \left[s_{\tilde{d}_2}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_4}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{4} \rho^2 (1 + \rho^4)(12 + 16\rho^2 + 11\rho^4 + 2\rho^6 + \rho^8),$$

$$\begin{aligned}
\text{cov} \left[s_{\tilde{d}_2}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] &= \text{cov} \left[s_{\tilde{d}_4}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] \\
&= \frac{1}{8} \sqrt{\frac{3}{2}} \rho [24 + \rho^2 \{-65 + \rho^2 [128 + \rho^2 (-133 + 68\rho^2)]\}] \\
&\quad + \frac{1}{8} \sqrt{\frac{3}{2}} \rho^9 [3\rho^2 (\rho^2 - 5)(3 + \rho^2)]
\end{aligned}$$

and

$$\text{cov} \left[s_{\tilde{d}_3}^{np}(\mathbf{y}, \rho), s_{\tilde{d}_5}^{np}(\mathbf{y}, \rho) \right] = \frac{1}{8} \rho^2 (1 - \rho^2) (160 - 144\rho^2 + 292\rho^4 - 43\rho^6 + 57\rho^8 + 27\rho^{10} - 9\rho^{12}).$$

E.3.3 The trivariate case with known marginals

Generalized Hyperbolic

The variances are

$$V[s_\eta(y_1, y_2, y_3, \boldsymbol{\rho})] = \frac{3(\rho_{12}^4 - \rho_{13}\rho_{23}\rho_{12}^3 + 3\rho_{12}^2 - \rho_{13}\rho_{23}(\rho_{13}^2 + \rho_{23}^2 + 5)\rho_{12})}{4(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1)} + \frac{3(\rho_{13}^4 + \rho_{23}^4 + 3\rho_{13}^2 + 3\rho_{23}^2 - 4)}{4(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1)},$$

and

$$V[m_{b_i}(y_1, y_2, y_3, \boldsymbol{\rho})] = \frac{4(\rho_{13}\rho_{23}\rho_{12}^3 - ((\rho_{23}^2 + 1)\rho_{13}^2 + \rho_{23}^2 - 1)\rho_{12}^2)}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1} + \frac{4(\rho_{13}\rho_{23}(\rho_{13}^2 + \rho_{23}^2 - 1)\rho_{12} - (\rho_{13}^2 - 1)(\rho_{23}^2 - 1))}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1},$$

for $i = 1, 2, 3$. In turn, the covariances are

$$\text{cov}[m_{b_1}(y_1, y_2, y_3, \boldsymbol{\rho}), m_{b_2}(y_1, y_2, y_3, \boldsymbol{\rho})] = \frac{2(\rho_{12}^5 - \rho_{13}\rho_{23}\rho_{12}^4 - \rho_{13}\rho_{23}(\rho_{13}^2 + \rho_{23}^2)\rho_{12}^2)}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1} + \frac{2(((\rho_{13}^2 + \rho_{23}^2)^2 - 1)\rho_{12} - \rho_{13}\rho_{23}(\rho_{13}^2 + \rho_{23}^2 - 1))}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1},$$

$$\text{cov}[m_{b_1}(y_1, y_2, y_3, \boldsymbol{\rho}), m_{b_3}(y_1, y_2, y_3, \boldsymbol{\rho})] = \frac{2(\rho_{13}^5 - \rho_{12}\rho_{23}\rho_{13}^4 - \rho_{12}\rho_{23}(\rho_{12}^2 + \rho_{23}^2)\rho_{13}^2)}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1} + \frac{2(((\rho_{12}^2 + \rho_{23}^2)^2 - 1)\rho_{13} - \rho_{12}\rho_{23}(\rho_{12}^2 + \rho_{23}^2 - 1))}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1},$$

$$\text{cov}[m_{b_2}(y_1, y_2, y_3, \boldsymbol{\rho}), m_{b_3}(y_1, y_2, y_3, \boldsymbol{\rho})] = \frac{2(\rho_{23}^5 - \rho_{12}\rho_{13}\rho_{23}^4 - \rho_{12}\rho_{13}(\rho_{12}^2 + \rho_{13}^2)\rho_{23}^2)}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1} + \frac{2(((\rho_{12}^2 + \rho_{13}^2)^2 - 1)\rho_{23} - \rho_{12}\rho_{13}(\rho_{12}^2 + \rho_{13}^2 - 1))}{\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1},$$

and

$$\text{cov}[s_\eta(y_1, y_2, y_3, \boldsymbol{\rho}), m_{b_k}(y_1, y_2, y_3, \boldsymbol{\rho})] = 0, \quad \text{for } k = 1, 2, 3.$$

E.3.4 The trivariate case: Accounting for non-parametric estimation of the marginals

Generalized Hyperbolic

The variances are

$$\begin{aligned}
V[s_{\eta}^{np}(y_1, y_2, y_3, \boldsymbol{\rho})] = & \frac{1}{16(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1)^2} \\
& \times (3\rho_{12}^{12} - 6\rho_{13}\rho_{23}\rho_{12}^{11} + 3\rho_{13}^2\rho_{23}^2\rho_{12}^{10} - 6\rho_{12}^{10} - 3\rho_{13}\rho_{23}^3\rho_{12}^9 - 3\rho_{13}^3\rho_{23}\rho_{12}^9 \\
& + 18\rho_{13}\rho_{23}\rho_{12}^9 + 3\rho_{13}^4\rho_{12}^8 + 3\rho_{13}^2\rho_{23}^4\rho_{12}^8 + 3\rho_{23}^4\rho_{12}^8 - 3\rho_{13}^2\rho_{12}^8 + 3\rho_{13}^4\rho_{23}^2\rho_{12}^8 \\
& - 12\rho_{13}^2\rho_{23}^2\rho_{12}^8 - 3\rho_{23}^2\rho_{12}^8 + 6\rho_{12}^8 - 3\rho_{13}\rho_{23}^5\rho_{12}^7 + 6\rho_{13}\rho_{23}^3\rho_{12}^7 - 3\rho_{13}^5\rho_{23}\rho_{12}^7 \\
& + 6\rho_{13}^3\rho_{23}\rho_{12}^7 - 18\rho_{13}\rho_{23}\rho_{12}^7 - 3\rho_{13}^4\rho_{12}^6 + 3\rho_{13}^4\rho_{23}^4\rho_{12}^6 - 6\rho_{13}^2\rho_{23}^4\rho_{12}^6 - 3\rho_{23}^4\rho_{12}^6 \\
& + 3\rho_{13}^2\rho_{12}^6 - 6\rho_{13}^4\rho_{23}^2\rho_{12}^6 + 15\rho_{13}^2\rho_{23}^2\rho_{12}^6 + 3\rho_{23}^2\rho_{12}^6 + 6\rho_{12}^6 - 3\rho_{13}\rho_{23}^7\rho_{12}^5 \\
& - 6\rho_{13}^3\rho_{23}^5\rho_{12}^5 + 12\rho_{13}\rho_{23}^5\rho_{12}^5 - 6\rho_{13}^5\rho_{23}^3\rho_{12}^5 + 6\rho_{13}^3\rho_{23}^3\rho_{12}^5 - 9\rho_{13}\rho_{23}^3\rho_{12}^5 \\
& - 3\rho_{13}^7\rho_{23}\rho_{12}^5 + 12\rho_{13}^5\rho_{23}\rho_{12}^5 - 9\rho_{13}^3\rho_{23}\rho_{12}^5 - 18\rho_{13}\rho_{23}\rho_{12}^5 + 3\rho_{13}^8\rho_{12}^4 \\
& + 3\rho_{13}^2\rho_{23}^8\rho_{12}^4 + 3\rho_{23}^8\rho_{12}^4 - 3\rho_{13}^6\rho_{12}^4 + 3\rho_{13}^4\rho_{23}^6\rho_{12}^4 - 6\rho_{13}^2\rho_{23}^6\rho_{12}^4 - 3\rho_{23}^6\rho_{12}^4 \\
& + 3\rho_{13}^4\rho_{12}^4 + 3\rho_{13}^6\rho_{23}^4\rho_{12}^4 + 9\rho_{13}^4\rho_{23}^4\rho_{12}^4 - 3\rho_{13}^2\rho_{23}^4\rho_{12}^4 + 3\rho_{23}^4\rho_{12}^4 + 9\rho_{13}^2\rho_{12}^4 \\
& + 3\rho_{13}^8\rho_{23}^2\rho_{12}^4 - 6\rho_{13}^6\rho_{23}^2\rho_{12}^4 - 3\rho_{13}^4\rho_{23}^2\rho_{12}^4 + 15\rho_{13}^2\rho_{23}^2\rho_{12}^4 + 9\rho_{23}^2\rho_{12}^4 + 27\rho_{12}^4 \\
& - 3\rho_{13}\rho_{23}^9\rho_{12}^3 + 6\rho_{13}\rho_{23}^7\rho_{12}^3 - 6\rho_{13}^5\rho_{23}^5\rho_{12}^3 + 6\rho_{13}^3\rho_{23}^5\rho_{12}^3 - 9\rho_{13}\rho_{23}^5\rho_{12}^3 \\
& + 6\rho_{13}^5\rho_{23}^3\rho_{12}^3 - 18\rho_{13}\rho_{23}^3\rho_{12}^3 - 3\rho_{13}^9\rho_{23}\rho_{12}^3 + 6\rho_{13}^7\rho_{23}\rho_{12}^3 - 9\rho_{13}^5\rho_{23}\rho_{12}^3 \\
& - 18\rho_{13}^3\rho_{23}\rho_{12}^3 - 132\rho_{13}\rho_{23}\rho_{12}^3 + 3\rho_{13}^2\rho_{23}^{10}\rho_{12}^2 - 3\rho_{13}^8\rho_{12}^2 + 3\rho_{13}^4\rho_{23}^8\rho_{12}^2 \\
& - 12\rho_{13}^2\rho_{23}^8\rho_{12}^2 - 3\rho_{23}^8\rho_{12}^2 + 3\rho_{13}^6\rho_{12}^2 - 6\rho_{13}^4\rho_{23}^6\rho_{12}^2 + 15\rho_{13}^2\rho_{23}^6\rho_{12}^2 + 3\rho_{23}^6\rho_{12}^2 \\
& + 9\rho_{13}^4\rho_{12}^2 + 3\rho_{13}^8\rho_{23}^4\rho_{12}^2 - 6\rho_{13}^6\rho_{23}^4\rho_{12}^2 - 3\rho_{13}^4\rho_{23}^4\rho_{12}^2 + 15\rho_{13}^2\rho_{23}^4\rho_{12}^2 + 9\rho_{23}^4\rho_{12}^2 \\
& + 75\rho_{13}^2\rho_{12}^2 + 3\rho_{13}^{10}\rho_{23}^2\rho_{12}^2 - 12\rho_{13}^8\rho_{23}^2\rho_{12}^2 + 15\rho_{13}^6\rho_{23}^2\rho_{12}^2 + 15\rho_{13}^4\rho_{23}^2\rho_{12}^2 \\
& + 138\rho_{13}^2\rho_{23}^2\rho_{12}^2 + 75\rho_{23}^2\rho_{12}^2 - 84\rho_{12}^2 - 6\rho_{13}\rho_{23}^{11}\rho_{12} - 3\rho_{13}^3\rho_{23}^9\rho_{12} \\
& + 18\rho_{13}\rho_{23}^9\rho_{12} - 3\rho_{13}^5\rho_{23}^7\rho_{12} + 6\rho_{13}^3\rho_{23}^7\rho_{12} - 18\rho_{13}\rho_{23}^7\rho_{12} - 3\rho_{13}^7\rho_{23}^5\rho_{12} \\
& + 12\rho_{13}^5\rho_{23}^5\rho_{12} - 9\rho_{13}^3\rho_{23}^5\rho_{12} - 18\rho_{13}\rho_{23}^5\rho_{12} - 3\rho_{13}^9\rho_{23}^3\rho_{12} + 6\rho_{13}^7\rho_{23}^3\rho_{12} \\
& - 9\rho_{13}^5\rho_{23}^3\rho_{12} - 18\rho_{13}^3\rho_{23}^3\rho_{12} - 132\rho_{13}\rho_{23}^3\rho_{12} - 6\rho_{13}^{11}\rho_{23}\rho_{12} + 18\rho_{13}^9\rho_{23}\rho_{12} \\
& - 18\rho_{13}^7\rho_{23}\rho_{12} - 18\rho_{13}^5\rho_{23}\rho_{12} - 132\rho_{13}^3\rho_{23}\rho_{12} + 156\rho_{13}\rho_{23}\rho_{12} + 3\rho_{13}^{12} \\
& + 3\rho_{23}^{12} - 6\rho_{13}^{10} - 6\rho_{23}^{10} + 6\rho_{13}^8 + 3\rho_{13}^4\rho_{23}^8 - 3\rho_{13}^2\rho_{23}^8 + 6\rho_{23}^8 + 6\rho_{13}^6 - 3\rho_{13}^4\rho_{23}^6 \\
& + 3\rho_{13}^2\rho_{23}^6 + 6\rho_{23}^6 + 27\rho_{13}^4 + 3\rho_{13}^8\rho_{23}^4 - 3\rho_{13}^6\rho_{23}^4 + 3\rho_{13}^4\rho_{23}^4 + 9\rho_{13}^2\rho_{23}^4 \\
& + 27\rho_{23}^4 - 84\rho_{13}^2 - 3\rho_{13}^8\rho_{23}^2 + 3\rho_{13}^6\rho_{23}^2 + 9\rho_{13}^4\rho_{23}^2 + 75\rho_{13}^2\rho_{23}^2 - 84\rho_{23}^2 + 48),
\end{aligned}$$

$$\begin{aligned}
V \left[m_{b_1}^{np} (y_1, y_2, y_3, \boldsymbol{\rho}) \right] &= \frac{1}{3 \left(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1 \right)^2} \\
&\times \left(4\rho_{12}^{10} - 8\rho_{13}\rho_{23}\rho_{12}^9 + 4\rho_{13}^2\rho_{12}^8 - 6\rho_{12}^8 + 4\rho_{13}^3\rho_{23}^3\rho_{12}^7 + 4\rho_{13}\rho_{23}^3\rho_{12}^7 \right. \\
&\quad - 8\rho_{13}^3\rho_{23}\rho_{12}^7 + 12\rho_{13}\rho_{23}\rho_{12}^7 + 8\rho_{13}^4\rho_{12}^6 - 4\rho_{13}^2\rho_{23}^4\rho_{12}^6 - 8\rho_{13}^2\rho_{12}^6 \\
&\quad - 2\rho_{13}^2\rho_{23}^2\rho_{12}^6 + 2\rho_{12}^6 - 4\rho_{13}\rho_{23}^5\rho_{12}^5 + 8\rho_{13}^5\rho_{23}^3\rho_{12}^5 - 4\rho_{13}^3\rho_{23}^3\rho_{12}^5 \\
&\quad - 16\rho_{13}^5\rho_{23}\rho_{12}^5 + 12\rho_{13}^3\rho_{23}\rho_{12}^5 + 12\rho_{13}\rho_{23}\rho_{12}^5 + 8\rho_{13}^6\rho_{12}^4 + 4\rho_{13}^2\rho_{23}^6\rho_{12}^4 \\
&\quad + 4\rho_{23}^6\rho_{12}^4 - 12\rho_{13}^4\rho_{12}^4 - 8\rho_{13}^4\rho_{23}^4\rho_{12}^4 + 2\rho_{13}^2\rho_{23}^4\rho_{12}^4 - 6\rho_{13}^2\rho_{12}^4 \\
&\quad + 12\rho_{13}^4\rho_{23}^2\rho_{12}^4 - 50\rho_{13}^2\rho_{23}^2\rho_{12}^4 - 12\rho_{23}^2\rho_{12}^4 + 10\rho_{12}^4 + 4\rho_{13}^3\rho_{23}^7\rho_{12}^3 \\
&\quad - 4\rho_{13}\rho_{23}^7\rho_{12}^3 - 4\rho_{13}^3\rho_{23}^5\rho_{12}^3 + 4\rho_{13}^7\rho_{23}^3\rho_{12}^3 - 4\rho_{13}^5\rho_{23}^3\rho_{12}^3 + 32\rho_{13}^3\rho_{23}^3\rho_{12}^3 \\
&\quad + 48\rho_{13}\rho_{23}^3\rho_{12}^3 - 8\rho_{13}^7\rho_{23}\rho_{12}^3 + 12\rho_{13}^5\rho_{23}\rho_{12}^3 + 48\rho_{13}^3\rho_{23}\rho_{12}^3 - 52\rho_{13}\rho_{23}\rho_{12}^3 \\
&\quad + 4\rho_{13}^8\rho_{12}^2 - 8\rho_{13}^2\rho_{23}^8\rho_{12}^2 - 8\rho_{13}^6\rho_{12}^2 + 4\rho_{13}^4\rho_{23}^6\rho_{12}^2 + 8\rho_{13}^2\rho_{23}^6\rho_{12}^2 - 2\rho_{23}^6\rho_{12}^2 \\
&\quad - 6\rho_{13}^4\rho_{12}^2 - 4\rho_{13}^6\rho_{23}^4\rho_{12}^2 + 2\rho_{13}^4\rho_{23}^4\rho_{12}^2 - 36\rho_{13}^2\rho_{23}^4\rho_{12}^2 - 12\rho_{23}^4\rho_{12}^2 + 32\rho_{13}^2\rho_{12}^2 \\
&\quad - 2\rho_{13}^6\rho_{23}^2\rho_{12}^2 - 50\rho_{13}^4\rho_{23}^2\rho_{12}^2 + 16\rho_{13}^2\rho_{23}^2\rho_{12}^2 + 36\rho_{23}^2\rho_{12}^2 - 22\rho_{12}^2 + 4\rho_{13}\rho_{23}^9\rho_{12} \\
&\quad - 4\rho_{13}^3\rho_{23}^7\rho_{12} - 4\rho_{13}^5\rho_{23}^5\rho_{12} + 12\rho_{13}\rho_{23}^5\rho_{12} + 4\rho_{13}^7\rho_{23}^3\rho_{12} + 48\rho_{13}^3\rho_{23}^3\rho_{12} \\
&\quad - 52\rho_{13}\rho_{23}^3\rho_{12} - 8\rho_{13}^9\rho_{23}\rho_{12} + 12\rho_{13}^7\rho_{23}\rho_{12} + 12\rho_{13}^5\rho_{23}\rho_{12} - 52\rho_{13}^3\rho_{23}\rho_{12} \\
&\quad + 36\rho_{13}\rho_{23}\rho_{12} + 4\rho_{13}^{10} - 6\rho_{13}^8 + 2\rho_{13}^6 + 4\rho_{13}^4\rho_{23}^6 - 2\rho_{13}^2\rho_{23}^6 + 10\rho_{13}^4 \\
&\quad \left. - 12\rho_{13}^2\rho_{23}^4 + 12\rho_{23}^4 - 22\rho_{13}^2 - 12\rho_{13}^4\rho_{23}^2 + 36\rho_{13}^2\rho_{23}^2 - 24\rho_{23}^2 + 12 \right),
\end{aligned}$$

$$\begin{aligned}
V \left[m_{b_2}^{np} (y_1, y_2, y_3, \boldsymbol{\rho}) \right] &= \frac{1}{3 \left(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1 \right)^2} \\
&\times \left(4\rho_{12}^{10} - 8\rho_{13}\rho_{23}\rho_{12}^9 + 4\rho_{23}^2\rho_{12}^8 - 6\rho_{12}^8 + 4\rho_{13}^3\rho_{23}^3\rho_{12}^7 - 8\rho_{13}\rho_{23}^3\rho_{12}^7 \right. \\
&\quad + 4\rho_{13}^3\rho_{23}\rho_{12}^7 + 12\rho_{13}\rho_{23}\rho_{12}^7 + 8\rho_{23}^4\rho_{12}^6 - 4\rho_{13}^4\rho_{23}^2\rho_{12}^6 - 2\rho_{13}^2\rho_{23}^2\rho_{12}^6 \\
&\quad - 8\rho_{23}^2\rho_{12}^6 + 2\rho_{12}^6 + 8\rho_{13}^3\rho_{23}^5\rho_{12}^5 - 16\rho_{13}\rho_{23}^5\rho_{12}^5 - 4\rho_{13}^3\rho_{23}^3\rho_{12}^5 + 12\rho_{13}\rho_{23}^3\rho_{12}^5 \\
&\quad - 4\rho_{13}^5\rho_{23}\rho_{12}^5 + 12\rho_{13}\rho_{23}\rho_{12}^5 + 4\rho_{13}^6\rho_{12}^4 + 8\rho_{23}^6\rho_{12}^4 - 8\rho_{13}^4\rho_{23}^4\rho_{12}^4 \\
&\quad + 12\rho_{13}^2\rho_{23}^4\rho_{12}^4 - 12\rho_{23}^4\rho_{12}^4 - 12\rho_{13}^2\rho_{12}^4 + 4\rho_{13}^6\rho_{23}^2\rho_{12}^4 + 2\rho_{13}^4\rho_{23}^2\rho_{12}^4 \\
&\quad - 50\rho_{13}^2\rho_{23}^2\rho_{12}^4 - 6\rho_{23}^2\rho_{12}^4 + 10\rho_{12}^4 + 4\rho_{13}^3\rho_{23}^7\rho_{12}^3 - 8\rho_{13}\rho_{23}^7\rho_{12}^3 - 4\rho_{13}^3\rho_{23}^5\rho_{12}^3 \\
&\quad + 12\rho_{13}\rho_{23}^5\rho_{12}^3 + 4\rho_{13}^7\rho_{23}^3\rho_{12}^3 - 4\rho_{13}^5\rho_{23}^3\rho_{12}^3 + 32\rho_{13}^3\rho_{23}^3\rho_{12}^3 + 48\rho_{13}\rho_{23}^3\rho_{12}^3 \\
&\quad - 4\rho_{13}^7\rho_{23}\rho_{12}^3 + 48\rho_{13}^3\rho_{23}\rho_{12}^3 - 52\rho_{13}\rho_{23}\rho_{12}^3 + 4\rho_{23}^8\rho_{12}^2 - 2\rho_{13}^6\rho_{12}^2 \\
&\quad - 4\rho_{13}^4\rho_{23}^6\rho_{12}^2 - 2\rho_{13}^2\rho_{23}^6\rho_{12}^2 - 8\rho_{23}^6\rho_{12}^2 - 12\rho_{13}^4\rho_{12}^2 + 4\rho_{13}^6\rho_{23}^4\rho_{12}^2 \\
&\quad + 2\rho_{13}^4\rho_{23}^4\rho_{12}^2 - 50\rho_{13}^2\rho_{23}^4\rho_{12}^2 - 6\rho_{23}^4\rho_{12}^2 + 36\rho_{13}^2\rho_{12}^2 - 8\rho_{13}^8\rho_{23}^2\rho_{12}^2 \\
&\quad + 8\rho_{13}^6\rho_{23}^2\rho_{12}^2 - 36\rho_{13}^4\rho_{23}^2\rho_{12}^2 + 16\rho_{13}^2\rho_{23}^2\rho_{12}^2 + 32\rho_{23}^2\rho_{12}^2 - 22\rho_{12}^2 \\
&\quad - 8\rho_{13}\rho_{23}^9\rho_{12} + 4\rho_{13}^3\rho_{23}^7\rho_{12} + 12\rho_{13}\rho_{23}^7\rho_{12} - 4\rho_{13}^5\rho_{23}^5\rho_{12} + 12\rho_{13}\rho_{23}^5\rho_{12} \\
&\quad - 4\rho_{13}^7\rho_{23}^3\rho_{12} + 48\rho_{13}^3\rho_{23}^3\rho_{12} - 52\rho_{13}\rho_{23}^3\rho_{12} + 4\rho_{13}^9\rho_{23}\rho_{12} + 12\rho_{13}^5\rho_{23}\rho_{12} \\
&\quad - 52\rho_{13}^3\rho_{23}\rho_{12} + 36\rho_{13}\rho_{23}\rho_{12} + 4\rho_{23}^{10} - 6\rho_{23}^8 + 2\rho_{23}^6 + 12\rho_{13}^4 + 4\rho_{13}^6\rho_{23}^4 \\
&\quad \left. - 12\rho_{13}^2\rho_{23}^4 + 10\rho_{23}^4 - 24\rho_{13}^2 - 2\rho_{13}^6\rho_{23}^2 - 12\rho_{13}^4\rho_{23}^2 + 36\rho_{13}^2\rho_{23}^2 - 22\rho_{23}^2 + 12 \right),
\end{aligned}$$

and

$$\begin{aligned}
V \left[m_{b_3}^{np} (y_1, y_2, y_3, \boldsymbol{\rho}) \right] &= \frac{1}{3 \left(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1 \right)^2} \\
&\times \left(4\rho_{13}^{10} - 8\rho_{12}\rho_{23}\rho_{13}^9 + 4\rho_{23}^2\rho_{13}^8 - 6\rho_{13}^8 + 4\rho_{12}^3\rho_{23}^3\rho_{13}^7 - 8\rho_{12}\rho_{23}^3\rho_{13}^7 \right. \\
&\quad + 4\rho_{12}^3\rho_{23}\rho_{13}^7 + 12\rho_{12}\rho_{23}\rho_{13}^7 + 8\rho_{23}^4\rho_{13}^6 - 4\rho_{12}^4\rho_{23}^2\rho_{13}^6 - 2\rho_{12}^2\rho_{23}^2\rho_{13}^6 \\
&\quad - 8\rho_{23}^2\rho_{13}^6 + 2\rho_{13}^6 + 8\rho_{12}^3\rho_{23}^5\rho_{13}^5 - 16\rho_{12}\rho_{23}^5\rho_{13}^5 - 4\rho_{12}^3\rho_{23}^3\rho_{13}^5 + 12\rho_{12}\rho_{23}^3\rho_{13}^5 \\
&\quad - 4\rho_{12}^5\rho_{23}\rho_{13}^5 + 12\rho_{12}\rho_{23}\rho_{13}^5 + 4\rho_{12}^6\rho_{13}^4 + 8\rho_{23}^6\rho_{13}^4 - 8\rho_{12}^4\rho_{23}^4\rho_{13}^4 \\
&\quad + 12\rho_{12}^2\rho_{23}^4\rho_{13}^4 - 12\rho_{23}^4\rho_{13}^4 - 12\rho_{12}^2\rho_{13}^4 + 4\rho_{12}^6\rho_{23}^2\rho_{13}^4 + 2\rho_{12}^4\rho_{23}^2\rho_{13}^4 \\
&\quad - 50\rho_{12}^2\rho_{23}^2\rho_{13}^4 - 6\rho_{23}^2\rho_{13}^4 + 10\rho_{13}^4 + 4\rho_{12}^3\rho_{23}^7\rho_{13}^3 - 8\rho_{12}\rho_{23}^7\rho_{13}^3 \\
&\quad - 4\rho_{12}^3\rho_{23}^5\rho_{13}^3 + 12\rho_{12}\rho_{23}^5\rho_{13}^3 + 4\rho_{12}^7\rho_{23}^3\rho_{13}^3 - 4\rho_{12}^5\rho_{23}^3\rho_{13}^3 + 32\rho_{12}^3\rho_{23}^3\rho_{13}^3 \\
&\quad + 48\rho_{12}\rho_{23}^3\rho_{13}^3 - 4\rho_{12}^7\rho_{23}\rho_{13}^3 + 48\rho_{12}^3\rho_{23}\rho_{13}^3 - 52\rho_{12}\rho_{23}\rho_{13}^3 + 4\rho_{23}^8\rho_{13}^2 \\
&\quad - 2\rho_{12}^6\rho_{13}^2 - 4\rho_{12}^4\rho_{23}^6\rho_{13}^2 - 2\rho_{12}^2\rho_{23}^6\rho_{13}^2 - 8\rho_{23}^6\rho_{13}^2 - 12\rho_{12}^4\rho_{13}^2 + 4\rho_{12}^6\rho_{23}^4\rho_{13}^2 \\
&\quad + 2\rho_{12}^4\rho_{23}^4\rho_{13}^2 - 50\rho_{12}^2\rho_{23}^4\rho_{13}^2 - 6\rho_{23}^4\rho_{13}^2 + 36\rho_{12}^2\rho_{13}^2 - 8\rho_{12}^8\rho_{23}^2\rho_{13}^2 \\
&\quad + 8\rho_{12}^6\rho_{23}^2\rho_{13}^2 - 36\rho_{12}^4\rho_{23}^2\rho_{13}^2 + 16\rho_{12}^2\rho_{23}^2\rho_{13}^2 + 32\rho_{23}^2\rho_{13}^2 - 22\rho_{13}^2 - 8\rho_{12}\rho_{23}^9\rho_{13} \\
&\quad + 4\rho_{12}^3\rho_{23}^7\rho_{13} + 12\rho_{12}\rho_{23}^7\rho_{13} - 4\rho_{12}^5\rho_{23}^5\rho_{13} + 12\rho_{12}\rho_{23}^5\rho_{13} - 4\rho_{12}^7\rho_{23}^3\rho_{13} \\
&\quad + 48\rho_{12}^3\rho_{23}^3\rho_{13} - 52\rho_{12}\rho_{23}^3\rho_{13} + 4\rho_{12}^9\rho_{23}\rho_{13} + 12\rho_{12}^5\rho_{23}\rho_{13} - 52\rho_{12}^3\rho_{23}\rho_{13} \\
&\quad + 36\rho_{12}\rho_{23}\rho_{13} + 4\rho_{23}^{10} - 6\rho_{23}^8 + 2\rho_{23}^6 + 12\rho_{12}^4 + 4\rho_{12}^6\rho_{23}^4 - 12\rho_{12}^2\rho_{23}^4 + 10\rho_{23}^4 \\
&\quad \left. - 24\rho_{12}^2 - 2\rho_{12}^6\rho_{23}^2 - 12\rho_{12}^4\rho_{23}^2 + 36\rho_{12}^2\rho_{23}^2 - 22\rho_{23}^2 + 12 \right).
\end{aligned}$$

In turn, the covariances are

$$\begin{aligned}
cov \left[m_{b_1}^{np}(\mathbf{y}, \boldsymbol{\rho}), m_{b_2}^{np}(\mathbf{y}, \boldsymbol{\rho}) \right] &= \frac{1}{3 \left(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1 \right)^2} \\
&\times \left(2\rho_{12}^{11} - 4\rho_{13}\rho_{23}\rho_{12}^{10} + 2\rho_{13}^2\rho_{23}^2\rho_{12}^9 - 2\rho_{12}^9 - 2\rho_{13}\rho_{23}^3\rho_{12}^8 - 2\rho_{13}^3\rho_{23}\rho_{12}^8 \right. \\
&+ 8\rho_{13}\rho_{23}\rho_{12}^8 + 2\rho_{13}^4\rho_{12}^7 + 2\rho_{13}^2\rho_{23}^4\rho_{12}^7 + 2\rho_{23}^4\rho_{12}^7 + 2\rho_{13}^4\rho_{23}^2\rho_{12}^7 \\
&- 10\rho_{13}^2\rho_{23}^2\rho_{12}^7 + 8\rho_{12}^7 - 4\rho_{13}\rho_{23}^5\rho_{12}^6 + 4\rho_{13}^3\rho_{23}^3\rho_{12}^6 + 6\rho_{13}\rho_{23}^3\rho_{12}^6 \\
&- 4\rho_{13}^5\rho_{23}\rho_{12}^6 + 6\rho_{13}^3\rho_{23}\rho_{12}^6 - 30\rho_{13}\rho_{23}\rho_{12}^6 + 2\rho_{13}^6\rho_{12}^5 + 2\rho_{13}^2\rho_{23}^6\rho_{12}^5 \\
&+ 2\rho_{23}^6\rho_{12}^5 + 4\rho_{13}^4\rho_{23}^4\rho_{12}^5 - 14\rho_{13}^2\rho_{23}^4\rho_{12}^5 + 8\rho_{13}^2\rho_{12}^5 + 2\rho_{13}^6\rho_{23}^2\rho_{12}^5 \\
&- 14\rho_{13}^4\rho_{23}^2\rho_{12}^5 + 26\rho_{13}^2\rho_{23}^2\rho_{12}^5 + 8\rho_{23}^2\rho_{12}^5 - 12\rho_{12}^5 - 2\rho_{13}\rho_{23}^7\rho_{12}^4 \\
&+ 4\rho_{13}\rho_{23}^5\rho_{12}^4 + 18\rho_{13}^3\rho_{23}^3\rho_{12}^4 - 26\rho_{13}\rho_{23}^3\rho_{12}^4 - 2\rho_{13}^7\rho_{23}\rho_{12}^4 + 4\rho_{13}^5\rho_{23}\rho_{12}^4 \\
&- 26\rho_{13}^3\rho_{23}\rho_{12}^4 + 24\rho_{13}\rho_{23}\rho_{12}^4 + 2\rho_{13}^2\rho_{23}^8\rho_{12}^3 + 2\rho_{13}^4\rho_{23}^6\rho_{12}^3 - 6\rho_{13}^2\rho_{23}^6\rho_{12}^3 \\
&+ 6\rho_{13}^4\rho_{12}^3 + 2\rho_{13}^6\rho_{23}^4\rho_{12}^3 - 8\rho_{13}^4\rho_{23}^4\rho_{12}^3 + 22\rho_{13}^2\rho_{23}^4\rho_{12}^3 + 6\rho_{23}^4\rho_{12}^3 \\
&- 4\rho_{13}^2\rho_{12}^3 + 2\rho_{13}^8\rho_{23}^2\rho_{12}^3 - 6\rho_{13}^6\rho_{23}^2\rho_{12}^3 + 22\rho_{13}^4\rho_{23}^2\rho_{12}^3 + 6\rho_{13}^2\rho_{23}^2\rho_{12}^3 \\
&- 4\rho_{23}^2\rho_{12}^3 - 2\rho_{12}^3 - 4\rho_{13}\rho_{23}^9\rho_{12}^2 + 2\rho_{13}^3\rho_{23}^7\rho_{12}^2 + 8\rho_{13}\rho_{23}^7\rho_{12}^2 \\
&- 4\rho_{13}^5\rho_{23}^5\rho_{12}^2 - 2\rho_{13}^3\rho_{23}^5\rho_{12}^2 - 22\rho_{13}\rho_{23}^5\rho_{12}^2 + 2\rho_{13}^7\rho_{23}^3\rho_{12}^2 - 2\rho_{13}^5\rho_{23}^3\rho_{12}^2 - \\
&52\rho_{13}^3\rho_{23}^3\rho_{12}^2 + 12\rho_{13}\rho_{23}^3\rho_{12}^2 - 4\rho_{13}^9\rho_{23}\rho_{12}^2 + 8\rho_{13}^7\rho_{23}\rho_{12}^2 - 22\rho_{13}^5\rho_{23}\rho_{12}^2 \\
&+ 12\rho_{13}^3\rho_{23}\rho_{12}^2 + 6\rho_{13}\rho_{23}\rho_{12}^2 + 2\rho_{13}^{10}\rho_{12} + 2\rho_{23}^{10}\rho_{12} - 2\rho_{13}^8\rho_{12} - 6\rho_{13}^2\rho_{23}^8\rho_{12} \\
&- 2\rho_{23}^8\rho_{12} + 6\rho_{13}^6\rho_{12} + 4\rho_{13}^2\rho_{23}^6\rho_{12} + 6\rho_{23}^6\rho_{12} - 8\rho_{13}^4\rho_{12} + 8\rho_{13}^4\rho_{23}^4\rho_{12} \\
&+ 30\rho_{13}^2\rho_{23}^4\rho_{12} - 8\rho_{23}^4\rho_{12} - 4\rho_{13}^2\rho_{12} - 6\rho_{13}^8\rho_{23}^2\rho_{12} + 4\rho_{13}^6\rho_{23}^2\rho_{12} \\
&+ 30\rho_{13}^4\rho_{23}^2\rho_{12} - 24\rho_{13}^2\rho_{23}^2\rho_{12} - 4\rho_{23}^2\rho_{12} + 6\rho_{12} + 2\rho_{13}\rho_{23}^9 + 2\rho_{13}^3\rho_{23}^7 \\
&- 2\rho_{13}\rho_{23}^7 - 2\rho_{13}^3\rho_{23}^5 - 6\rho_{13}\rho_{23}^5 + 2\rho_{13}^7\rho_{23}^3 - 2\rho_{13}^5\rho_{23}^3 - 10\rho_{13}^3\rho_{23}^3 \\
&\left. + 10\rho_{13}\rho_{23}^3 + 2\rho_{13}^9\rho_{23} - 2\rho_{13}^7\rho_{23} - 6\rho_{13}^5\rho_{23} + 10\rho_{13}^3\rho_{23} - 4\rho_{13}\rho_{23} \right),
\end{aligned}$$

$$\begin{aligned}
\text{cov} \left[m_{b_1}^{np}(\mathbf{y}, \boldsymbol{\rho}), m_{b_3}^{np}(\mathbf{y}, \boldsymbol{\rho}) \right] &= \frac{1}{3 \left(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1 \right)^2} \\
&\times \left(2\rho_{13}^{11} - 4\rho_{12}\rho_{23}\rho_{13}^{10} + 2\rho_{12}^2\rho_{23}^2\rho_{13}^9 - 2\rho_{13}^9 - 2\rho_{12}\rho_{23}^3\rho_{13}^8 - 2\rho_{12}^3\rho_{23}\rho_{13}^8 \right. \\
&\quad + 8\rho_{12}\rho_{23}\rho_{13}^8 + 2\rho_{12}^4\rho_{13}^7 + 2\rho_{12}^2\rho_{23}^4\rho_{13}^7 + 2\rho_{23}^4\rho_{13}^7 + 2\rho_{12}^4\rho_{23}^2\rho_{13}^7 \\
&\quad - 10\rho_{12}^2\rho_{23}^2\rho_{13}^7 + 8\rho_{13}^7 - 4\rho_{12}\rho_{23}^5\rho_{13}^6 + 4\rho_{12}^3\rho_{23}^3\rho_{13}^6 + 6\rho_{12}\rho_{23}^3\rho_{13}^6 \\
&\quad - 4\rho_{12}^5\rho_{23}\rho_{13}^6 + 6\rho_{12}^3\rho_{23}\rho_{13}^6 - 30\rho_{12}\rho_{23}\rho_{13}^6 + 2\rho_{12}^6\rho_{13}^5 + 2\rho_{12}^2\rho_{23}^6\rho_{13}^5 \\
&\quad + 2\rho_{23}^6\rho_{13}^5 + 4\rho_{12}^4\rho_{23}^4\rho_{13}^5 - 14\rho_{12}^2\rho_{23}^4\rho_{13}^5 + 8\rho_{12}^2\rho_{13}^5 + 2\rho_{12}^6\rho_{23}^2\rho_{13}^5 \\
&\quad - 14\rho_{12}^4\rho_{23}^2\rho_{13}^5 + 26\rho_{12}^2\rho_{23}^2\rho_{13}^5 + 8\rho_{23}^2\rho_{13}^5 - 12\rho_{13}^5 - 2\rho_{12}\rho_{23}^7\rho_{13}^4 \\
&\quad + 4\rho_{12}\rho_{23}^5\rho_{13}^4 + 18\rho_{12}^3\rho_{23}^3\rho_{13}^4 - 26\rho_{12}\rho_{23}^3\rho_{13}^4 - 2\rho_{12}^7\rho_{23}\rho_{13}^4 + 4\rho_{12}^5\rho_{23}\rho_{13}^4 \\
&\quad - 26\rho_{12}^3\rho_{23}\rho_{13}^4 + 24\rho_{12}\rho_{23}\rho_{13}^4 + 2\rho_{12}^2\rho_{23}^8\rho_{13}^3 + 2\rho_{12}^4\rho_{23}^6\rho_{13}^3 - 6\rho_{12}^2\rho_{23}^6\rho_{13}^3 \\
&\quad + 6\rho_{12}^4\rho_{13}^3 + 2\rho_{12}^6\rho_{23}^4\rho_{13}^3 - 8\rho_{12}^4\rho_{23}^4\rho_{13}^3 + 22\rho_{12}^2\rho_{23}^4\rho_{13}^3 + 6\rho_{23}^4\rho_{13}^3 \\
&\quad - 4\rho_{12}^2\rho_{13}^3 + 2\rho_{12}^8\rho_{23}^2\rho_{13}^3 - 6\rho_{12}^6\rho_{23}^2\rho_{13}^3 + 22\rho_{12}^4\rho_{23}^2\rho_{13}^3 + 6\rho_{12}^2\rho_{23}^2\rho_{13}^3 \\
&\quad - 4\rho_{23}^2\rho_{13}^3 - 2\rho_{13}^3 - 4\rho_{12}\rho_{23}^9\rho_{13}^2 + 2\rho_{12}^3\rho_{23}^7\rho_{13}^2 + 8\rho_{12}\rho_{23}^7\rho_{13}^2 - 4\rho_{12}^5\rho_{23}^5\rho_{13}^2 \\
&\quad - 2\rho_{12}^3\rho_{23}^5\rho_{13}^2 - 22\rho_{12}\rho_{23}^5\rho_{13}^2 + 2\rho_{12}^7\rho_{23}^3\rho_{13}^2 - 2\rho_{12}^5\rho_{23}^3\rho_{13}^2 - 52\rho_{12}^3\rho_{23}^3\rho_{13}^2 \\
&\quad + 12\rho_{12}\rho_{23}^3\rho_{13}^2 - 4\rho_{12}^9\rho_{23}\rho_{13}^2 + 8\rho_{12}^7\rho_{23}\rho_{13}^2 - 22\rho_{12}^5\rho_{23}\rho_{13}^2 + 12\rho_{12}^3\rho_{23}\rho_{13}^2 \\
&\quad + 6\rho_{12}\rho_{23}\rho_{13}^2 + 2\rho_{12}^{10}\rho_{13} + 2\rho_{23}^{10}\rho_{13} - 2\rho_{12}^8\rho_{13} - 6\rho_{12}^2\rho_{23}^8\rho_{13} - 2\rho_{23}^8\rho_{13} \\
&\quad + 6\rho_{12}^6\rho_{13} + 4\rho_{12}^2\rho_{23}^6\rho_{13} + 6\rho_{23}^6\rho_{13} - 8\rho_{12}^4\rho_{13} + 8\rho_{12}^4\rho_{23}^4\rho_{13} + 30\rho_{12}^2\rho_{23}^4\rho_{13} \\
&\quad - 8\rho_{23}^4\rho_{13} - 4\rho_{12}^2\rho_{13} - 6\rho_{12}^8\rho_{23}^2\rho_{13} + 4\rho_{12}^6\rho_{23}^2\rho_{13} + 30\rho_{12}^4\rho_{23}^2\rho_{13} \\
&\quad - 24\rho_{12}^2\rho_{23}^2\rho_{13} - 4\rho_{23}^2\rho_{13} + 6\rho_{13} + 2\rho_{12}\rho_{23}^9 + 2\rho_{12}^3\rho_{23}^7 - 2\rho_{12}\rho_{23}^7 \\
&\quad - 2\rho_{12}^3\rho_{23}^5 - 6\rho_{12}\rho_{23}^5 + 2\rho_{12}^7\rho_{23}^3 - 2\rho_{12}^5\rho_{23}^3 - 10\rho_{12}^3\rho_{23}^3 + 10\rho_{12}\rho_{23}^3 \\
&\quad \left. + 2\rho_{12}^9\rho_{23} - 2\rho_{12}^7\rho_{23} - 6\rho_{12}^5\rho_{23} + 10\rho_{12}^3\rho_{23} - 4\rho_{12}\rho_{23} \right),
\end{aligned}$$

$$\begin{aligned}
\text{cov} \left[m_{b_2}^{np}(\mathbf{y}, \boldsymbol{\rho}), m_{b_3}^{np}(\mathbf{y}, \boldsymbol{\rho}) \right] &= \frac{1}{3 \left(\rho_{12}^2 - 2\rho_{13}\rho_{23}\rho_{12} + \rho_{13}^2 + \rho_{23}^2 - 1 \right)^2} \\
&\times \left(2\rho_{23}^{11} - 4\rho_{12}\rho_{13}\rho_{23}^{10} + 2\rho_{12}^2\rho_{13}^2\rho_{23}^9 - 2\rho_{23}^9 - 2\rho_{12}\rho_{13}^3\rho_{23}^8 - 2\rho_{12}^3\rho_{13}\rho_{23}^8 \right. \\
&\quad + 8\rho_{12}\rho_{13}\rho_{23}^8 + 2\rho_{12}^4\rho_{23}^7 + 2\rho_{12}^2\rho_{13}^4\rho_{23}^7 + 2\rho_{13}^4\rho_{23}^7 + 2\rho_{12}^4\rho_{13}^2\rho_{23}^7 \\
&\quad - 10\rho_{12}^2\rho_{13}^2\rho_{23}^7 + 8\rho_{23}^7 - 4\rho_{12}\rho_{13}^5\rho_{23}^6 + 4\rho_{12}^3\rho_{13}^3\rho_{23}^6 + 6\rho_{12}\rho_{13}^3\rho_{23}^6 \\
&\quad - 4\rho_{12}^5\rho_{13}\rho_{23}^6 + 6\rho_{12}^3\rho_{13}\rho_{23}^6 - 30\rho_{12}\rho_{13}\rho_{23}^6 + 2\rho_{12}^6\rho_{23}^5 + 2\rho_{12}^2\rho_{13}^6\rho_{23}^5 \\
&\quad + 2\rho_{13}^6\rho_{23}^5 + 4\rho_{12}^4\rho_{13}^4\rho_{23}^5 - 14\rho_{12}^2\rho_{13}^4\rho_{23}^5 + 8\rho_{12}^2\rho_{23}^5 + 2\rho_{12}^6\rho_{13}^2\rho_{23}^5 \\
&\quad - 14\rho_{12}^4\rho_{13}^2\rho_{23}^5 + 26\rho_{12}^2\rho_{13}^2\rho_{23}^5 + 8\rho_{13}^2\rho_{23}^5 - 12\rho_{23}^5 - 2\rho_{12}\rho_{13}^7\rho_{23}^4 \\
&\quad + 4\rho_{12}\rho_{13}^5\rho_{23}^4 + 18\rho_{12}^3\rho_{13}^3\rho_{23}^4 - 26\rho_{12}\rho_{13}^3\rho_{23}^4 - 2\rho_{12}^7\rho_{13}\rho_{23}^4 + 4\rho_{12}^5\rho_{13}\rho_{23}^4 \\
&\quad - 26\rho_{12}^3\rho_{13}\rho_{23}^4 + 24\rho_{12}\rho_{13}\rho_{23}^4 + 2\rho_{12}^2\rho_{13}^8\rho_{23}^3 + 2\rho_{12}^4\rho_{13}^6\rho_{23}^3 - 6\rho_{12}^2\rho_{13}^6\rho_{23}^3 \\
&\quad + 6\rho_{12}^4\rho_{23}^3 + 2\rho_{12}^6\rho_{13}^4\rho_{23}^3 - 8\rho_{12}^4\rho_{13}^4\rho_{23}^3 + 22\rho_{12}^2\rho_{13}^4\rho_{23}^3 + 6\rho_{13}^4\rho_{23}^3 \\
&\quad - 4\rho_{12}^2\rho_{23}^3 + 2\rho_{12}^8\rho_{13}^2\rho_{23}^3 - 6\rho_{12}^6\rho_{13}^2\rho_{23}^3 + 22\rho_{12}^4\rho_{13}^2\rho_{23}^3 + 6\rho_{12}^2\rho_{13}^2\rho_{23}^3 \\
&\quad - 4\rho_{13}^2\rho_{23}^3 - 2\rho_{23}^3 - 4\rho_{12}\rho_{13}^9\rho_{23}^2 + 2\rho_{12}^3\rho_{13}^7\rho_{23}^2 + 8\rho_{12}\rho_{13}^7\rho_{23}^2 - 4\rho_{12}^5\rho_{13}^5\rho_{23}^2 \\
&\quad - 2\rho_{12}^3\rho_{13}^5\rho_{23}^2 - 22\rho_{12}\rho_{13}^5\rho_{23}^2 + 2\rho_{12}^7\rho_{13}^3\rho_{23}^2 - 2\rho_{12}^5\rho_{13}^3\rho_{23}^2 - 52\rho_{12}^3\rho_{13}^3\rho_{23}^2 \\
&\quad + 12\rho_{12}\rho_{13}^3\rho_{23}^2 - 4\rho_{12}^9\rho_{13}\rho_{23}^2 + 8\rho_{12}^7\rho_{13}\rho_{23}^2 - 22\rho_{12}^5\rho_{13}\rho_{23}^2 + 12\rho_{12}^3\rho_{13}\rho_{23}^2 \\
&\quad + 6\rho_{12}\rho_{13}\rho_{23}^2 + 2\rho_{12}^{10}\rho_{23} + 2\rho_{13}^{10}\rho_{23} - 2\rho_{12}^8\rho_{23} - 6\rho_{12}^2\rho_{13}^8\rho_{23} - 2\rho_{13}^8\rho_{23} \\
&\quad + 6\rho_{12}^6\rho_{23} + 4\rho_{12}^2\rho_{13}^6\rho_{23} + 6\rho_{13}^6\rho_{23} - 8\rho_{12}^4\rho_{23} + 8\rho_{12}^4\rho_{13}^4\rho_{23} + 30\rho_{12}^2\rho_{13}^4\rho_{23} \\
&\quad - 8\rho_{13}^4\rho_{23} - 4\rho_{12}^2\rho_{23} - 6\rho_{12}^8\rho_{13}^2\rho_{23} + 4\rho_{12}^6\rho_{13}^2\rho_{23} + 30\rho_{12}^4\rho_{13}^2\rho_{23} \\
&\quad - 24\rho_{12}^2\rho_{13}^2\rho_{23} - 4\rho_{13}^2\rho_{23} + 6\rho_{23} + 2\rho_{12}\rho_{13}^9 + 2\rho_{12}^3\rho_{13}^7 - 2\rho_{12}\rho_{13}^7 \\
&\quad - 2\rho_{12}^3\rho_{13}^5 - 6\rho_{12}\rho_{13}^5 + 2\rho_{12}^7\rho_{13}^3 - 2\rho_{12}^5\rho_{13}^3 - 10\rho_{12}^3\rho_{13}^3 + 10\rho_{12}\rho_{13}^3 \\
&\quad \left. + 2\rho_{12}^9\rho_{13} - 2\rho_{12}^7\rho_{13} - 6\rho_{12}^5\rho_{13} + 10\rho_{12}^3\rho_{13} - 4\rho_{12}\rho_{13} \right),
\end{aligned}$$

and

$$\text{cov} \left[s_{\eta}^{np}(\mathbf{y}, \boldsymbol{\rho}), m_{b_k}^{np}(\mathbf{y}, \boldsymbol{\rho}) \right] = 0, \quad \text{for } k = 1, 2, 3.$$

CEMFI WORKING PAPERS

- 0801 *David Martinez-Miera and Rafael Repullo*: "Does competition reduce the risk of bank failure?".
- 0802 *Joan Llull*: "The impact of immigration on productivity".
- 0803 *Cristina López-Mayán*: "Microeconometric analysis of residential water demand".
- 0804 *Javier Mencía and Enrique Sentana*: "Distributional tests in multivariate dynamic models with Normal and Student t innovations".
- 0805 *Javier Mencía and Enrique Sentana*: "Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation".
- 0806 *Dante Amengual and Enrique Sentana*: "A comparison of mean-variance efficiency tests".
- 0807 *Enrique Sentana*: "The econometrics of mean-variance efficiency tests: A survey".
- 0808 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla*: "Are joint negotiations in standard setting "reasonably necessary"?".
- 0809 *Rafael Repullo and Javier Suarez*: "The procyclical effects of Basel II".
- 0810 *Ildefonso Mendez*: "Promoting permanent employment: Lessons from Spain".
- 0811 *Ildefonso Mendez*: "Intergenerational time transfers and internal migration: Accounting for low spatial mobility in Southern Europe".
- 0812 *Francisco Maeso and Ildefonso Mendez*: "The role of partnership status and expectations on the emancipation behaviour of Spanish graduates".
- 0813 *Rubén Hernández-Murillo, Gerard Llobet and Roberto Fuentes*: "Strategic online-banking adoption".
- 0901 *Max Bruche and Javier Suarez*: "The macroeconomics of money market freezes".
- 0902 *Max Bruche*: "Bankruptcy codes, liquidation timing, and debt valuation".
- 0903 *Rafael Repullo, Jesús Saurina and Carlos Trucharte*: "Mitigating the procyclicality of Basel II".
- 0904 *Manuel Arellano and Stéphane Bonhomme*: "Identifying distributional characteristics in random coefficients panel data models".
- 0905 *Manuel Arellano, Lars Peter Hansen and Enrique Sentana*: "Underidentification?".
- 0906 *Stéphane Bonhomme and Ulrich Sauder*: "Accounting for unobservables in comparing selective and comprehensive schooling".
- 0907 *Roberto Serrano*: "On Watson's non-forcing contracts and renegotiation".
- 0908 *Roberto Serrano and Rajiv Vohra*: "Multiplicity of mixed equilibria in mechanisms: a unified approach to exact and approximate implementation".
- 0909 *Roland Pongou and Roberto Serrano*: "A dynamic theory of fidelity networks with an application to the spread of HIV / AIDS".
- 0910 *Josep Pijoan-Mas and Virginia Sánchez-Marcos*: "Spain is different: Falling trends of inequality".
- 0911 *Yusuke Kamishiro and Roberto Serrano*: "Equilibrium blocking in large quasilinear economies".
- 0912 *Gabriele Fiorentini and Enrique Sentana*: "Dynamic specification tests for static factor models".

- 0913 *Javier Mencía and Enrique Sentana*: "Valuation of VIX derivatives".
- 1001 *Gerard Llobet and Javier Suarez*: "Entrepreneurial innovation, patent protection and industry dynamics".
- 1002 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla*: "An economic take on patent licensing: Understanding the implications of the "first sale patent exhaustion" doctrine.
- 1003 *Max Bruche and Gerard Llobet*: "Walking wounded or living dead? Making banks foreclose bad loans".
- 1004 *Francisco Peñaranda and Enrique Sentana*: "A Unifying approach to the empirical evaluation of asset pricing models".
- 1005 *Javier Suarez*: "The Spanish crisis: Background and policy challenges".
- 1006 *Enrique Moral-Benito*: "Panel growth regressions with general predetermined variables: Likelihood-based estimation and Bayesian averaging".
- 1007 *Laura Crespo and Pedro Mira*: "Caregiving to elderly parents and employment status of European mature women".
- 1008 *Enrique Moral-Benito*: "Model averaging in economics".
- 1009 *Samuel Bentolila, Pierre Cahuc, Juan J. Dolado and Thomas Le Barbanchon*: "Two-tier labor markets in the Great Recession: France vs. Spain".
- 1010 *Manuel García-Santana and Josep Pijoan-Mas*: "Small Scale Reservation Laws and the misallocation of talent".
- 1101 *Javier Díaz-Giménez and Josep Pijoan-Mas*: "Flat tax reforms: Investment expensing and progressivity".
- 1102 *Rafael Repullo and Jesús Saurina*: "The countercyclical capital buffer of Basel III: A critical assessment".
- 1103 *Luis García-Álvarez and Richard Luger*: "Dynamic correlations, estimation risk, and portfolio management during the financial crisis".
- 1104 *Alicia Barroso and Gerard Llobet*: "Advertising and consumer awareness of new, differentiated products".
- 1105 *Anatoli Segura and Javier Suarez*: "Dynamic maturity transformation".
- 1106 *Samuel Bentolila, Juan J. Dolado and Juan F. Jimeno*: "Reforming an insider-outsider labor market: The Spanish experience".
- 1201 *Dante Amengual, Gabriele Fiorentini and Enrique Sentana*: "Sequential estimation of shape parameters in multivariate dynamic models".
- 1202 *Rafael Repullo and Javier Suarez*: "The procyclical effects of bank capital regulation".
- 1203 *Anne Layne-Farrar, Gerard Llobet and Jorge Padilla*: "Payments and participation: The incentives to join cooperative standard setting efforts".
- 1204 *Manuel Garcia-Santana and Roberto Ramos*: "Dissecting the size distribution of establishments across countries".
- 1205 *Rafael Repullo*: "Cyclical adjustment of capital requirements: A simple framework".

- 1206 *Enzo A. Cerletti and Josep Pijoan-Mas*: "Durable goods, borrowing constraints and consumption insurance".
- 1207 *Juan José Ganuza and Fernando Gomez*: "Optional law for firms and consumers: An economic analysis of opting into the Common European Sales Law".
- 1208 *Stéphane Bonhomme and Elena Manresa*: "Grouped patterns of heterogeneity in panel data".
- 1209 *Stéphane Bonhomme and Laura Hospido*: "The cycle of earnings inequality: Evidence from Spanish Social Security data".
- 1210 *Josep Pijoan-Mas and José-Víctor Ríos-Rull*: "Heterogeneity in expected longevity".
- 1211 *Gabriele Fiorentini and Enrique Sentana*: "Tests for serial dependence in static, non-Gaussian factor models".
- 1301 *Jorge De la Roca and Diego Puga*: "Learning by working in big cities".
- 1302 *Monica Martinez-Bravo*: "The role of local officials in new democracies: Evidence from Indonesia".
- 1303 *Max Bruche and Anatoli Segura*: "Debt maturity and the liquidity of secondary debt markets".
- 1304 *Laura Crespo, Borja López-Noval and Pedro Mira*: "Compulsory schooling, education and mental health: New evidence from SHARELIFE".
- 1305 *Lars Peter Hansen*: "Challenges in identifying and measuring systemic risk".
- 1306 *Gabriele Fiorentini and Enrique Sentana*: "Dynamic specification tests for dynamic factor models".
- 1307 *Diego Puga and Daniel Trefler*: "International trade and institutional change: Medieval Venice's response to globalization".
- 1308 *Gilles Duranton and Diego Puga*: "The growth of cities".
- 1309 *Roberto Ramos*: "Banning US foreign bribery: Do US firms win?".
- 1310 *Samuel Bentolila, Marcel Jansen, Gabriel Jiménez and Sonia Ruano*: "When credit dries up: Job losses in the Great Recession".
- 1401 *Felipe Carozzi and Luca Repetto*: "Sending the pork home: Birth town bias in transfers to Italian municipalities".
- 1402 *Anatoli Segura*: "Why did sponsor banks rescue their SIVs? A signaling model of rescues".
- 1403 *Rosario Crinò and Laura Ogliari*: "Financial frictions, product quality, and international trade".
- 1404 *Monica Martinez-Bravo*: "Educate to lead? The local political economy effects of school construction in Indonesia".
- 1405 *Pablo Lavado*: "The effect of a child on female work when family planning may fail".

- 1406 *Gabriele Fiorentini and Enrique Sentana*: "Neglected serial correlation tests in UCARIMA models".
- 1407 *Julio Galvez and Javier Mencía*: "Distributional linkages between European sovereign bond and bank asset returns".
- 1408 *Laurent Clerc, Alexis Derviz, Caterina Mendicino, Stéphane Moyen, Kalin Nikolov, Livio Stracca, Javier Suarez and Alexandros P. Vardoulakis*: "Capital regulation in a macroeconomic model with three layers of default".
- 1409 *Gerard Llobet and Jorge Padilla*: "The optimal scope of the royalty base in patent licensing".
- 1410 *Dante Amengual and Luca Repetto*: "Testing a large number of hypotheses in approximate factor models".
- 1411 *Gabriele Fiorentini, Alessandro Galesi and Enrique Sentana*: "A spectral EM algorithm for dynamic factor models".
- 1501 *Javier Mencía and Enrique Sentana*: "Volatility-related exchange traded assets: An econometric investigation".
- 1502 *Gabriele Fiorentini, Alessandro Galesi and Enrique Sentana*: "Fast ML estimation of dynamic bifactor models: An application to European inflation".
- 1503 *Samuel Bentolila, Ildefonso Mendez and Francisco Maeso*: "Leaving home with a partner".
- 1504 *Dante Amengual and Enrique Sentana*: "Is a normal copula the right copula?".